Signaling in Retrospect and the Informational Structure of Markets†

By MICHAEL SPENCE*

When I was a graduate student in economics at Harvard, I had the privilege of serving as rapporteur for a faculty seminar in the then-new Kennedy School of Government. Among other distinguished scholars, it included all of my thesis advisers—Kenneth Arrow, Thomas Schelling, and Richard Zeckhauser. In the course of that seminar there were discussions of statistical discrimination and many other subjects that relate to the incompleteness of information in markets. One of my advisers came in one day with the strong suggestion that I read a paper he had just read called "The Market for 'Lemons'" by George A. Akerlof (1970). I always did what my advisers told me to do and hence followed up immediately. It was quite electrifying. There we all found a wonderfully clear and plausible analysis of the performance characteristics of a market with incomplete and asymmetrically located information. That, combined with my puzzlement about several aspects of the discussion of the consequences of incomplete information in job markets, pretty much launched me on a search for things that I came to call signals, that would carry information persistently in equilibrium from sellers to buyers, or more generally from those with more to those with less information.[1] The issue, of course, was that signals are not terribly complicated things in games where the parties have the same incentives, i.e., where there is a commonly understood desire to communicate accurate information to each other. Even in that case (sometimes called the pure coordination case), however, there are potential problems of choosing among equilibria as illustrated in Thomas Schelling's (1960) brilliant analysis of the use of focal points and contextual information to solve communication/coordination problems when the parties have been deprived of the ability to communicate directly. In markets where the issue is often undetectable or imperfectly detectable quality differentials, the alignment of incentives is typically imperfect and the incentive of the high-quality product owners to distinguish themselves and the incentive of the low-quality owners to imitate the signal so as to obscure the distinction is fairly clear. Of course there is more to it, as one needs to know such things as who is in the market persistently and hence who has an incentive to establish a reputation through repeated plays of the game. I am going to devote a good portion of this lecture to these issues, in a sense to revisit signaling, and then turn to some other aspects of the informational structure of markets that are raised by the parameter shifts caused by the proliferation of the Internet as a communication medium in the past few years.[2]

† This paper is based on a lecture delivered in Stockholm, Sweden, on the occasion of the 100th anniversary of the founding of the Nobel Prize, when the author received the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel.

* Stanford Business School, Stanford University, 518 Memorial Way, Stanford, CA 94305. I thank my fellow recipients of the Nobel Prize in Economics this year, George Akerlof and Joseph Stiglitz, for their work and their inspiration, and my thesis advisers, Kenneth Arrow, Thomas Schelling, and Richard Zeckhauser whose ideas and guidance got me launched on the study of market structure (particularly informational structure) and performance. My colleagues Edward Lazear and Mark Wolfson gave me a great deal of constructive input. I also owe a great debt to James Rosse and Bruce Owen, with whom I learned and taught industrial organization and applied microeconomic theory at Stanford. It was a great group of young people and a wonderful time to be in that part of the field.

[1] I believe it was Robert Jervis who introduced the terms "indices" and "signals." Indices are attributes over which one has no control, like gender, race, etc. Think of them as unalterable attributes of something, not necessarily a person. Signals are things one does that are visible and that are in part designed to communicate. In a sense they are alterable attributes. I thought it was a useful set of distinctions and terminology and I still do.

[2] In economics and other social sciences, models abstract from reality and focus on structural features of organizations or markets that drive outcomes. There are embedded and usually unstated parameters that are not the focus of attention, as they do not normally change. When I refer to

I was asked recently by a somewhat incredulous questioner (actually, a journalist) whether it was true that you could be awarded the Nobel Prize in Economics for simply noticing that there are markets in which certain participants do not know certain things that others in the market do know. I thought it was pretty funny. It was as if this had somehow been a closely guarded secret up until about 1970, at least in economics. I clearly cannot speak for those who make the decisions about the Nobel Prize, but I suspect that the correct answer to that question is no. What did blossom at that time was a serious attempt by many talented economists to capture in applied microeconomic theory a whole variety of aspects of market structure and performance. That work produced a partial melding of theory, industrial organization, labor economics, finance, and other fields. An important early part of that effort was the attempt to capture *informational* aspects of market structure to study the ways in which markets adapt, and the consequences of informational gaps for market performance.

Therefore, in answer to the question, we noticed that there are many markets with informational gaps. These include most consumer durables, virtually all job markets, many financial markets, markets for various types of food and pharmaceuticals, and many more. These informational gaps were widely acknowledged and those of us who taught applied microeconomic theory freely admitted that these gaps might change some of the performance characteristics, not to mention the institutional structure, of markets in which they appear. But I think it is fair to say that we did not have much systematic knowledge based on theory of what those changes might be. And so we thought that applied microeconomic theory deserved an attempt to build these informational characteristics into models that capture the structure and performance of these markets with reasonably accurate assumptions about the *ex ante* informational conditions. It was a very exciting time for all of us who were involved. One of the wonderful aspects of receiving this award is that it has triggered in me, and I think others, fond memories of the sense of discovery and excitement. And in that context, I would like to take this opportunity to express my admiration and deep gratitude to the many extant young colleagues with whom I worked and shared these ideas. They should, and certainly in my mind do, share in the recognition triggered by the awarding of the distinguished Nobel Prize for what was accomplished during those years.

The plan for this lecture is as follows. The overriding goal is not to look too stupid to the next generation of students and scholars. More seriously, following the advice that my advisers once gave me, I will first discuss the simplest model that I can devise that illustrates in reasonably general form the definitions and properties of signaling equilibria.[3] Next, we will allow the signal (in this case education in the job market context) to contribute directly to the productivity of the individual as well as functioning as a signal. Following that, the paper examines a market in which there is signaling and both separating and pooling in the equilibrium.

The section after that examines a fairly general partial-equilibrium model of signaling and discusses competitive equilibria and certain kinds of "optimal" responses to signals.[4] Those

parameter shifts caused by the Internet, I mean changes in structural elements that formerly did not change. For example, most markets have geographic dimensions and boundaries that are implicitly understood and are not the focus of attention. The Internet however, has moved those boundaries by collapsing time and distance in the information/communication dimensions of markets.

[3] John von Neumann, who deserves to be on a short list of the greatest minds of the twentieth century, is reported to have said that you do not understand a theory or an abstract structure until you have seen and worked through hundreds of examples of it. Even if he did not say that (and I think he did), I agree with it. Few people can say that they had the idea that calculating machines did not have to be hardwired, that you could store instructions in memory and then execute them in order: that is, you could build a programmable computer.

[4] These optimal responses problems are selection problems. In insurance markets, they can be moral hazard or adverse selection problems, and in other contexts they are agency problems or optimal taxation problems. You want to confront people with schedules that cause them to make appropriate choices and in so doing to reveal themselves and the information that they privately hold *ex ante*. Generally the results are below what is achievable with perfect information, because with perfect information, you do not have to worry about the revelation problem. These problems have the economic and mathematical properties of second-best taxation problems, of which perhaps the first to be analyzed was the "optimal" income tax problem, by the Reverend Frank Plumpton Ramsey (1927), though the anal-

who are more interested in the general idea and less in the general case can skip over this section without risk of missing any important properties. However this is the section in which the formal relationship between signaling models and models of self-selection and optimal taxation with imperfect information is examined, and that may be of interest to some readers. I also want to introduce in that context a new possibility that I discovered only recently and that is that one can have a signaling equilibrium in which the costs of the signal appear to vary with the unseen ability characteristics in the wrong way. That is to say the costs of education (absolutely and at the margin) rise with ability, or more generally, with the unobserved attribute that contributes positively to productivity. For the most part, I will use the job market case for expositional purposes and comment on some other signaling situations more briefly later in the essay. There is a risk in using job markets to illustrate signaling. In order to illustrate the fundamental properties of signaling models, I have stripped away other features of the market, particularly in the simpler models. But I have noticed in the past that there is a tendency for the simpler models set in the job market context to convey unintended messages such as (1) education does not contribute to productivity, or (2) the information contained in the signal does not increase efficiency. This essay is mainly about the theory of signaling and information transfer in markets, and not mainly about human capital and labor markets. There are many who are more knowledgeable about the latter than I.

ysis seemed to have largely escaped notice until Anthony Atkinson (1971), James A. Mirrlees (1971), and then others rediscovered it. The idea is that the taxing authority cannot directly observe an individual's earning power (think of it as maximum potential hourly wage) so it taxes what it can see and monitor, which is income. This produces a disincentive with respect to work that could have been avoided if the tax were based on earning potential rather than actual earnings. The second-best optimizing calculation then trades off in the most efficient way disincentives for work and the need to meet a revenue goal for the taxing authority. The mathematics of these problems are sometimes referred to as optimal control problems and sometimes the calculus of variations. In later work, Bengt Holmström (1979) applies these models to a labor setting, incorporating Robert Wilson's (1968) work on syndicates (optimal risk sharing). Here the second-best optimizing calculation sacrifices optimal risk-sharing employees and owners to provide incentives for employees to work diligently.

The paper concludes with two sections. The first of these discusses the ubiquitous use of time and the allocation of time as a signal and as a screening device. The second focuses on the potentially rather large information parameter shifts that the Internet may have caused and speculates a little on how applied models of markets, organizations, and boundaries between them may have changed and on what is needed to capture the effects of these parameter shifts in models of markets and the economy.

I. The Simplest Job Market Signaling Model

The idea behind the job market signaling model is that there are attributes of potential employees that the employer cannot observe and that affect the individual's subsequent productivity and, hence, value to the employer on the job. Let us suppose that there are just two groups of people. Group 1 has productivity or value to any employer of 1, and group 2 has productivity of 2. In this example, these productivity values do not depend on the level of investment in the signal. If there were no way to distinguish between people in these two groups then if both groups stay in the market, the average wage would be $2 - \alpha$, where α is the fraction of the population in group 1, and everyone would get that wage. If the higher productivity group through dissatisfaction or for any other reason exits this labor market, the average productivity and the wage drop to 1. This phenomenon when it occurs is sometimes called the adverse selection problem, a label most commonly applied to insurance markets. It is structurally the same problem that Akerlof (1970) described in his famous paper on used cars (lemons).

Now suppose that there is something called education, which we will denote by E, that can be acquired or invested in. It is assumed to be visible, and its acquisition costs differ for the two types. Let us suppose that the cost of E years of education for group 1 is E, and the cost for group 2 individuals is $E/2$. For this example I am going to assume that education does not affect the individual's productivity. I do this purely to keep it simple and not to suggest that human capital, including that acquired through education, is somehow not relevant. In later sections, the assumption will be relaxed.

Equilibrium in a situation like this, and in general, has two components. First, given the returns to and the costs of investing in education, individuals make rational investment choices with respect to education. Second, employers have beliefs about the relation between the signal and the individual's underlying productivity. These beliefs are based on incoming data from the marketplace. In equilibrium, the beliefs must be consistent, that is, they *must not be disconfirmed* by the incoming data and the subsequent experience. Therefore, one could say that the beliefs must be *accurate*. But one should also notice that employers' beliefs/expectations determine the wage offers that are made at various levels of education. These wage offers in turn determine the returns to individuals from investments in education, and finally, those returns determine the investment decisions that individuals make with respect to education, and hence the actual relationship between productivity and education that is observed by employers in the marketplace. This is a complete circle. Therefore it is probably more accurate to a say that in equilibrium, the employers' beliefs are *self-confirming*. This may sound like a minor restatement but it is important. It is the self-confirming nature of the beliefs that gives rise to the potential presence of *multiple equilibria* in the market.

In the example at hand, suppose that group 1 individuals set $E_1 = 0$ and individuals in group 2 set $E_2 = E^*$. Suppose further that employers, none of whom individually influence investment decisions by individuals,[5] believe that if $E < E^*$ then productivity is 1 and if $E \geq E^*$, then productivity is 2. With these assumptions, individuals in group 1 will rationally set $E = 0$ provided that

$$2 - E^* < 1.$$

Members of group 2 will rationally set $E = E^*$ provided that

$$2 - \frac{E^*}{2} > 1.$$

[5] Investment decisions are made by individuals in advance of knowing with whom they will work and employers are such that none of them individually influences materially the perceived offers in the market place.

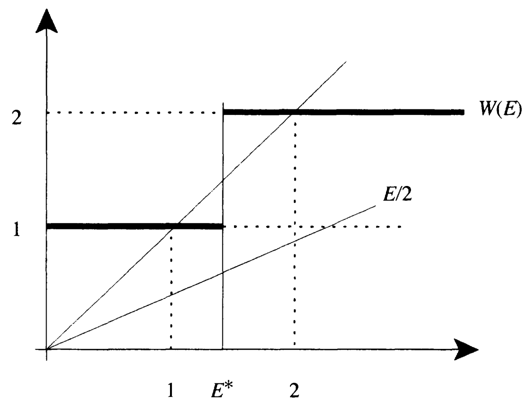

FIGURE 1. THE TWO-GROUP SIGNALING EQUILIBRIUM

Therefore the choices will be rational and the expectations confirmed in the market if

$$1 < E^* < 2.$$

Though it is a highly stylized example it has many of the general properties of signaling equilibria. There is a continuum of equilibria, in each of which there is more investment in the signal than there would be in a world of full information. Because investment in the signal dissipates resources without improving productivity, the result is inefficient. Moreover, the equilibria are orderable by the Pareto criterion, that is, as you move from one equilibrium to another, everyone is either worse off or no better off. The signal actually does distinguish low- and high-productivity people and the reason it is able to do so is that the cost of the signal is negatively correlated with the unseen characteristic that is valuable to employers, in this case productivity itself. The equilibrium is characterized by a schedule that gives the returns to education, that is a wage for each level of education, and optimizing choices given that schedule by individuals in the two groups.

The equilibrium is depicted in Figure 1. The wage schedule is the dark line and jumps from a wage level of 1 to a wage level of 2 at e^*. E^* is between 1 and 2 so that group 1 chooses the education level 0 and group 2 chooses E^*. It is not essential that net income for group 1 at e^* is negative as it is in this example. We could make the productivity levels for the two groups 3 and 4. The same signaling result would occur with suitable adjustments in the level of the wage schedule.

There are in fact other signaling equilibria. For example, there could be a minimal level of education that group 1 invests in. But unless it is a productive investment, there is no reason to think that over time the market would not discover this and eliminate it as an inefficiency that can be removed costlessly. In the spectrum of equilibria described above, probably the most interesting is the most efficient one. That equilibrium is the one in which $E^* = 1 + \delta$, where δ is to be thought of as a small positive number. In this equilibrium,

$$w_1 = N_1 = 1$$

$$w_2 = 2 \qquad N_2 = 1.5 - \delta/2$$

where N_i is income net of signaling costs for group i.

On the assumption that the market will find equilibria that are Pareto efficient, one can ask whether there is a better equilibrium in the Pareto sense (everyone is better off) than the one described above. The answer is sometimes yes. It involves pooling and it depends on the relative sizes of the two groups. What is really going on here is that the market is, in a sense, acting as if it is trying to maximize the net income of the higher productivity group. In some cases that is accomplished by recognizing that it is too expensive for that group to distinguish itself. The alternative is to appear undifferentiated in a pool. That clearly benefits group 1 as they are mistaken for the average productivity, which is above 1. As a reminder, α is the fraction of the population in group 1. In a pooling situation the average productivity is $2 - \alpha > 1$. Group 1 would clearly prefer it. Group 2 prefers it as well if

$$2 - \alpha > 1.5 - \delta/2, \text{ or}$$

$$\alpha < 0.5(1 - \delta).$$

Since δ can be made as small as we want, there is a preferred pooling equilibrium (in the Pareto sense) if in this example, group 1 is less than half the population. In general with discrete groups, pooling with lower-level groups becomes attractive if they are relatively small, because the higher-level groups do not give up too much by being mistaken for the average and they avoid the signaling costs. In a later section I will describe briefly a case in which one can observe pooling and separating in the same equilibrium.

If you wanted to make the result in this market more efficient than the equilibria described above, you would tax education (assuming that such a tax is not costly to impose and administer), making it more expensive for group 2, and thereby make it possible to lower the level of education without losing the informational content of the signal. The revenues generated could be distributed equally to all participants in both groups independent of their education choice. The lump-sum component of the tax will therefore not affect their signaling behavior.

Let t be the tax rate on investment in education and let k be the lump-sum distribution from the tax revenues. This distribution goes to everyone. Let the signaling level of education be E^*. Group 1 individuals will rationally choose not to send the signal if

$$2 - (1 + t)E^* + k < 1 + k$$

and group 2 will rationally send the signal if

$$2 - (0.5 + t)E^* + k > 1 + k.$$

We have an equilibrium then if

$$\frac{1}{0.5 + t} > E^* > \frac{1}{1 + t}.$$

We will pick the efficient end of the signaling spectrum by setting

$$E^* = \frac{1 + \delta}{1 + t}.$$

The lump-sum distribution is equal to the tax revenues so that

$$k = \frac{t(1 + \delta)(1 - \alpha)}{1 + t}.$$

Thus the equilibrium net incomes are

$$N_1 = 1 + k$$

$$N_2 = 2 - \frac{(0.5 + t)(1 + \delta)}{1 + t} + k.$$

As δ becomes small and t becomes very large,

k approaches $(1 - \alpha)$, and hence both net incomes approach $(2 - \alpha)$. This, the reader will recall, is the pooling equilibrium outcome in terms of net income. Here however, there is no upper limit on the size of group 1, namely α. Signaling costs (socially) are negligible as E^* approaches zero, though the private marginal costs of signaling (including the tax) approach 1.

To summarize, one can get rid of the inefficiency and keep the signaling and the informational content of the signals with an appropriate tax on the signaling activity. The effect is to redistribute income in such a way as to reproduce the results of the pooling equilibrium. It could be that the information the signal carries is itself productive. In that case it is important that the equilibrium actually be a separating one.[6] That is the case here. Separating is retained in this equilibrium. Groups are correctly identified. The use of a tax on the signaling activity reduces the level of the signal required to distinguish the groups, and hence reduces the inefficiency of the untaxed signaling equilibrium.

II. The Two-Group Model When Education Enhances Human Capital and also Serves as a Potential Signal

It will naturally occur to the reader to ask what is the effect on the market signaling that we have just examined, at least in the context of labor markets, when the signal also contributes directly to productivity for the individual worker.[7] This section addresses the subject by modifying the two-group model of the preceding section to allow education to be productive. I will make the model slightly more general in terms of functional forms. Let $s_i(E)$ be the value of a worker of type I with education E to an employer. Here $I = 1, 2$ and we assume that $s_2(E) > s_1(E)$, and $s_2^{\prime}(E) > s_1^{\prime}(E)$. Let $c_i(E)$ be the cost of investing in E units of education for group 1 and assume that $c_1(e) > c_2(e)$, and

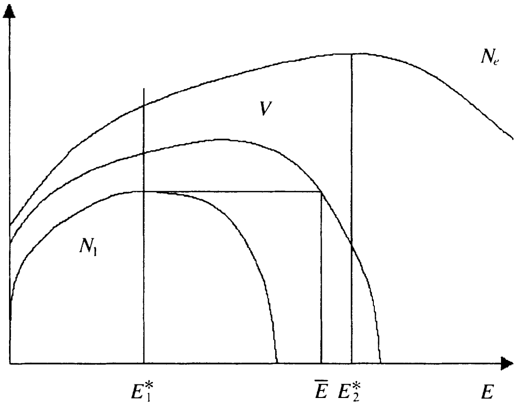

FIGURE 2. TWO-GROUP SIGNALING WITH EDUCATION AFFECTING PRODUCTIVITY

that $c_1^{\prime}(E) > c_2^{\prime}(E)$. These assumptions simply capture the idea that the group with the higher productivity also has the lower costs of signaling. As before, the individual's type is not directly observable. We will also assume that $s_i(E)$ is concave, that $c_i(E)$ is convex and hence that the net income function $N_i(E) = s_i(E) - c_i(E)$ is concave.

There are three qualitatively different kinds of equilibria in this market. The first is a fully efficient separating equilibrium. It is easiest to see the types of equilibrium pictorially. In Figure 2, we have various forms of net income plotted as a function of E. Let $V_1(E) = s_2(E) - c_1(E)$. Our interest in the function V clearly relates to the question of whether group 1 will adopt the group 2 signal and hence create a problem with the separating equilibrium.

In Figure 2, you will see the functions N_1, N_2, and V. The points E_1^* and E_2^* are the points that maximize N_1 and N_2. The point $\bar{E}$ is the largest value of E such that V is larger than N_1^*, the maximum value of N_1. In Figure 1, you can see that E_2^* is to the right of $\bar{E}$. This means that as long as the wage offer schedule jumps up from $s_1(E)$ to $s_2(E)$ to the right of $\bar{E}$, individuals in group 1 will have no incentive to imitate the signaling behavior of group 2.

This is the fully efficient separating equilibrium. The signal carries accurate information, and the investments in education are the efficient ones. The outcome is as if there was perfect information in the market place. In essence, the two groups are sufficiently different

[6] Information is directly productive, for example, if there is a job allocation or training decision that is made more efficiently with accurate knowledge of the type. In that case the pooling equilibrium would be inefficient.

[7] Actually this issue does not just arise in labor markets. Generally signals can change the value of the product. For example, in the used car case, a warrantee changes the value of the bundle, the car, and the warrantee in addition to carrying information about the product itself.

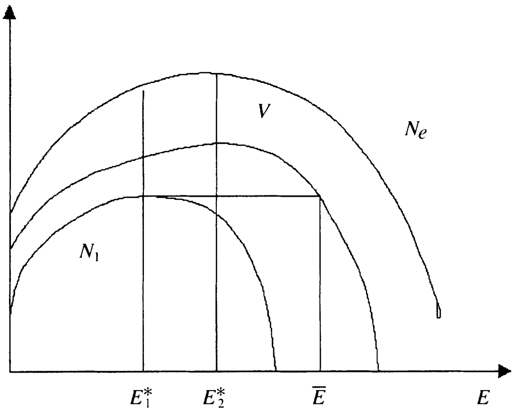

FIGURE 3. TWO-GROUP MODEL WITH OVERINVESTMENT IN EDUCATION

in terms of some combination of the productivity of education as human capital and the costs of education that the fully efficient outcome survives as an equilibrium. Note however, that the negative relation between productivity and costs really matters. If signaling costs were the same, then the functions N_2 and V would be the same and we could not have this market outcome. When signaling costs are the same one can never prevent group 1 from imitating group 2 behavior when it is to their advantage to do so.

Now turn to Figure 3. It is essentially the same picture as Figure 2 except that E_2^* is to the left of $\bar{E}$. This means that if the wage schedule jumps to $s_2(E)$ before E_2^*, group 1 will imitate group 2 behavior and dismantle the separating equilibrium. To prevent this we will specify that the wage schedule jumps up to $s_2(E)$ at $\bar{E} + \delta$. Subject to a qualification that I will come to in a minute, this is a separating equilibrium. To achieve the separation, the group 2 investment in education is pushed up above the efficient level, in order for the individuals in that group to achieve the signaling effect and avoid imitation by group 1. So here we have the same kind of result that we had in the no-human-capital case. Here the return to investing in education has a signaling component and a human-capital component. The former can push the levels of investment above the full information optimum. In the first model the human-capital effect was zero, leaving only the signaling effect, an assumption that ensures that overinvestment will occur in any separating equilibrium in which signaling occurs.

There remain two issues. One is whether and under what circumstances pooling will destroy the hypothesized separating equilibrium, and the second is: "Can one tax the signal in such a way as to improve market performance?" Pooling first. Let α be the fraction of the total population in group 1. In a pooling equilibrium, the wage for all will be $\alpha s_1(E) + (1 - \alpha)s_2(E)$ for any level of E that emerges. This cannot break a fully efficient separating equilibrium because group 2 cannot be better off. Thus if pooling breaks a separating equilibrium it can only be in the case where the more productive group has been forced, for signaling purposes, above its optimal level of E from the standpoint of investment in human capital. Let $W(E) = \alpha s_1(E) + (1 - \alpha)s_2(E)$. By making α arbitrarily small we can make W as close to $s_2(E)$ as we want. Now let us look again at Figure 2. If you start at $\bar{E}$ and reduce E and move to the left and up the curves N_2 and V, it is clear that both groups are better off. For small α, the net income functions $W - c_1$ and $W - c_2$ are close to V and N_2. Hence for a range of α at the low end, pooling will break the separating outcome. For these cases, the natural choices of equilibrium levels of E are those that lie between the education levels that maximize $W - c_1$ and $W - c_2$.

As α rises, eventually the function $W - c_2$ will fall below the level of group 2 net income at the separating point $\bar{E}$, at which point the pooling outcome is not capable of breaking the separating equilibrium. To summarize, in the case of an inefficient separating equilibrium, there exist pooling equilibria that are Pareto superior to the separating equilibrium provided the size of the lower productivity group is below some threshold level. If the market has a way of finding the pooling equilibria, it will break the separating one.

We turn now to the problem of improving the efficiency of the market. We saw in the preceding section that in the case where there is no human-capital component of productivity, one can in principle come arbitrarily close to the efficient outcome through appropriate taxes. The same is true when the signal adds to the individual's productivity as well as functioning as a signal. We have already seen that in some cases the market equilibrium itself is efficient in the sense of maximizing total net income as part of a separating equilibrium. We now want to show that one can achieve the same result in the case where the higher productivity group has been forced by signaling considerations to move above its efficient level of investment in human capital.

To achieve an efficient outcome, we want to maximize total net income. To do that we need to have $s_i^t = c_i^t$ for $i = 1, 2$. To accomplish that if it is possible, we need a wage function that induces these choices. The wage function can be thought of as a schedule of taxes on education superimposed on a wage function that sets wages equal to productivity. Since each individual in each group considers all possible levels of E in making its education investment decision, it turns out to be easier to solve this problem indirectly rather than directly. Let us suppose therefore that counter to fact, there is a continuum of types specified by the parameter z, where productivity for type z is $zs_1 + (1 - z)s_2$ and costs for type z are $zc_1 + (1 - z)c_2$. Let $E(z)$ maximize

$$z(s_1 - c_1) + (1 - z)(s_2 - c_2).$$

Note that $E(z)$ declines as a function of z. Invert $E(z)$ to get $Z(E)$. If the wage schedule we are looking for is $w(e)$, then individuals will have maximized by setting

$$w^t = zc_1^t + (1 - z)c_2^t.$$

Substitute $Z(E)$ into the differential equation above and integrate with respect to E, setting the constant so that total wages equal total output. This wage schedule and/or the implied tax schedule will induce the efficient choices of E.[8]

The reader will notice that the optimal wage schedule above does not depend on the distribution of z in the population.[9] Therefore let us assume that almost all the weight in the distribution is at either $z = 0$ or $z = 1$. These are the two group cases we are studying.

To summarize, with two groups and education as productive human capital, one can have a signaling equilibrium with full efficiency or overinvestment in education by the more productive group. You can also have a pooling equilibrium that dominates the separating equilibrium provided that the less productive group is not too large. And finally there exists a tax/subsidy scheme that produces a fully efficient separating outcome as an equilibrium.

III. A Model with Signaling, Selection, and Pooling[10]

The model we are going to look at briefly in this section is of interest for two reasons. First, it illustrates a case in which there is both pooling and separating components of the equilibrium and, second, it shows that the critical criterion for having a separating component in an equilibrium is that the net benefits of issuing the signal are positively correlated with an unseen attribute that contributes positively to productivity. This positive correlation can result, as in preceding examples, from signaling costs that are negatively correlated with the valued attribute. That is a sufficient but not a necessary condition for a separating equilibrium. In this very nice example the information is acquired at a fixed cost and the positive correlation comes from subsequent discovery of the attribute post-employment.

The idea behind the model is that the value of individuals to firms is not directly observed, at least at the time of hiring. The value which we will denote by q is distributed on the interval $[q_{\min}, q_{\max}]$. Let the distribution of q in the population be $f(q)$. Individuals have a choice. They can work for firms that do not distinguish among them and hence pay everyone the same amount. Or they can work for firms that incur an expense, denoted by e, as a result of which the firm eventually learns the value of q for the individual and pays the person accordingly. In equilibrium this cost, which is the same for everyone, is passed on to the individual in the form of a reduction in compensation. Labor markets are assumed to be competitive. For

[8] Note that $w^{tt} - [zc_1^{tt} + (1 - z)c_2] = z^t(E)[c_1^t - c_2^t] < 0$, so that the second-order condition for a maximum is satisfied. The optimal tax schedule will actually consist of a rising tax on education at all levels of education combined with a lump-sum distribution to everyone so that net tax revenues are zero. In this example, group 1 will receive a net subsidy and group 2 will pay a net tax.

[9] This is a general property of the continuous case, when the objective is, as it is here, to maximize net income by inducing the efficient levels of investment in human capital.

[10] This example was developed by Edward P. Lazear (1998). It illustrates nicely the coexistence of pooling and separating in equilibrium and also a different structural condition that permits the signaling and selection to occur.

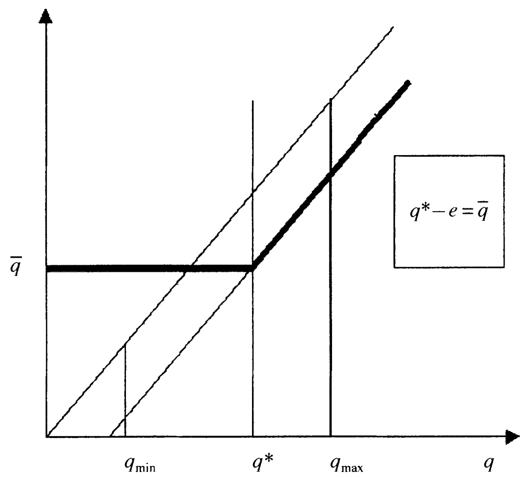

FIGURE 4. EQUILIBRIUM IN THE SEPARATING AND POOLING MODEL

those individuals who choose the firms where they are distinguished, the wage is $q - e$. If the individual chooses a firm that does not incur the cost and does not distinguish among workers, then the compensation is the average value of the individuals who work for that firm.

Let us suppose that the average value of the workers in the pooling firms is $\bar{q}$. If we consider the optimizing decisions of individuals, it is clear that if

$$q - e > \bar{q}$$

then the individual will choose to work for the separating firm and conversely. Let $q^* = \bar{q} + e$. In equilibrium, $\bar{q}$ is the actual average value of those who work in the pooling firms, so that

$$q^* - e = \frac{\int_{q\,\min}^{q^*} qf(q)\, dq}{F(q^*)} = E(q|q \leq q^*)$$

where $F(q)$ is the cumulative distribution function for the attribute q. The equilibrium is illustrated in Figure 4.

Those with $q \leq q^*$ receive $\bar{q}$ and are pooled, while those with $q > q^*$ receive $q - e$ and are identified and are in the separating part of the equilibrium. It is possible that everyone will be pooled. This occurs when e, the cost of discovery, is large enough that $q_{\max} - e$ is less than the unconditional mean of the whole distribution. Another view of the equilibrium is provided in Figure 5.

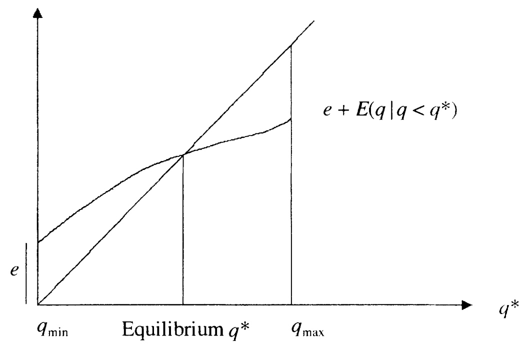

FIGURE 5. AN ALTERNATIVE VIEW OF THE EQUILIBRIUM

Here we plot the functions q^*, and $E(q|q \leq q^*) + e$, as a function of q^*. The second function is upward sloping and will cross the 45-degree line, unless e is large enough to keep the second function above the 45-degree line for all q^*. The crossover point is the equilibrium level of q^*.[11]

The signal is the choice of which type of firm to go to and it is partially, but not completely, informative about the individual's value to the firm. I noted in the first model that the signal retains its informational content in equilibrium if the cost of the signal is negatively correlated with the (hard-to-observe) valued attribute. In this case, the cost of the signal is a constant across people but because of the subsequent discovery and adjustment, the net benefit to the individual of issuing the signal is positively correlated with the valued attribute.

It is possible that the information carried by the signal increases efficiency. That circumstance increases the return to going into the separating side of the market and has the overall effect of increasing the size of the group that sends the signal. I should also note that it has been implicitly assumed here that the information about productivity, once it is acquired, is

[11] While the conditional expected value function is upward sloping, its slope is not necessarily less than 1, and depends on $f(q)$. It is therefore in principle possible to have an odd number of multiple crossovers. If there are three, the middle one will be unstable and the ones at the outside will cross over with a slope of less than 1 and be stable. To see this, note that if at some point of time, $q^* - e > \bar{q}$, then q^* will rise, and conversely.

public. That would force the employer to pay the employee his or her productivity minus the discovery cost. If, on the other hand, the discovery is private, there will be a negotiation between the employer and the employee over the net income, and that will have the effect of lowering the net income of the employee relative to the case in which the information is public. Thus one should expect the size of the pooling component of the market equilibrium to rise when the discovery is private.

IV. The General Continuous Model[12]

I would now like to establish some reasonably general properties of signaling models.[13] Following that, we will look at improving the equilibrium performance of the market through taxes and subsidies, using the approach of optimal taxation with imperfect information. The variables are n, standing for some attribute that is (a) not directly observable, and (b) valuable to employers, and y, which is years of education. The latter is observable and may be valuable to employers. The functions with need are $S(n, y)$ which specifies an individual's productivity or value to an employer as a function of education and n, which I will henceforth refer to as ability. The remaining functions are $c(y, n)$, which specifies the cost of education as a function of the same two variables and $w(y)$, which is the wage offered to an individual who presents him or herself in the market with education of y. Since this is quite familiar territory, I will proceed fairly quickly. The equilibrium is defined by two conditions. First, given $w(y)$, individuals maximize income net of education costs with respect to y. Thus they maximize $w(y) - c(y, n)$ by setting

$$w'(y) = c_y(y, n).$$

This holds for all n. The second-order condition, $w''(y) - c_{yy}(y, n) < 0$, must hold. The second condition is that employers' experience in the market over time must be consistent with the offers they are making, so that for all n,

$$w(y) = s(n, y).$$

Ignore the second-order conditions for the moment. Since by assumption $s_n > 0$, one can in principle solve the equation above for n in terms of w and y, say, $n = N(w, y)$. Substituting in the first-order conditions, we have

$$w'(y) = c_y(y, N(w, y)). \qquad (1)$$

This is a first-order, ordinary differential equation. It has a one-parameter family of solutions that do not cross each other. In principle each member of this one-parameter family can be part of a market-signaling equilibrium.[14]

If $c(0, n) = 0$ for all n and if $c_{yn} < 0$, then assuming $c_y > 0$, $c(y, n)$ will be declining in n and that combined with an upward-sloping $w(y)$ (without which no one would invest in education), ensures that y is an increasing or at least a nondecreasing function of n. Let the net income function be

$$N(y, n) = w(y) - c(y, n).$$

This function has the property that N_y is an increasing function of n, which ensures that if $N_y = 0$ for a particular value of n, then if you raise n, N_y will be greater than zero and the maximum will occur at a higher value of y.

Differentiating (1) with respect to y we have

$$w'' - c_{yy} = c_{yn}\frac{dn}{dy} < 0$$

which means that the second-order condition is in fact satisfied.

Without being too formal about it, education

[12] This section is somewhat technical. It is not difficult mathematics but does require a general knowledge of differential equations and the calculus of variations. It is not essential for those interested in learning the basic properties of signaling equilibria and can be skipped by readers who are willing to accept that the properties cited in the earlier examples survive in the general case. The second-best optimizing calculations are of some interest in interpreting what is going on in the market.

[13] Much of the material in this section is taken from parts of Spence (1974a).

[14] The solutions cannot cross because if they did, then the derivative at a given point in (w, y) space would have two values, which contradicts the statement that w' has a well-defined value $F(w, y)$ at every point. As a consequence, let us assume that the family of solutions is $w(y, K)$ where K is a parameter. If $w_K > 0$ anywhere, then it is true everywhere, since to assume the opposite would mean that the solutions cross.

or any potential signal transmits information in equilibrium if its costs absolutely and at the margin decline as the unseen valued attribute increases. This is a sufficient condition. In a later section I will show by example that it is not a necessary condition.

Since for every level of n, in equilibrium $w(y) \equiv s(n, y)$, differentiating we have

$$w'(y) = s_y + s_n \frac{dn}{dy} > s_y$$

because $s_n > 0$ and $dn/dy > 0$. This implies that in equilibrium, the private return to education is higher than its direct contribution to productivity, because of the second term in the equation above. This second term is the signaling effect. It is the part of the private return to the investment in education that is tied to the unobserved level of n. Since $w' = c_y$, the above inequality implies that $s_y - c_y < 0$, so that in equilibrium for all n, the investment in education is higher than it would be with perfect information. If n were observable, then individuals would be paid $s(n, y)$, and they would select education to maximize $s - c$ by setting $s_y = c_y$. Note that if $s_n \equiv 0$ for all n, then the signaling effect goes away. The attribute n is still unobservable to employers, but individuals make efficient education investment decisions as they have the information they need to make the investment decision optimally, and they are the only ones who need that information.

As noted earlier, there is a one-parameter family of equilibrium wage schedules that do not cross. Let the parameter be k and denote the schedules by $w(y, k)$. As the solutions do not cross, without loss of generality we can assume that $w_k > 0$. Net income to the individual is $N = w - c$. Differentiating partially with respect to k, with n held constant, we have

$$(2) \quad N_k = w_k + (w_y - c_y)(\partial y/\partial k) = w_k > 0.$$

Thus a shift from one equilibrium to another makes everyone better or worse off together. The equilibria are orderable by the Pareto criterion. What happens as one moves from one equilibrium to another is that the extent of overinvestment in the signal is increasing or declining for everyone. I have to confess that at the time, even after discovering that there might be multiple equilibria, I did not expect that there would be such a simple relationship among them in terms of market performance.

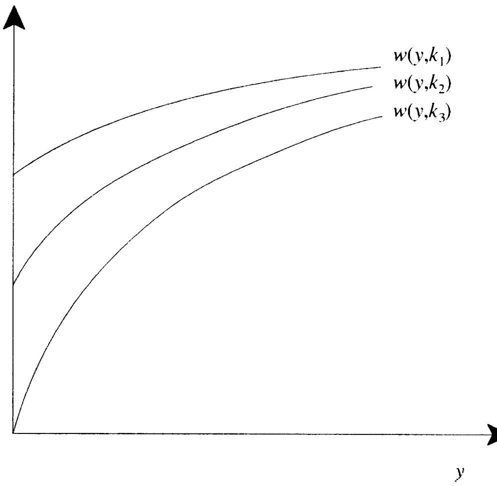

FIGURE 6. A FAMILY OF EQUILIBRIUM WAGE SCHEDULES

We can use (2) above and the fact that $w \equiv s$ to derive the effect of a shift in the equilibrium on the levels of investment in education. $N = w - c = s - c$. Differentiating with respect to k, holding n constant, and using (2) we have

$$\frac{\partial y}{\partial k} = \frac{w_k}{(s_y - c_y)} < 0.$$

Finally, from the equilibrium condition $w' = c_y$, differentiating with respect to k with n fixed, we have

$$w_{yk} = -\left[\frac{\partial y}{\partial k}\right](w_{yy} - c_{yy}) < 0.$$

This gives us a pretty complete picture of what happens when the equilibrium wage schedule shifts up. It becomes flatter, inducing lower levels of investment in education. Average wages decline and average education costs decline more, so that net incomes rise and everyone is better off. Figure 6 represents the family of equilibrium wage schedules. As they rise, they become flatter and the range of investment in education shifts to the left. Figure 6 shows the range of investment in education for each of the schedules. As the schedules rise and become flatter, the ranges of investment in education move to the left, that is, they fall.

While it is perhaps interesting to note that the signaling effect causes overinvestment in the signal by the standard of a world in which there is perfect information, that is not the world we live in. Thus it may be more interesting to examine the equilibrium outcomes in comparison to various second-best outcomes that acknowledge that there is an informational gap and asymmetry that cannot simply be removed by a wave of the pen.

I will begin by looking at the maximization of total net income. Here as we shall see, one can achieve the efficient levels of investment in education for all n. The required tax in effect undoes the signaling effect. The reason that this is a clean case is because a dollar of net income is of equal value regardless of the recipient; we do not get into the business of trading off income distribution against the objective of efficiency. Following this brief analysis, we will look at social welfare functions that are not linear in net income. For these, there will be a trade-off between efficiency and distribution. I will show that the market outcome is more akin to the outcome of maximizing a convex (that is to say antiegalitarian) social welfare function. This is not a surprising result since the function of signaling is to distinguish low- from high-productivity individuals, which results in a move away from egalitarian outcomes. I do also want to emphasize at the outset that the purpose of this analysis is not to propose or promote taxing or subsidizing signals but rather to help shed additional light on the properties of the competitive equilibrium.

V. Maximizing Total Net Income

We will assume that the unobservable attribute n is distributed in the population according to the density function $f(n)$, and the cumulative distribution will be represented by $F(n)$. The total net income then is

$$\int_{\underline{n}}^{\bar{n}} N(y, n) f(n)\, dn. \tag{3}$$

Here $N(y, n) = w(y) - c(y, n)$ as before. The constraints are two. Individuals choose rationally so that $w^t = c_y$ and total gross wages have to equal total productivity or value generated to employers:

$$\int_{\underline{n}}^{\bar{n}} w(y) f(n)\, dn = \int_{\underline{n}}^{\bar{n}} s(y, n) f(n)\, dn. \tag{4}$$

Our goal here is to select a schedule $w(y)$ that causes education choices $y(n)$ for each n, that maximize total net income, net, that is, of education costs. This is actually pretty straightforward. Using (4), we can replace $w(y)$ in (3) and choose $y(n)$ to maximize

$$\int_{\underline{n}}^{\bar{n}} (s(y, n) - c(y, n)) f(n)\, dn.$$

This is the simplest form of a calculus of variations problem.[15] The solution is a schedule $y(n)$ that causes $s_y = c_y$, for all levels of n. The reader will note that the implied levels of investment in education are those which would occur if n were observable, that is, in the hypothetical world of perfect information. They are therefore also the efficient levels of investment in education for each n: education is invested in up to the point where at the margin its direct contribution to productivity (the human-capital effect) is equal to the marginal cost. That is to say, to maximize net income, the wage schedule is set so as to remove the signaling effect, though education still carries the information and acts as a signaling. The private part of the return to the signal (which has to do with distinguishing one individual from another and hence falls into the zero-sum/redistribution aspect of the market mechanism) is simply removed by the optimal tax or wage schedule. To find the required wage schedule, you invert the optimal schedule $y(n)$ to $r(y)$, substitute that in the optimizing condition $w^t(y) = c_y(y, r(y))$, and integrate to get

$$w(y) = w(0) + \int_0^y c_y(v, r(v))\, dv. \tag{5}$$

The parameter $w(0)$ is then set to cause the breakeven condition (4) to be met. Individuals

[15] A good treatment of the calculus of variations can be found in Courant. You displace y by $\delta\theta(n)$, differentiate with respect to δ and insist that the result be equal to zero for all $\theta(n)$. This gives us $\int_{\underline{n}}^{\bar{n}} (s_y - c_y)\theta(n) f(n)\, dn = 0$. That in turn implies that $s_y = c_y$ for all n.

are likely being paid their productivity in the marketplace, $s(y, r(y))$. Therefore one can get to the desired result by imposing a tax/subsidy on education of $t(y) = s(y, r(y)) - w(y)$, where $w(y)$ is defined by (5) above. If you differentiate $t(y)$ with respect to y you have

$$\begin{aligned} t'(y) &= s_y + s_n r' - w' \\ &= s_y - c_y + s_n r' = s_n r'. \end{aligned}$$

The last term in the equation is the signaling effect at the margin of a change in education. This then says that the optimal tax at the margin is equal to the signaling effect. It takes that part of the private return to education away, leaving only the human-capital effect or the direct contribution to productivity.

One can also achieve the same result with an income tax. Let income be I. Its relation to education is given by $I = s(y, r(y))$. Invert that function to get $y = Y(I)$ and then set the income tax $T(I)$ equal to $I - w(Y(I))$, where again $w(y)$ is defined by (5). Individuals will select income so as to maximize $I - T(I) - c(Y(I), n) = w(Y(I) - c(Y(I), n)$, by setting $w' = c_y$ which is the desired result.[16]

It is interesting for those of us who live in the United States that to a first crude approximation, our education costs are roughly similar in form to the efficiency-inducing tax schedule. Lower levels of education are subsidized and these subsidies decline at college and university levels because of the existence of a large private sector in higher education.

VI. Second-Best Optima with Asymmetric Information

We turn now to a brief examination of the effects of investment in education and on net incomes of maximizing some social welfare function. As before, $N = w(y) - c(y, n)$ is net income and n is distributed in the population according to $f(n)$. Let $V(N)$ be the social value of net income of N to any individual. For the moment, we will place no restrictions on $V(N)$ other than that it is upward sloping—more net income is better. We will adopt an additive social welfare function:

$$Z = \int_{\underline{n}}^{\bar{n}} V(N(n))f(n)\, dn.$$

For the moment, we will make no assumptions about the shape of $V(N)$ except that it is upward sloping. The objective is to maximize Z subject to two constraints. One is that individuals make a rational choice of education

$$w'(y) = c_y(y, n).$$

The second constraint is that gross wages add up to total output or productivity:

$$\int_{\underline{n}}^{\bar{n}} (w - s)f(n)\, dn = 0. \tag{6}$$

Let us suppose that the particular function $w(y)$ that is chosen induces a choice of $y = r(n)$ for each level of the unobserved characteristic. Differentiating $N(y, n)$ totally with respect to n gives

$$\frac{dN}{dn} = (w' - c_y)r' - c_n = -c_n > 0. \tag{7}$$

Let $N(\underline{n}) = K$. Integrating (7) with respect to n we have

$$N(n) = K - \int_{\underline{n}}^{n} c_n(r(u), u)\, du. \tag{8}$$

Noting that $w = N + c$, and substituting in (6) the condition that total wages equal total productivity across the whole population, we have

$$K = \int_{\underline{n}}^{\bar{n}} \left[s - c + \int_{\underline{n}}^{n} c_n(r(u), u)\, du \right] f(n)\, dn. \tag{9}$$

The point of all this is simply to get rid of the function $w(y)$. Thus we imagine selecting an

[16] This analysis assumes that there are no adverse incentive effects of an income tax in terms of the choice between work and leisure. If there are such effects, then taxing income and taxing the signal are not equivalent in terms of the outcome.

upward-sloping function $y = r(n)$ to maximize Z. By using (8) and (9) we are assured that both constraints in the problem are satisfied, and the wage or tax schedule is nowhere in sight. We can calculate it later using the optimal schedule $r(n)$ by inverting that schedule to $n = h(y)$, substituting for n in $w^t = c_y$ and integrating to get the $w(y)$ that induce the signal choices that solve the optimizing problem.

The solution to the problem is:[17]

$$(10)\quad (s_y - c_y)f = c_{yn} \times \left[F(n) - 1 + \frac{\int_n^{\bar{n}} V^t f\, du}{\int_{\underline{n}}^{\bar{n}} V^t f\, du} \right].$$

If $V^t(N)$ is a constant so that $V(N)$ is linear, then the right side of (10) is zero. This is the case we have just looked at, where total net income is maximized by inducing the efficient choices of investment in education. It is just a special case of the more general problem and in that case, the optimal schedule is defined by $s_y = c_y$, for all n.

If $V(N)$ is very concave so that the first derivative falls rapidly, then at the limit the only thing that matters is the net income of the people with the lowest net income. This case is normally called the maximin case. In this case (10) becomes

$$(11)\qquad (s_y - c_y)f = -c_{yn}(1 - F).$$

The right side of (11) is positive, which means that investment in education is below the efficient and full information level. At the top of the range of n, $(1 - F)/f$ approaches zero and the investment in education is efficient.[18]

The maximin outcome in terms of investment in the signal is also the result that would obtain if there were a monopsonist purchaser of labor services. The reason is that the monopsonist wants to maximize the difference between productivity and gross wages, subject to the constraint that the net income of the individuals with the lowest net income not fall below some prespecified level. Thus the solution to the monopsonist's problem is to maximize the net income of the lowest level and then take it all away by lowering $w(y)$ uniformly across all education levels. Mathematically you are just reversing the objective function and one of the constraints.

The opposite extreme occurs when the welfare function is extremely convex so that its slope rises rapidly. In the limit that would mean valuing only the net incomes of those with the highest net incomes, the maximax case. In this case, the optimizing condition (10) becomes

$$(12)\qquad (s_y - c_y)f = c_{yn}F.$$

In this case, the right side of (12) is negative, implying that like the market equilibrium with imperfect information, the investment in education goes beyond the point at which the direction contribution to productivity is reached. At the bottom end of the range of n, F/f approaches 0 and thus at the lowest levels the investment in education is efficient.[19]

For the cases where $v(N)$ is not approaching an extreme (very convex or concave), one can see from (10) that the right side approaches zero when n approaches its minimum and its maximum values. Provided that $f(n)$ does not go to zero at the extremes, this would mean that at the end values of n, investment in education is efficient, that is $s_y - c_y = 0$. This is generally not a feature of the competitive equilibrium, and thus we may conclude that the competitive equilibrium is generally not the solution to some optimizing problem with this form of social welfare function. It is possible to rewrite the optimizing condition (10) in the following form:

[17] The derivation of this is another application of the calculus of variations. You displace the function $r(n)$ by $\delta\theta(n)$, differentiate the whole thing with respect to δ, and set the result equal to zero. After some reversing of orders of integration, you end up with an equation of the form $\int_{\underline{n}}^{\bar{n}} Q(r(n), n)\theta(n)\, dn = 0$. Since this must hold for all functions $\theta(n)$, the optimizing condition is $Q(r(n), n) = 0$. The application of all that to this case yields condition (10).

[18] This follows from the fact that $(1 - F)/f$ is the reciprocal of the derivative of $-\log(1 - F)$.

[19] This follows from the fact that F/f is the reciprocal of the derivative of $\log(F)$.

$$(s_y - c_y)f = -c_{yn}[1 - F] \times \left[1 - \frac{E(V'|u > n)}{E(V')}\right]$$

where $E(*|-)$ is the conditional expected value. For concave functions the term in the large square brackets on the right is always positive and thus investment in education is below the efficient level for all n except possibly at the end points. For convex welfare functions, the derivative is rising and thus the term in square brackets is negative. The levels of investment in education exceed the efficient ones.[20]

If you stand back from all this, the general pattern is fairly clear. The market equilibrium produces overinvestment in the signal because of the signaling effect, which is a private benefit to the investor, but yields no social benefit as its function is purely redistributive. The redistribution occurs in the direction of increasing the gross and net incomes of those with higher levels of education and productivity and hence income, the market equilibria tend to look more like the solutions to a second-best optimizing problem with a convex social welfare function. These, not surprisingly, are the welfare functions that weight higher net incomes more highly than lower net incomes.

VII. The Case of Education Costs Rising with the Unseen Attribute: Signaling Costs Vary the Wrong Way with Respect to Productivity

The standard case of signaling in which the signal has the capacity to survive and retain its informational content occurs when there is an unobservable attribute that is valuable to buyers (in the examples we are looking at, employers) and the costs of undertaking some activity that is observable are negatively correlated with the valued attribute. The labor-market examples, however, are slightly more complicated in that the unobserved attribute contributes to the individual's productivity and so does the signal. Thus, whatever the unobservable attribute is, it has two sources of value. One is the direct effect on productivity and the other is that lowering the costs of acquiring human capital also has value. Up to this point we have assumed that all of this works in the same direction. But it is possible that attributes that lower the cost of acquiring education might not be those that enhance productivity or might even be attributes that have a negative effect on productivity. We know that if $S_n = 0$, then you can get a signaling equilibrium, though in that case the signal is not needed: it simply identifies *ex post* those with lower education costs and hence higher levels of education.

The question I pose now is: can you get a signaling equilibrium when $c_{yn} > 0$, assuming that $s_n > 0$? From the analysis of equilibria, we know that the second-order condition for the individual's choice of education is satisfied only if $w'' - c_{yy} = c_{yn}\, dn/dy < 0$. This means that if $c_{yn} > 0$ one can have an equilibrium with signaling only if $dn/dy < 0$. So the question is: can that happen? The only way it could happen is if the human-capital effect is large in the sense that it overrides the negative signaling effect. Thus if $s_y \equiv 0$, it certainly cannot happen. The signal makes no direct contribution to productivity and sends the wrong message about the unobserved attribute. Everyone will set $y = 0$. But the answer to the question posed above is: yes, provided that the human-capital effect is large enough to override the negative signaling effect. There can be signaling equilibria in which signaling costs rise with the level of

[20] Those familiar with optimal tax problems will recognize the form of this optimizing condition. In the income tax problem, individual welfare measured in dollars is $u(w, l) = wl - t(wl) - h(\bar{l} - l)$, where w is income per hour worked, l is working hours, and $t(wl)$ is a tax on income, and $\bar{l}$ is a constant, which one can think of it as the total time available. Individuals maximize u with respect to l by setting $w - t'w - h' = 0$. The government has a revenue constraint so that $\int_{\underline{n}}^{\bar{n}} tf(w)\, dw = K$. If $v(u)$ is the welfare function then we want to maximize $\int_{\underline{n}}^{\bar{n}} v(u)f(w)\, dw$. Here you change the optimizing schedule to $y(w)$, get rid of tax function $t(y)$ by noting that $du/dw = h'y/w^2$, impose the revenue constraint to determine $u(0)$, and the displace the function $y(w)$ to find the optimum. The optimum is achieved when

$$f\left[1 - \frac{h'}{w}\right] = \left[1 - \frac{yh''}{h'}\right]\frac{h'}{w^2} \times (1 - F)\left(1 - \frac{E(v'|u > w)}{E(v')}\right).$$

Note that if $v(u)$ is linear, then the right side of this equation is zero. Work leisure choices are efficient and the government revenue is raised by a lump-sum tax, which does not distort the work-leisure choice.

the unobserved attribute that contributes to productivity. I will demonstrate this by example in a moment. Intuitively this should make sense if the human-capital effect is big enough, because then the attribute that is of real value to the individual and to the employer is the one that drives education costs down. The key is that the wage schedule has to be upward sloping in the signal, and this can happen if s_y is big enough even if $s_n < 0$. In this context, the correct statement about the condition that allows for a signaling equilibrium is that the net benefit of acquiring the signal has to be positively correlated with the gross effect on productivity. This could happen in segments of the population if very talented people face high opportunity costs associated with spending time on education.

This case may or may not be interesting from an empirical point of view, but it does illustrate that the more general formulation of the conditions for signaling are in terms of gross and net benefits and not just signaling costs. I did not realize this when I first worked on signaling equilibria. I thought then that the absence of the intuitively plausible negative cost correlation condition would destroy a signaling equilibrium. It is also of potential interest because the signaling effect is reversed. If $c_{yn} > 0$, then in equilibrium if there is signaling, $dn/dy < 0$, and thus

$$w^{t} = s_y + s_n \frac{dn}{dy} < s_y.$$

Therefore the negative signaling effect causes the private return to education to fall short of the social return and hence causes underinvestment in education.[21] It remains to show by a constructive example that this can happen at least in the theory, and perhaps also in the world.

We will show that this can happen with an example. Let $s(n, y) = ny^{\theta}$, and let $c(y, n) = n^{\alpha}y^{\beta}$, where $\alpha < 1$. Following the standard equilibrium analysis, we have

$$w^{t} = c_y = \beta n^{\alpha} y^{\beta - 1}. \tag{13}$$

In addition $w = ny^{\theta}$, or $n = wy^{-\theta}$. Substituting in (13), we have the differential equation in $w(y)$ that defines the equilibrium wage schedules:

$$w^{-\alpha}w^{t} = \beta y^{(\beta - 1 - \alpha\theta)}.$$

One of the solutions to this differential equation is:[22]

$$w(y) = Ky^{(\beta - \alpha\theta)/(1 - \alpha)}$$

where K is a constant. If you then proceed to determine equilibrium choices of education, they are given by

$$y(n) = Tn^{(1 - \alpha)/(\beta - \theta)}$$

where T is another constant. The second-order condition is satisfied and this all works if $dy/dn < 0$ and this will in fact be the case provided that $\alpha < 1$, and $\beta < \theta$. That is to say, signaling occurs and there is a separating equilibrium, if the elasticity of productivity with respect to education is larger than the elasticity of education costs with respect to education. Or in simple ordinary language, education is productive enough to justify its costs and override the negative signaling effect.

Another way to think about these relationships and this case is to change the unobserved variable from n to $t = \bar{n} - n$, where $\bar{n}$ is the highest level of n. In effect this defines the unobserved characteristic to be that which lowers education costs. Now investment in education will rise with t, provided there is a separating equilibrium, because $c_{yt} < 0$. But the effect of t on productivity is now negative, and if that effect is strong enough relative to the human-capital effect, then that will prevent a separating version of the signaling equilibrium. The reason is that unless the human-capital effect is large enough, the signaling effect will cause the wage function to be downward sloping in y. There will be a pooling equilibrium at $y = 0$ and that of course entails in the hypothetical example, reasonably dramatic underinvestment in the potential signal even though it is productive human capital.

[21] I thought Gary Becker whose pioneering work in human capital, household production functions, and much more, might appreciate this result.

[22] As this is simply an illustrative example, there is not much point in studying all the equilibria in the example.

VIII. The Information Contained in the Signal Can Improve Productivity

It should be noted that the information carried by the signal can be productive itself. This will occur if there is a decision that is made better or with greater efficiency, with better information. In the job market context, one could build this in as follows. Let productivity be $v(n, y, d)$, where d is a decision that the employer makes. It could be a decision about what type of job to assign to someone or a decision about required training. One can interpret the preceding analysis then, at least for equilibria in which the signal carries information in equilibrium, as $s(n, y)$ being the maximum of $v(n, y, d)$ with respect to d for each level of n and y. Note that this component of value is different from the human-capital effect. Unlike the human-capital effect, this element of value only exists if the signal is carrying information in equilibrium. You would lose it in a partial or complete pooling equilibrium even if that equilibrium included investment in education.

In situations like the Lazear model that we examined earlier, in which there is both pooling and separating components of the equilibrium, the addition of a decision that is made better with information will cause the value of the signal to rise, or more accurately the net benefits to rise. Those net benefits are passed on to individuals by virtue of competition on the employer side. Thus the relative sizes of the separating and pooling components of the equilibrium will shift in favor of the separating part. More people will opt to go to the firms that incur the expense of monitoring and learning productivity over time.

IX. Time and the Allocation of Time as a Signal and a Screening Device[23]

The principles that govern the survival of signals in markets can be applied to other contexts. The kind of mixed or imperfectly aligned incentive structure that characterizes markets is common in many situations. In particular, situations in which individuals have an unknown or imperfectly perceived level of interest in someone or something are ubiquitous. In these situations, one frequently observes that time spent by the individual is taken as a signal of interest. Literature and everyday experience suggest that the allocation of time is a ubiquitous and persistent signal sometimes used deliberately as a screening device.

The persistence of spending time as a signal of interest in something is a reflection of the fact that time is in short supply and everyone knows it. There is a nonzero shadow price on time, and hence the use of it in pursuit of some person, goal, or interest must mean that some implicit net benefit test has been passed for those who spent the time and failed for others who did not.

The use and interpretation of time as a signal is complicated and enriched by the accurate observation that there are both perceived and actual differences across people in the shadow price on time. These differences are used to interpret the signal. The allocation of a small amount of time by an individual with a high perceived shadow price on time has the same weight as a larger allocation of time by an individual with a lower perceived shadow price. As an example, those who have had visible leadership positions in virtually any organization know that the events they attend are taken as signals of interest and support and, conversely, when they do not show up, they deliberately and sometimes inadvertently signal a lack of interest, support, or enthusiasm. It is noteworthy that the informational content of these common signals is not necessarily related to whether the individual (the signaler) is actually needed or has a role at the event. In fact, having a role at the event can dilute the signal because it complicates the interpretation of the reason for the leader's presence.

Time is also routinely used as a screening device or as part of the price of admission to events at which the number of places is in short supply. For entertainment and sporting events it is common to find that a mixed system that involves both price and time is used to allocate the limited number of places at the event. The question that I address briefly in this section is why the mixed system would be used in preference to the pure price system. The intuitive answer is that willingness or ability to pay may be a poor signal of real interest in the event, not withstanding the fact that failure to use the price

[23] Part of this section is based on a nearly unknown (probably for good reason) paper (Spence, 1973b).

system is inefficient for two reasons.[24] The mixed system allocates places or access to those who would (and often do, if there is a secondary market) sell the place to someone with a higher willingness to pay—when that happens both parties are better off. Second, the time that is used to acquire the places is a deadweight loss under most circumstances. Notwithstanding the clear problems with the failure to use the price system, the use of the mixed system is very common and must mean that some objective is being pursued that is not respectful of the current distribution of income. For if the distribution of income is viewed as optimal, it is hard to conjure up a rationale for incurring the costs of the mixed system mentioned above.

I am going to use a simple example to look at the effects of using a mixed system for this class of resource allocation problems. It is not meant to be a general model, but rather to suggest what parameters are important in determining the performance characteristics of the outcome.

Let the value attached to an event by an individual be θ. One can think of this as a dollar valuation based on a roughly equal distribution of income, or one can think of it as a measure of value in a pure time system. What it is not is a current willingness to pay for access to the event. It is uniformly distributed in the population, on the interval [0, 1]. The value of money is λ. It also is distributed uniformly in the population on the interval [0, 1]. The assumption about the value of time at the margin is that it is equal to $G - a\lambda$. If the parameter a is zero, then it is a constant, if $a > 0$, then the marginal value of time is negatively correlated with the marginal value of income, and the converse is also true. The fraction of the population that can have access to the even is N, and we will assume that $N < 0.5$. Finally let p be the monetary price for access to the event and let T be the time price.

The set of people who attend the event is the people with combinations of (θ, λ) that satisfy the inequality

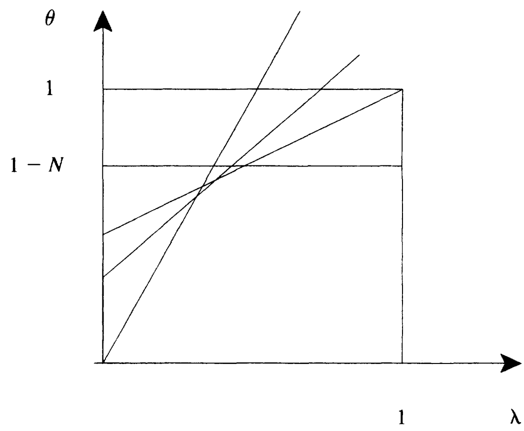

FIGURE 7. COMBINATIONS OF PRICE AND TIME TO RATION THE SCARCE RESOURCE

$$\theta - \lambda p - (G - a\lambda)T$$
$$= \theta - (p - aT)\lambda - GT \geq 0.$$

Let $S = GT$ and $R = p - aT$. Then the inequality above becomes

$$\theta - R\lambda - S \geq 0.$$

Let $A = \{(\theta, \lambda): \theta - R\lambda - S \geq 0\}$. Then because of the fixed supply, the area of the region associated with A has to be N.

Figure 7 is useful in analyzing the choices. Each of the lines inside the unit square represent combinations of R and S, so that the area above the line is N. The line through the origin represents the pure price system. As the time price T is raised, S rises and R falls. It is easy to see without a complicated proof that as T rises, the average value of θ is rising while the total numbers stay the same, so that the total gross value to consumers is rising. This continues until the line goes flat, at which point $R = 0$. It is also clear that the increases are largest for low levels of T or S, because initially one is trading low levels of θ for much higher levels.

It is possible that the mixed time and price system is commonly used simply to achieve this result, namely allocating the spots to those who attach the highest value to them, and that whoever makes these decisions does not think about or care about the deadweight cost of the time

[24] There are performers who want people with a genuine interest, and not just those with an interest and high incomes, at the events, and who feel that the quality of the event is affected by interaction between the performer and among members of the audience.

component of the price. But perhaps a more interesting question is whether the use of the mixed system ever increases the net benefits. We will assume that the dollar revenue is distributed in a neutral way so that the price component of the access charge generates neither net benefits or costs. The time component of the access charge however generates a deadweight cost.

The net benefits are

$$Z = \int_A \theta - SN + \frac{aS}{G} \int_A \lambda.$$

The second two terms are the deadweight cost of the time component of the access charge. We have already observed that the first term, the gross benefits, are a rising function of $S = GT$, reflecting the use of time as part of the allocation mechanism. This is true independent of the sign or size of the parameter "a." Moreover, one can see from Figure 7 that starting at $S = 0$, the gains from increasing S, or T, are largest at the start (that is near $S = 0$) because as the lines rise and rotate right, low average levels of the valuation parameter θ are being knocked out and high ones are being brought in. The difference between the exits and the entries steadily declines as S increases, and the lines become flatter.

This brings us to the question of whether the net benefits also increase at least over some range as s increases. It is somewhat tedious but not difficult to show that in this example that the slope of gross benefits at $S = 0$ is

$$\left. \frac{d \int_A \theta}{dS} \right|_{s=0} = \frac{1}{3} N.$$

The derivative of net benefits at $S = 0$ is[25]

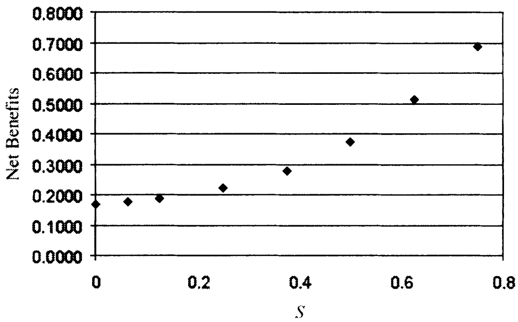

FIGURE 8. NET BENEFITS AS A FUNCTION OF THE USE OF TIME

$$\left. \frac{dZ}{dS} \right|_{S=0} = \frac{2}{3} N \left[N \frac{a}{G} - 1 \right].$$

If the parameter a is zero, then one can see that in this example, the deadweight costs outweigh the gain in gross benefits. However if "a" is positive and large enough, then the net benefits rise at least over some range. Why should this occur intuitively? Here price is a badly flawed signal of value. Time, if its marginal cost is uncorrelated with price, is a better signal of value, but not good enough to overcome the deadweight cost. But if the marginal cost of time is negatively related to the marginal value of income, then the combination of price and time is a much better signal of value. And if marginal time costs fall quite rapidly as the marginal value of income increases, then the cost of the time that is needed to *screen out* the high-income–low-value purchasers is relatively low.

To illustrate, Figure 8 shows the value of net benefits as a function of S, for a case in which $G = 2, N = 0.25$ and $a = 14$. Here the reader should ignore the higher levels of S, as they are associated with negative prices. Above a certain level of S, or time cost, the model starts redistributing income, which is counter to the spirit of the analysis.[26]

[25] These results follow from the following facts. For $R + S > 1, R = (1 - S)^2/2N$, gross benefits are $(1/R)[(1/3) \times (1 - S^3) - (S/2)(1 - S^2)]$, and the deadweight time cost is $(2/3)N^2[aS/G(1 - S)] - SN$. Similarly, when $R + S < 1$, which occurs as S becomes larger, $R = 2(1 - S - N)$, gross benefits are $(1/2)[(1 - S^2) - RS - (R^2/3)]$, and the deadweight cost of the use of time is $[(1 - s)/2] - R/3$. When $R + S = 1$, the line that separates those who are admitted from those who are not goes through the point (1, 1) which is why the formulae shift at that point.

[26] The existence of secondary market will to some extent reduce the benefit of the mixed system because it will draw into the primary allocation system entrepreneurs who are there to use their relatively low valuation on time to make

To summarize, while this simple example hardly disposes of the issue of how markets go about trying to solve complex allocation problems in a world of imperfect information, it does, I hope, suggest that the kinds of mixed systems that we observe in reality may be quite innovative solutions to these problems and probably should not be dismissed in all cases as mistakes made by people who do not understand the virtues of the price system. There may even be more interesting and less wasteful alternative currencies, such as hours spent in public service work, that could be used in conjunction with the price system to allocate highly valued access to important events. While such alternatives might not be as effective as time in undoing the effect of the price system, they may also have considerably less deadweight cost associated with them.

X. The Internet and the Changing Informational Structure of Markets

The proposition that the Internet (or more accurately, its relatively recent accessibility to a broad spectrum of users) has changed the informational structure of many markets, industries, and economies is probably not controversial. One could argue about how rapid the change has been from quite sudden on the one hand to a gradual evolution from the telegraph (which was the first time that communication over distance did not require the physical movement of something and hence became, in a certain sense, nearly instantaneous) through radio, the telephone, television, fax, etc., on the other. It has also been easy for some to dismiss all of this as a fad, relying on the boom-and-bust cycle that we have just observed. But in my judgment that would be a mistake. We know from the economics of information that there are periods in which enough has changed so that we operate in a data-free environment, that is, one in which there is temporarily little or no relevant data to constrain expectation. The data trickle in and the beliefs and expectations start to line up with reality again.

The more fundamental point is that investors and others were probably not wrong about the ultimate effect of this technology on markets and the economy, but almost everyone overestimated the speed with which individuals and organizations change their behavior, and we also underestimated the amount of difficult technical infrastructure that needed to be built in order to have the foreseen outcomes become a reality. We would undoubtedly have been better off if we had reminded ourselves of the important work of the late Zvi Griliches (1977) on the diffusion of innovation, as that would have caused some questioning of the assumption that what very intelligent people can foresee will come about quickly. I wanted also to recognize Zvi's work as an early and very important example of behavioral economics.

Nevertheless, there are powerful forces driving the outcomes and changing the informational structure of markets. Three of the most important are Moore's law, Metcalfe's law, and the dramatic reduction of the noise to signal ration in fiber-optic cable resulting in very large expansions in the throughput capacity of these pipes measured in signal processing terms. Moore's law is well known. It is an empirical observation that the number of transistors on a chip doubles every 18 months to 24 months. To a first approximation this has produced roughly a 10-billion-times cost reduction in the first 50 years of the computer age, dated from about 1950. Or to put it another way, things that were imaginable but unimaginably expensive in 1950 are essentially costless today. Economic historians will be better able than I to say whether there were comparable periods of change in the past with respect to costs of doing something important.

Metcalfe's law states that the value of a network to the entities attached to it is proportional to the square of the number of connected entities.[27] In economic terms this probably means

an arbitrage profit. They will bid up the price required to clear the primary market and hence drive some of those with high gross valuations of the event from that market.

[27] This should be thought of as an empirical regularity as well. It can be derived as follows. Suppose that the value a person to each other person on a network is x. If there are n people on the network and we add another person, the n people experience added value of nx and the new person experiences value of nx. If $V(n)$ is the value of a network, being the sum of its value to all connected people, then $V(n + 1) = V(n) + 2nx$. Therefore, $V(n) = 2x[1 + 2 + \cdots + (n - 1)]$. Using Gauss's formula for the sum of the integers from 1 to n, we have $V(n) = 2xn(n - 1)$, which is the quadratic relationship that Metcalfe's law asserts.

that the value and hence the speed of connecting accelerates as the numbers increase. This is sometimes referred to as the network effect.

From signal processing theory we know that the capacity of a channel is proportional to the log of one plus the ratio of signal to noise. Through scientific and technical advances (in addition to being able to use lasers with multiple wavelengths), the noise generated when the light reflects off the side of the fiber-optic cable is being reduced dramatically, causing very large increases in data rates on existing fiber. In economic terms this translates into large cost reductions in providing bandwidth.

These three effects interact with each other over time to produce accelerating economic effects. Much of the recent economic impact of the Internet is associated with reducing transaction costs of a variety of kinds. I will return to this subject in a moment. For most of these effects, one needs not just computing power, but also a reasonably ubiquitous reliable network that has standardized protocols, and is more or less always on. After all, transactions involve, almost by definition, more than one party, though increasingly one or more of the parties are machines. It is interesting and not surprising, therefore, that for the first 40 years of the proliferation of computers, not much measurable productivity increase was detectable in macroeconomic data. In the past ten years, the period in which the expansion of the network occurred, there are noticeable measurable productivity gains. It seems very likely that the productivity gains (the ones we can detect) are associated with reduced transaction costs, improved performance of markets, the ability to create new markets that were too expensive to create without the technology, and the important ability to take time and cost out of the coordination of economic activity, inside the firm and in the supply or value-added chain. That is what the network, with high and reliable capacity, with growing numbers connected to it and with enough computing power to run the endpoints and the nodes has permitted. It is a cumulative effect that we are just beginning to see.

Building models in economics is an art and a science. The science part consists of determining analytically the consequences of the assumption that create the structure of the model. The art consists in deciding what to put into the structure and what to leave out. Putting in too much makes the models intractable and of little use in illuminating the determinants of market performance. Putting in either too little or the wrong structural features makes the results, though tractable, uninteresting. A natural consequence of this is that parameters that do not change much from market to market or over time tend to be suppressed, as a matter of good practice in applied microeconomic theory. I mention all of this because it is possible and even likely that some parameters related to search costs, transaction costs, acquiring information, and geography have shifted reasonably rapidly in the past few years (or are in the process of shifting) as a result of the increasing speed, ubiquity, and connectivity of the Internet. This potential shift in parameters creates the opportunity to revisit aspects of the informational structure of markets and organizations, and in a sense, reintroduce the suppressed parameters as they have become interesting again. I would like to conclude this essay by suggesting a few areas in which such an inquiry might yield some interesting results.

One might have the impression from a first course in economic theory that markets combine supply and demand curves to produce prices, quantities, and to determine who buys the good. There is nothing wrong with that view. But in addition, a market economy performs lots of other functions. Potential buyers and sellers need to find each other. Often buyers and sellers need to acquire information about each other and about the product. If the buyers, sellers, and products are differentiated, then there is a matching problem that in one form or another needs to be resolved in the marketplace. If sellers are charging different prices, then buyers need to consider engaging in some sort of reasonable search for the lower prices. These functions are not performed costlessly, and the costs of performing them are frequently lumped together under the heading of transaction costs. Broadly these costs are being changed and lowered by the Internet.

XI. Buyers and Sellers Finding Each Other

Probably the most obvious example of reduced transaction costs is that associated with the rapid expansion of the markets for collectibles and used products of all kinds. In this arena

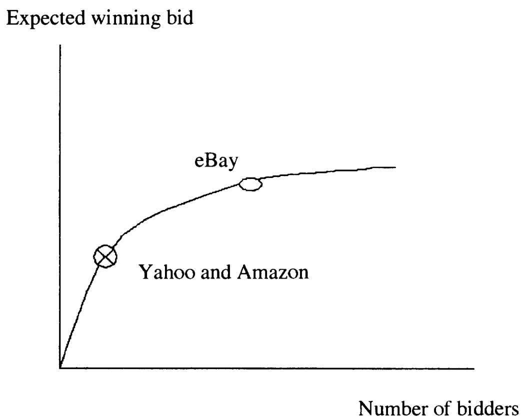

FIGURE 9. THE LIQUIDITY EFFECT IN AUCTION MARKETS

the overwhelmingly largest market making entity is eBay. There are on average 750,000 transactions on each day with an average total daily dollar volume of 30 million.[28] Most of these markets simply did not exist previously, and those that did exist were more costly to create, less efficient, and less liquid. It is worth noting that the drop in the cost of buyers and sellers finding each other is largely independent of physical geography (recall that both distance and time are simultaneously compressed with the Internet technology). This partial removal of the geographic bounds of markets makes them more liquid and in a certain sense more competitive. However, there are elements of natural monopoly in the market-making function. As the number of bidders increases, the expected winning bid rises in these auctions. Hence the marketplace that is in the lead will attract the sellers, and the variety of products will attract the buyers.[29] Figure 9 shows the relationship between the number of bidders and the expected selling price in an ordinary auction.

[28] Figures are from Jeff Skoll, cofounder of eBay. There are of course more markets open in any given day than there are transactions.

[29] There is a potential qualification to this tendency. Intelligent software agents are in principle capable of searching for low prices across marketplaces. If individuals were prepared to use these agents, that would eliminate the liquidity advantage of the larger market-makers. However, this does not appear to have happened, at least not to the extent that it significantly diminishes eBay's advantage. The making of online markets is pretty clearly a potential area for research.

XII. Searching for the Lowest Price

The late George J. Stigler recognized that finding the lowest price was an activity that required resources and that there was a trade-off between incurring costs of further search and the expected benefits of finding even lower prices. For prices that are posted in an Internet environment, the cost of finding the lowest price is pretty close to zero. In principle, this should eliminate price dispersion by eliminating one side of the trade-off. There was a kind of natural partial protection from price competition, the magnitude of which is a function of the search costs. The reduction or elimination of these search costs in the first instance increases competition. However, there is probably more to this story. The decision to post a price is a strategic decision and in the face of negligible search costs, it is possible that sellers' willingness to post prices will decline and that there will be more negotiated prices or prices that are tailored to the individual buyer.

XIII. The Boundaries of the Firm

There is a well-known and important literature in economics associated with Ronald H. Coase (1937, 1960) and Oliver E. Williamson (1970, 1971) and others. It deals in part with the fundamental question of what economic processes are contained within a firm and which are mediated by markets, that is, transaction between firms. One aspect of this general set of questions has to do with outsourcing: the decision to conduct a certain set of activities in-house or to contract with another entity to have them conducted for the firm. Generally on the side of outsourcing is the likelihood that certain functions can be performed better by specialists who have the advantages of economies of focus and scale. Countering that are the contracting, monitoring, and implementation costs of having a second party provide the service. Some of these countervailing costs are just transaction costs associated with complexity in communication. It is becoming reasonably clear that the Internet platform is reducing some of these transaction costs and hence tipping the balance in a number of areas in the direction of outsourcing. Many, if not most, informationally based services are efficiently deliverable and monitorable on the platform.

We have in a sense, for the period of transition, a natural laboratory in which transaction cost parameters are being shifted, resulting in new experimental behavior on the part of firms. In general, the range of economic activity that can be effectively coordinated across a complex multifirm supply chain is just starting to be explored by companies and those who do research on these subjects. The boundaries of the firm, transaction costs, supply-chain architecture and coordination, and outsourcing are all facets of a large mosaic in which incentives, communication and coordination, and the boundaries of the firm are worked out. The outsourcing extends to the employment relationship, where again the relative costs of in-house and the outsourced resources may have shifted. I hasten to add that these issues are far from being settled in the world of practice and in the world of economic research.

REFERENCES[30]

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, August 1970, *84*(3), pp. 488–500.

______. "The Economics of Caste and of the Rat Race and Other Woeful Tales." *Quarterly Journal of Economics*, November 1976, *90*(4), pp. 599–617.

______. "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics*, November 1982, *97*(4), pp. 543–69.

______. "Gift Exchange and Efficiency-Wage Theory: Four Views." *American Economic Review*, May 1984 (*Papers and Proceedings*), *74*(2), pp. 79–83.

Alchian, Armen A. and Demsetz, Harold. "Production, Information Costs, and Economic Organization." *American Economic Review*, December 1972, *62*(5), pp. 777–95.

Arnott, Richard J. and Stiglitz, Joseph E. "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review*, March 1991, *81*(1), pp. 179–90.

Arrow, Kenneth J. "The Role of Securities in the Optimal Allocation of Risk-Bearing." *Review of Economic Studies*, April 1964, *31*, pp. 91–96.

______. "The Organization of Economic Activity," in The Analysis and Evaluation of Public Expenditures: The PPB System, 1, Joint Economic Committee, 91st Congress, Washington, DC, 1969.

______. *Essays in the theory of risk bearing*. Chicago: Markham, 1971.

______. "Models of Job Discrimination," in A. Pascal, ed., *Racial discrimination in economic life*. Lexington, MA: Lexington Books, 1972.

______. "Higher Education as a Filter." *Journal of Public Economy*, July 1973, *2*(3), pp. 193–216.

Atkinson, Anthony. "Maximin and Optimal Income Taxation." Working paper, circa 1971.

Baker, George; Gibbs, Michael and Holmström, Bengt. "The Internal Economics of the Firm: Evidence from Personnel Data." *Quarterly Journal of Economics*, November 1994, *109*(4), pp. 881–919.

Becker, Gary S. *The economics of discrimination*. Chicago: University of Chicago Press, 1957.

______. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy*, October 1962, *70*, pp. 9–49.

______. "Human Capital and the Personal Distribution of Income." W. S. Wotinsky Lecture No. 1, University of Michigan, 1967.

______. *Human capital: A theoretical and empirical analysis, with special reference to education*, 2nd Ed. New York: Columbia University Press (for National Bureau of Economic Research), 1975.

______. *The economic approach to human behavior*. Chicago: University of Chicago Press, 1976.

Bergmann, Barbara R. "The Effect on White Incomes of Discrimination in Employment." *Journal of Political Economy*, March–April 1971, *79*(2), pp. 294–313.

______. "Occupational Segregation, Wages and Profits when Employers Discriminate by Sex." *Eastern Economic Journal*, 1976, *2*, pp. 103–110.

Bhattacharya, Sudipto and Guasch, J. Luis. "Heterogeneity, Tournaments, and Hierarchies." *Journal of Political Economy*, August 1988, *96*(4), pp. 867–81.

[30] *Editor's note:* The reference list includes some relevant works that are not cited in the text. This is an exception to usual journal style for the *AER*.

Bowles, Samuel. "Schooling and Inequality from Generation to Generation." *Journal of Political Economy*, May/June 1972, Pt. II, *80*(3), pp. S219–S51.

Cain, Glen G. "The Economic Analysis of Labor Market Discrimination: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 1987.

Carlton, Dennis W. "Planning and Market Structure," in John J. McCall, ed., *The economics of information and uncertainty*. Chicago: University of Chicago Press, 1982.

Coase, Ronald H. "The Nature of the Firm." *Economica*, November 1937, *4*, pp. 386–405.

______. "The Problem of Social Cost." *Journal of Law and Economics*, October 1960, pp. 1–44.

Cohen, K. J. and Cyert, R. M. *Theory of the firm: Resource allocation in a market economy*, 2nd Ed. Englewood Cliffs, NJ: Prentice-Hall, 1975.

Debreu, Gerard. *The theory of value*, Cowles Foundation Monograph No. 17. New Haven, CT: Yale University Press, 1959.

Diamond, Peter and Rothschild, Michael. *Uncertainty in economics*. New York: Academic Press, 1978.

Doeringer, P. and Piore, M. *Internal labor markets and manpower analysis*. Lexington, MA: D.C. Heath, 1971.

Downs, Anthony. *An economic theory of democracy*. New York: Harper & Row, 1957.

Ekern, Steinar and Wilson, Robert. "On the Theory of the Firm in an Economy with Incomplete Markets." *Bell Journal of Economics and Management Science*, Spring 1974, *5*(1), pp. 171–80.

Fama, Eugene F. "Agency Problems and the Theory of the Firm." *Journal of Political Economy*, April 1980, *88*(2), pp. 288–307.

Fogel, Robert William and Engermann, Stanley L. *Time on the cross: The economics of American Negro slavery*. New York: W. W. Norton, 1989.

Gibbons, Robert. "Piece-Rate Incentive Schemes." *Journal of Labor Economics*, Pt. 1, October 1987, *5*(4), pp. 413–29.

Gibbons, Robert and Katz, Lawrence F. "Layoffs and Lemons." *Journal of Labor Economics*, October 1991, *9*(4), pp. 351–80.

Griliches, Zvi. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica*, January 1977, *45*(1), pp. 1–22.

Grossman, Sanford J. and Stiglitz, Joseph E. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, June 1980, *70*(3), pp. 393–408.

Hart, Oliver D. "Optimal Labour Contracts under Asymmetric Information: An Introduction." *Review of Economic Studies*, January 1983, *50*(1), pp. 3–35.

Hirshleifer, Jack. "The Private and Social Value of Information and the Reward to Inventive Activity." *American Economic Review*, September 1961, *51*(4), pp. 561–74.

Holmström, Bengt. "Moral Hazard and Observability." *Bell Journal of Economics*, Spring 1979, *10*(1), pp. 74–91.

______. "Moral Hazard in Teams." *Bell Journal of Economics*, Autumn 1982, *13*(2), pp. 324–40.

Jencks, Christopher; et al. *Inequality*. New York: Basic Books, 1972.

Kreps, David M.; Milgrom, Paul; Roberts, John and Wilson, Robert. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory*, August 1982, *27*(2), pp. 245–52.

Lazear, Edward P. "Salaries and Piece Rates." *Journal of Business*, July 1986, *59*(3), pp. 405–31.

______. *Personnel economics for managers*. New York: Wiley, 1998.

______. "Educational Production." *Quarterly Journal of Economics*, August 2001, *116*(3), pp. 777–803.

Leibenstein, Harvey. "Allocation Efficiency vs. X-Efficiency." *American Economic Review*, June 1966, *56*(3), pp. 392–415.

Levhari, D. and Weiss, Yoram. "The Effect of Risk and Investment in Human Capital." *American Economic Review*, December 1974, *64*(5), pp. 950–63.

McCall, J. J. "Economics of Information and Job Search." *Quarterly Journal of Economics*, February 1970, *84*(1), pp. 113–26.

Mincer, Jacob. *Schooling, experience, and earnings*. New York: Columbia University Press (for NBER), 1974.

Mirrlees, James A. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, April 1971, *38*(114), pp. 175–208.

______. "The Optimum Town." *Swedish Journal of Economics*, March 1972, *74*(1), pp. 114–35.

______. "The Optimal Structure of Incentives with Authority within an Organization." *Bell Journal of Economics*, Spring 1976, *7*(1), pp. 105–31.

Modligiani, Franco and Miller, Merton. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review*, June 1958, *48*(3), pp. 261–97.

Nelson, Phillip. "Information and Consumer Behavior." *Journal of Political Economy*, March–April 1970, *78*(2), pp. 311–29.

Pencavel, John H. "Work Effort, On-the-Job Screening, and Alternative Methods of Remuneration," in Ronald Ehrenberg, ed., *Research in labor economics*, Vol. 1. Greenwich, CT: JAI Press, 1977.

Radner, Roy and Rothschild, Michael. "On the Allocation of Effort." *Journal of Economic Theory*, June 1975, *10*(3), pp. 358–76.

Ramsey, Frank Plumpton. "A Contribution to the Theory of Taxation." *Economic Journal*, March 1927, *37*, pp. 47–61.

Riley, John G. "Competitive Signalling." *Journal of Economic Theory*, April 1975, *10*(2), pp. 174–86.

Rogerson, William P. "Repeated Moral Hazard." *Econometrica*, January 1985, *53*(1), pp. 69–76.

Ross, Stephen A. "The Economic Theory of Agency: The Principal's Problem." *American Economic Review*, May 1973 (*Papers and Proceedings*), *63*(2), pp. 134–39.

Rothschild, Michael. "Models of Market Organization with Imperfect Information: A Survey." *Journal of Political Economy*, November–December 1973, *81*(6), pp. 1283–1308.

______. "Searching for the Lowest Price When the Distribution of Prices Is Unknown." *Journal of Political Economy*, July/August 1984, *82*(4), pp. 689–711.

Rothschild, Michael and Stiglitz, Joseph E. "Increasing Risk: I. A Definition." *Journal of Economic Theory*, September 1970, *2*(3), pp. 225–43.

______. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, November 1976, *90*(4), pp. 629–49.

Salop, Joanne and Salop, Steven. "Self-Selection and Turnover in the Labor Market." *Quarterly Journal of Economics*, November 1976, *90*(4), pp. 619–27.

Samuelson, Paul A. "Foundations of Economic Analysis." Harvard Economic Studies, HB 31 .H33 v.80

Schelling, Thomas. *The strategy of conflict*. Cambridge, MA: Harvard University Press, 1960.

Simon, Herbert A. *Models of man*. New York: Wiley & Sons, 1957.

Solow, Robert M. "Investment and Technical Progress," in Kenneth J. Arrow, S. Karlin, and P. Suppes, eds., *Mathematical models in the social sciences*. Stanford, CA: Stanford University Press, 1960.

Spence, A. Michael. "Job Market Signaling." *Quarterly Journal of Economics*, August 1973a, *87*(3), pp. 355–74.

______. "Time and Communication in Economic and Social Interaction." *Quarterly Journal of Economics*, November 1973b, *87*(4), pp. 651–60.

______. "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution." *Journal of Economic Theory*, March 1974a, *7*(3), pp. 296–332.

______. *Market signaling: Informational transfer in hiring and related processes*. Cambridge, MA: Harvard University Press, 1974b.

______. "An Economist's View of Information," in Carlos Cuadra, ed., *Annual review of information science and technology*, Vol. 9. American Society for Information Science, 1974c.

______. "The Economics of Internal Organization: An Introduction." *Bell Journal of Economics*, Spring 1975, *6*(1), pp. 163–72.

______. "Informational Aspects of Market Structure: An Introduction." *Quarterly Journal of Economics*, November 1976a, *90*(4), pp. 591–97.

______. "Competition in Salaries, Credentials, and Signaling Prerequisites for Jobs." *Quarterly Journal of Economics*, February 1976b, *90*(1), pp. 51–74.

______. "Consumer Misperceptions, Product Failure and Producer Liability." *Review of Economic Studies*, October 1977, *44*(3), pp. 561–72.

______. "Markets and Imperfect Information." *Portfolio, 5*(5).

______. "Signaling and Screening," in S.

Rosen, ed., *Low income labor markets*, National Bureau of Economic Research (Cambridge, MA) Working Paper.

Spence, Michael and Zeckhauser, Richard J. "Insurance, Information, and Individual Action." *American Economic Review*, May 1971, (*Papers and Proceedings*), *61*(2), pp. 380–87.

______. "The Effect of the Timing of Consumption Decisions and the Resolution of Uncertainty on the Choice of Lotteries." *Econometrica*, March 1972, *40*(2), pp. 401–03.

Stigler, George J. "The Economics of Information." *Journal of Political Economy*, June 1961, *69*, pp. 213–25.

______. "Information in the Labor Market." *Journal of Political Economy*, October 1962, *70*, pp. S94–S104.

Stiglitz, Joseph E. "Approaches to the Economics of Discrimination." *American Economic Review*, May 1973 (*Papers and Proceedings*), *63*(2), pp. 287–95.

______. "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies*, April 1974, *41*(2), pp. 219–56.

______. "Incentives, Risk, and Information: Notes Toward a Theory of Hierarchy." *Bell Journal of Economics and Management Science*, August 1975, *6*(2), pp. 552–79.

______. "The Theory of 'Screening', Education, and the Distribution of Income." *American Economic Review*, June 1975, *65*(3), pp. 283–300.

Stiglitz, Joseph E. and Weiss, Andrew. "Incentive Effects of Terminations: Applications to the Credit and Labor Markets." *American Economic Review*, December 1983, *73*(5), pp. 912–27.

Thurow, Lester. "Education and Economic Equality." *Public Interest*, Summer 1972, *28*, pp. 66–81.

Tversky, Amos and Kahneman, Daniel. "The Framing of Decisions and the Psychology of Choice." *Science*, 1981, *211*, pp. 453–58.

______. "Judgment under Uncertainty: Heuristics and Biases." *Science*, September 1974, *185*, pp. 1124–31.

Waldman, Michael. "Job Assignments, Signalling, and Efficiency." *Rand Journal of Economics*, Summer 1984, *15*(2), pp. 255–67.

______. "Up-or-Out Contracts: A Signaling Perspective." *Journal of Labor Economics*, April 1990, *8*(2), pp. 230–50.

Weiss, Andrew. "Job Queues and Layoffs in Labor Markets with Flexible Wages." *Journal of Political Economy*, June 1980, *88*, pp. 526–38.

______. *Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersion*. Princeton, NJ: Princeton University Press, 1990.

Wilson, Robert. "The Theory of Syndicates." *Econometrica*, January 1968, *68*(1), pp. 119–32.

Williamson, Oliver E. *Corporate control and business behavior*. Englewood Cliffs, NJ: Prentice Hall, 1970.

______. "The Vertical Integration of Production: Market Failure Considerations." *American Economic Review*, May 1971 (*Papers and Proceedings*), *61*(2), pp. 112–23.