Volume Title: Studies in Labor Markets

Volume Author/Editor: Sherwin Rosen, ed.

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-72628-2

Volume URL: http://www.nber.org/books/rose81-1

Publication Date: 1981

Chapter Title: Learning by Observing and the Distribution of Wages

Chapter Author: Stephen Ross, Paul Taubman, Michael L. Wachter

Chapter URL: http://www.nber.org/chapters/c8916

Chapter pages in book: (p. 359 - 386)

11 Learning by Observing and the Distribution of Wages

Stephen Ross, Paul Taubman,
and Michael Wachter

It is well known that the more educated have higher earnings. There are several possible explanations for this fact. In the framework of the human capital model, education produces skills that are rewarded in the marketplace. Of course, part of the observed differences in earnings by education level may arise because the more educated are also more able, but such an observation is not in conflict with the human capital production model. However, some economists have gone beyond the observation that ability and earnings are correlated to argue that the only or primary role of education is to signal who are the more able.[1]

It is extremely difficult to distinguish between the human capital and signaling models based on information confined to education and its returns, though there have been several papers on the subject (see Taubman and Wales 1973; Riley 1978). When one set of empirical information is not able to distinguish theories, an obvious approach is to test the theories against other types of information. In this paper we first consider what other information about the distribution of income is available. We summarize the human capital model explanation of these additional facts. We then present an alternative model based on two assumptions; the different skills and wage changes which occur over a

Steven Ross is Professor of Economics, Yale University Graduate School of Management.

Paul Taubman is Professor of Economics, University of Pennsylvania, and Research Associate, National Bureau of Economic Research.

Michael Wachter is Professor of Economics, University of Pennsylvania.

The research in this paper was supported by the National Science Foundation under grant Soc 77–08568. Additional support was obtained from funds granted to the Institute for Research on Poverty, University of Wisconsin-Madison, by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964.

person's life cycle do so because a firm monitors the worker's performance on a particular job. As a consequence of monitoring, the firm may alter the worker's job assignment to fit the updated evaluation of performance. Finally, we demonstrate that the observed earnings data are as consistent with our model as with the human capital model.

11.1 Background

The multitude of empirical work over the past two decades has generated many results about the distribution of annual earnings and its evolution as people age. Some of the results which are found in nearly all studies include the following:[2] (1) in general, average earnings increase with years of schooling, quality of schooling, and years of work experience; (2) the age-earnings profile slopes upward, but it does so at a decreasing rate; and (3) the variance of earnings, or of its log, is not invariant with respect to age. Often the variance of the log of earnings is U-shaped, indicating a larger dispersion for young workers and older workers. Diamond, et al. (1976), however, present some evidence that this variance decreases continuously as people age.

There are other empirical features which have been found in several studies based on specialized and as yet infrequently available samples. If these are substantiated in other samples, they should prove to be important elements in income distribution models. To begin with, the correlation between (natural log) earnings and variables such as years of schooling or IQ is lowest when years of work experience is small and increases with experience for at least the first 7–10 years of experience.[3] Second, Lillard and Willis (1977), using a nationwide random sample, find that 70%–80% of the variance in annual earnings is attributable to permanent income and that annual fluctuations about the permanent level have a serial correlation coefficient of about 0.3. Also Taubman (1975) finds that in the NBER-TH sample, the average percentage growth rate in earnings between 1955 and 1969, when the men averaged thirty-three and forty-seven years old, respectively, is the same regardless of their 1955 earnings level—except for men with the very highest and very lowest income. Diamond, et al. (1976), using Social Security data for 1957–72, find that when men are grouped by their permanent earnings level, the slopes of the age-earnings profiles are quite similar. Those with permanently low income, however, tend to reach their peak income at a slower rate.

Recently Mincer (1974) has modified and extended the human capital model to explain many features of the evolution of the distribution of annual earnings as people age. Mincer's model for earnings in year t can be written as

$$\ln Y_t = bS + cK_t - I_t \tag{1}$$

where Y is earnings, S is years of schooling, K is the stock of on-the-job training capital, and I is investment in on-the-job training. Both K and I in equation 1 would be unobserved but denominated in units of time. K and I are assumed to differ across individuals.

Mincer explains the age profiles for natural log earnings and their variance and for correlation coefficients by changes in the level and distribution of K_t and I_t. For example, consider the variance in ln Y_t in equation 2.

$$\sigma^2 \ln Y_t = b^2\sigma_S^2 + c^2\sigma_{K_t}^2 + \sigma_{I_t}^2 + 2bc\sigma_{S,K_t} - 2b\sigma_{S,I_t} - 2c\sigma_{K_tI_t} \tag{2}$$

Note first that this variance need not remain constant as a cohort ages since components such as $\sigma_{K_t}^2$ can change. Moreover equation 2 need not imply monotonic changes in $\sigma^2 \ln Y_t$. Consider for example three time periods. In the initial year of working, K is zero for all workers. In Mincer's so-called overtaking period, $cK_t = I_t$. Finally, in the peak period, I_t is zero. Under certain conditions all individuals will be in the overtaking period and subsequently in their peak period at the same time. In the overtaking period, $\sigma^2 \ln Y_t = b^2\sigma^2 \ln S$. The variances in the initial period will also involve the variances and covariance terms of I and K, respectively. Depending on the size of the coefficients b and c and the sign of σ_{SI_t}, $\sigma_{K_tI_t}$, and σ_{SK_t}, various age profiles for $\sigma^2 \ln Y_t$ can be generated.

In Mincer's (1975) model the time coefficient of schooling is a constant, b. The corresponding coefficient that would be estimated by ordinary least squares could change, however, because K_t and I_t would be omitted variables whose covariances with S could vary with t or age. Similarly, Mincer's model would suggest that the correlation of S with ln Y_t are functions of age.

In Mincer's work, it is generally assumed that individuals and firms always know a person's marginal product at each and every point in time, in each and every occupation, and with and without various training programs. There have been some attempts to incorporate uncertainty about future wages and/or marginal products into this framework using the expected utility approach (see, for example, Thaler and Rosen 1976; Levhari and Weiss 1974; Weiss 1972; and Fardoust 1978). We think, however, that informational uncertainty is best approached in a different fashion.

A basic problem that faces firms and individuals is matching the right person with the right job, an issue which inherently involves uncertainty. This matching process is made difficult because the particular job may not be well defined by a firm. Even if the job is defined, the requisite skills may not be easily measured in advance of hiring workers, and may not correlate well with any easily observed set of personal attributes.[4]

Recently economists have begun to examine in some detail how workers and firms solve this matching problem in the face of uncertainty. The Spence (1973) Arrow (1973) signaling model basically argues that in some instances individuals invest in signals so that firms can better distinguish among workers of different skills. Using this information, the firm separates workers into categories with differing marginal products and real wages. Inherent in much of this work is the notion that if a worker acquires a signal such as schooling, he is always thought of as a better worker and paid the average wage in that category. To make this assumption more palatable, it is argued that poor workers do not invest in the signal because the investment costs are higher for these workers.

The signaling literature appears to suggest different conclusions from the human capital model on two separate points. First, population-wide increases in investment in schooling need not lead to increases in earnings. Second, one-to-one correspondence of real wages and marginal productivity in the human capital model need not hold. Even if the signal is unbiased, under a range of assumptions, considerable latitude exists for randomness in the eventual income distribution. In models which assume that signals are biased, the randomness is considerably strengthened.

In this paper we provide a somewhat different critique of the human capital framework and, at the same time, of the signaling literature. Our model is based on the existence of an internal labor market in the firm. An important function of this market is to sort workers into jobs where they are most productive. To focus attention on the potential importance of sorting, we shall examine a model where the only function of the internal labor market is to sort workers. No on-the-job training is provided; this is a pure ability model. In our simplified model, we show that we also can explain the observed empirical facts concerning age-earnings profiles, changes in the variance of earnings, and the movements in the correlation coefficients by assuming that firms unravel the uncertainties about the abilities of the work force.

In addition, our model, as opposed to the traditional signaling model, implies that in the long run, the distribution of annual real wages may or may not ever equal marginal productivities. Moreover, the distribution of present discounted value of real wages will not equal the distribution of marginal products. If there is "incomplete" sorting, distinctions between marginal products and real wages will persist. If there is "complete" sorting, these differences will eventually disappear. How quickly these differences disappear will affect the extent to which the present discounted value of expected lifetime earnings departs from the distribution of real wages.

11.2 A Sorting Model

We assume that each worker comes to the firm with certain observable characteristics such as education and age. The exact skill of the worker is unknown, but his characteristics permit the assignment of a probability vector

$$p = \langle p_{1j} \ldots p_n \rangle j;\ \sum_j p_i = 1;\ jp_i \geq 0$$

The number p_i indicates the subjective market probability that a worker is of type or skill i. A cohort of workers is defined as a group of workers with the same signal p. Workers initially enter a firm with a p derived purely from external (to the firm) characteristics. Once workers are assigned to jobs, the vector p is altered to reflect the internal labor market experience of the workers.

The firm is defined by a job technology which describes its output as a function of a job structure. This job structure is the internal labor market of the firm. For purposes of simplicity, we assume that the internal labor market is "open" in that horizontal movements across firms can occur at any point along the job structure. A "perfectly closed" internal labor market would be one where horizontal interfirm mobility was possible only at the time when workers are first choosing a firm. After the initial assignment, interfirm movement would require a "demotion" in the job structure matrix.

In an open structure the firm never pays each cohort less than its expected marginal return. Hence, on a worker with characteristics p,

$$w(p) \geqq E(p)$$

where $w(p)$ denotes the wage structure. Implicit in the notation $w(p)$ is the assumption that the wage structure is functionally dependent on the workers' signals.[5]

At any point in time, the firm hires a group of workers. For simplicity we assume that these workers can be placed into a discrete number of categories where each category has a separate p. Clearly, the workers with the highest signals, e.g., the best education, will be placed in jobs with higher starting salaries. This follows from our assumption of an open internal labor market. Each firm must pay its various cohorts a wage no less than $E(p)$ or it will lose the group. This assumes, of course, that the market is rational in the way it processes information on signals p.

The basic construct of the firm is the job ladder or matrix. Firms are viewed as having a technology describing the output of particular types of workers across the job array. Suppose, for example, that there are n basic types of workers and m jobs at which workers can be employed. The symbol a_{ij} will denote the (marginal) output of a worker of type i assigned to job j.

If there were no uncertainty whatsoever about the market type, and with constant returns (or using the marginal job matrix), then the ith worker or cohort would be assigned to the job j^* with

$$\sup \max_j a_{ij} = a_{ij^*} \geq a_{ij} \text{ all } j.$$

Under competitive conditions, a_{ij^*} would also be the cohort's wage. With uncertainty, and under the conditions discussed above, the workers in the cohort would receive a wage equal to their expected product. Typically, the exact worker type is in fact unknown, and a category is defined by certain educational and personal characteristics which permit only the assignment of the probability vector p.

The expected return on a worker in this cohort in job j then is given by

$$E_j(p) = \sum_i p_i\, a_{ij}$$

and the firm will assign workers to the job so as to maximize this return. That is, given our assumption of an open internal labor market, workers with signal p will receive

$$w(p) = \max_j \sum_i p_i\, a_{ij}$$

11.2.1 Properties of the Wage Structure

The definition of the wage structure permits us to establish a close connection between wages and jobs. More specifically, the wage structure $w(p)$ contains all of the information on the job structure a_{ij} in the same way as the cost function embodies information on the firm's technology.

If we think of the job structure as representing a set A of jobs $a \varepsilon A$, then

$$w(p) = \sup_{a \varepsilon A} \sum_i p_i a_i = \sup_{a \varepsilon A} p.a.$$

Since $w(p)$ is a support function for the set A, it is well known from duality theory that the set

$$A^* = \{a \mid pa \leq w(p) \text{ for all } p \varepsilon S\}$$

contains A. Furthermore, if A is closed and convex and admits free disposal in that some of the worker's output can be thrown away, the set $A^* = A$. Even if not, the wage structure derived from A^* will be the same as that from A, and, thus, information on the wage structure alone will not permit us to infer more about the job matrix than that A^* is its convex hull.

A second property about such a wage structure is that it is a convex function of the probability vector p. Formally, for any two signals x and y

$$\begin{aligned} w\left(\frac{1}{2}x + \frac{1}{2}y\right) &= \sup_{a \varepsilon A} \left(\frac{1}{2}x + \frac{1}{2}y\right) \cdot a \\ &= \frac{1}{2} \sup_{a \varepsilon A} (xa + ya) \\ &\leq \frac{1}{2} \sup_{a \varepsilon A} xa + \frac{1}{2} \sup_{a \varepsilon A} ya \end{aligned}$$

$$=\frac{1}{2}w(x)+\frac{1}{2}w(y)$$

In other words, suppose an individual has a signal [1/2 1/2]; that is, he is thought to have a 1/2 chance of being a type x worker and a 1/2 chance of being a type y worker. Then his wage cannot be greater than the average of the wages for an x worker and a y worker. Indeed, under very general circumstances, as shall be shown below, the wage must be lower than the average for x and y workers. Although this may seem paradoxical in a risk neutral world, it has a simple explanation which is central to the sorting model. Knowledge that the worker is of type x or y will permit a more optimal job placement than in the 50–50 uncertain situation. For example, consider the job structure in figure 11.1. A worker with a signal [1/2 1/2] will be paid a wage of 5 and placed in J_1. In the next period, the firm will be able to tell whether a worker is an x or a y by whether he had produced 10 units of output or zero output. With this new information, the worker's signal changes to either [1 0] or [0 1]. If he produced 10, he will be labeled an x worker, left in job 1, and paid a wage of 10. If he produced 0, he will be labeled a y worker, changed to job 2, and paid a wage of 9. In either case, his wage increases.

The basic proposition, however, is not that all workers have wage increases, but rather that the average wage increases. For example, if the job structure is as shown in figure 11.2, the initial wage for the cohort is 5 and they are all placed in job 1. In the following period, the x workers receive a wage increase to 10 and the y workers receive a wage cut to 1. The average wage is 5.5 which is greater than the initial average wage. (The fact that 1/2 the workers are x while 1/2 are y follows from the assumption that the signal is unbiased. If this were not the case, the average wage could decline.)

11.2.2 Upward-Sloping Age-Earnings Profiles

The knowledge that the wage structure is a convex function of the signals permits us to derive an important result. Even in the absence of

	J_1	J_2
x	10	0
y	0	9

Figure 11.1

	J_1	J_2
x	10	0
y	0	1

Figure 11.2

any change in the intrinsic productivity of a worker over a lifetime, the market's perception of a worker's ability tends to alter with work experience. The job performance might reveal that a particular worker has been overvalued or undervalued, but on average his wage will increase over the worker's lifetime. This arises simply from the procedure of sorting workers over their lifetime.

Theorem 1: The average wage for a cohort will rise over time, i.e., age-earnings profiles rise.

Proof: Since we have assumed that the signal implies a probability vector about the true population proportion, it is sufficient to show that the expected wage increases with any initial signal. Let p^0 be the initial signal and p^1 the random signal at time 1 dependent on both p^0 and the information acquired in the first job. If a^0 is the initial job then

$$\begin{aligned} w(p^0) &= \sup_{a \varepsilon A} p^0\, a \\ &= p^0\, a^0 \end{aligned}$$

Now, if a worker is of type i, then let I_i denote the information such a worker gives in job a^0, and $p^1(I_i)$ be the probability vector for a worker of type i in job a^0.

It is important to realize that we do not always have full information about a worker simply by observing the worker on a job. For example, suppose in job a, that $a_1 = a_2 = , \ldots , = a_m$, i.e., all workers perform the same. The job matrix has a column of identical numbers. Clearly, if the only information is the productivity of the worker, then observing the worker in job a^0 provides no additional knowledge about the worker, and the future signal equals

$$p^1(I_i) = p^1 = p^0$$

the initial signal. It is in this case, which we call "incomplete" sorting, that the average wage is constant over time and does not increase.

In general, though, the market will obtain on-the-job information, and $I_k \neq I$ if $k = \ell$. Now, $p_k^1(I_i)$ is defined as the probability that the worker is of type k, conditional on having the information I_i. At time 0 the probability that a worker will be of type k will be given by

$$\begin{aligned} p_k^0 &= \sum_i p_k^1(I_i)\, p_i^0 \\ &\equiv E\{p_k^1\} \end{aligned}$$

Thus, p^1 must be a probability vector with expected value or, for the cohort, average value equal to p^0. This makes intuitive sense since at time 0 with p^0 the market cannot anticipate receiving information that will lead to a p^1 systematically biased from p^0; if such information were anticipated in a rational framework it would already be reflected in p^0.

The rising age-earnings profile now follows directly from the convexity of the wage function,

$$\begin{aligned} E\{w(p^1)\} &\geq w(E\{p^1\}) \\ &\geq w(p^0), \end{aligned}$$

with strict inequality if p^1 ranges over some nonlinear portion of the wage structure.

Theorem 1 verifies that the sorting, on-the-job ladder model implies the first stylized empirical observation of earnings profiles. This theorem is strikingly robust since no structure need be imposed on the job matrix. Sorting alone is sufficient to impart a positive slope to the age-earnings profile.

It is useful to define a sorting equilibrium as occurring when all workers or worker cohorts hold the identical jobs in periods $t+1$ as they did in t. The sorting process can be in equilibrium in two situations. The first, which we refer to as a complete sorting, occurs whenever all of the workers are placed optimally in the job structure. Incomplete sorting occurs when workers are not optimally placed, but the job structure does not permit further sorting. As indicated above, this results whenever a column in the job matrix has identical entries.

Let the job structure be given by figure 11.3, and suppose a type x worker belongs to a cohort whose initial signal is $p^0 = (1/2, 1/2)$. With this initial signal the cohort will be assigned to job J_1 at which all members will produce 2 units and in which no information will be obtained. The x workers in the cohort will now produce below what they could produce in job J_2. Of course, this result depends critically on the assumption that the worker knows only the initial signal p^0.

Suppose, for example, that the worker knew he was a type x, but only signaled $p^0 = (1/2, 1/2)$. Such a worker could volunteer to work for less than 2 units in job J_y to prove himself a type x. Even better, the worker could agree to a contingent on performance contract. If he produced 3 units in job J_y he would receive 2+ units with the remainder for the firm; otherwise, he would receive nothing. Nor is there any moral hazard dilemma with such a contract. Quite to the contrary, only those who

$$\begin{array}{ccc} & J_1 & J_2 \\ x & 2 & 3 \\ y & 2 & 0 \end{array}$$

Figure 11.3

knew themselves to be of type x would accept jobs J_y under such conditions; other workers would stay with J_x. We do encounter problems if we let the acceptance of the offer alone represent a signal, for then wages might be paid ahead of performance which would create a moral hazard. We will ignore such difficulties below, and return to our initial assumption that firms and workers have the same perception of worker signals.

11.2.3 The Diminishing Rate of Increase

The second stylized fact is that age-earnings profiles rise at eventually diminishing rates. While this hypothesis is as consistent with the job ladder model as with the human capital model, its derivation requires somewhat more structure than that required for theorem 1.

It is tempting, for example, to argue that the incremental value from additional information along the job ladder must be declining, and that wages, therefore, while rising must do so at a diminishing rate. That is, the initial jobs contribute a great deal of new information on a worker cohort, allowing for major revisions in this signal p. After several job changes, however, the new information flow decreases so that the increase in wages slows also. Although this is an attractive initial point, it is not sufficient to prove diminishing ratios of wage growth.

Suppose that the job ladder takes the form shown in figure 11.4, and that the initial signal is $p^0 = (p_x^0, p_y^0, p_z^0) = (1/2, 1/3, 1/6)$. The highest initial wage is attainable by placing this group in J_1.

$$\begin{aligned} w(p^0) &= \max_{\{x,\ y,\ z\}} \ p^0 a \\ &= p^0 J_x \\ &= (1/2 \times 5) + (1/3 \times 5) + (1/6 \times 10) \\ &= 5\ 5/6 \end{aligned}$$

If the worker produces 10 units, then he will be identified as a type z and left in J_1, but if 5 units are produced he can be either a type x or y. The x or y workers are placed on job J_2 and the z workers remain in J_1. The expected wage is thus

$$\begin{aligned} E\{w(p^1)\} &= (1/2 \times 10) + (1/3 \times 0) + (1/6 \times 10) \\ &= 6\ 2/3 \end{aligned}$$

	J_1	J_2	J_3
x	5	10	0
y	5	0	10
z	10	0	0

Figure 11.4

At time 2, all of the workers will be fully identified and placed in the correct job. Hence

$$E\{w(p^2)\} = w(p^2) = 10$$

The age-earnings profile increases at an increasing rate between periods 1 and 2. It initially increases from 5 5/6 to 6 2/3, but then it jumps to 10 in the final period.

This same job structure, though, can result in a concave age-earnings profile for a different cohort. If the initial signal is $p^0 = (1/12, 7/12, 4/12)$, then the initial job is still J_1, and $w(p^0) = (5 \times 1/12) + (5 \times 7/12) + (10 \times 4/12) = 6\ 2/3$. In the next period 4/12 of the workers remain in j_1. For those that can either be x or y workers, the optimal second job is J_3. Thus 8/12 of the workers are placed in that job. Of that group, 1/12 are misplaced; they are actually X workers and hence produce 0 in J_2. The remaining 7/12 are still in their optimal job and produce w.

$$\begin{aligned} E\{w(p^1)\} &= (1/12 \times 0) + (7/12 \times 10) + (4/12 \times 10) \\ &= 9\ 1/6 \end{aligned}$$

In the final period, all workers are again correctly placed and $w(p^2) = 10$. The age-earnings profile in this case does increase at a diminishing rate. Notice also that the job progressions are different with the two signals. For the workers that change jobs twice, the progression is from J_1 to J_2 to J_3. In the latter case, it is J_1 to J_3 to J_2.

These two examples indicate that without additional information, the job ladder provides no quantitative restrictions on the age-earnings profiles. Furthermore, even if such restrictions were put on the job structure, the issue would still be unresolved. Demonstrating that a plot of earnings against *jobs* is rising at a diminishing rate is neither necessary nor sufficient for an *age*-earnings profile to have the same shape. The reason is that the data are on *age*-earnings profiles, not *job*-earnings profiles.

By way of illustration, consider the example where $w(p^0) = 6\ 2/3$, $E\{w(p^1)\} = 9/16$, and $E\{w(p^2)\} = 10$. These points are plotted in figure 11.5. Suppose, now, that the length of time between job changes is fixed, either institutionally through sensitivity rules or by the nature of the information structure of the model. In addition, suppose that the length of time for the first move is over three times longer than that for the second. Figure 11.6 illustrates that the associated age-earnings profile is convex.

To derive an eventual leveling off of the age-earnings profile thus requires a theory of the rate at which job performance is generated.

If the bulk of the value of sorting occurred early in the worker's life span and the worker tended to remain in jobs for increasing time periods as the incremental value of sorting diminished, then the age-earnings

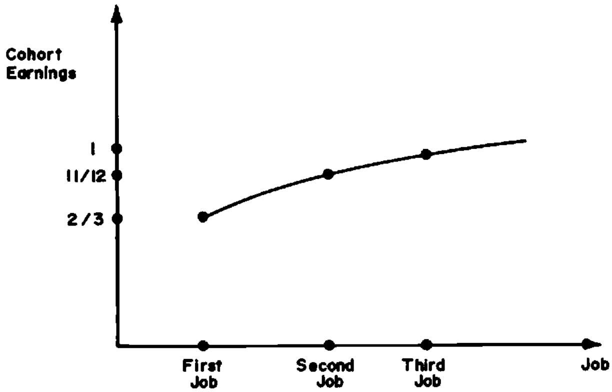

Figure 11.5 Job Earnings Profile

profile would level off. At the beginning, though, the shape of the profile would be somewhat indeterminate. If the job structure is one which has equilibrium sorting within the worker's lifetime, then the age-earnings profile must level off.

The data indicate that a cohort's age-earnings profiles become flat early in the workers' careers and that correlations of earnings with schooling increase with experience for seven to ten years and then level off. These findings are consistent with an equilibrium sorting model view. Indeed, if all were fully sorted, the "increasing variance" proposition would not hold.

11.3 Conclusions

The human capital model has provided explanations of the age profiles of earnings and its variance and correlation coefficients. We have shown

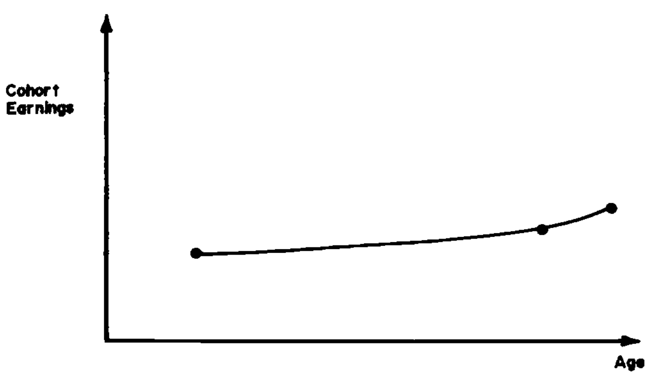

Figure 11.6 Age Earnings Profile

in this paper that a sequential sorting model operating in the presence of uncertainty can also explain all the available empirical evidence. Our explanation is based on the unobserved convexity of the wage function over jobs for workers with expected but uncertain skills. The human capital model explanation is based on unobserved variables with unobserved correlations with measured variables. The two alternative models have different implications for some purposes, and thus it would be useful to devise tests to distinguish them.

Notes

1. Taubman (1977) presents some evidence that the ability correlated with schooling is mostly though not exclusively cognitive. He also presents evidence that noncognitive skills (characteristics) or financing capability that flows from the family explains more of the variance in earnings around age 50 than cognitive skills.

2. We do not include the well-known fact that earnings or wage rates are not normally distributed. This characteristic can be explained by assuming that (unobserved) abilities are not normally distributed.

3. For a recent examination using single cross-section samples, see chapters 3, 4, and 6 in Jencks, et al. (1979). Similar results are found in panel data. See, for example, Fagerlind (1975); Hauser and Daymount (1976); Taubman (1975).

4. As Reder (1969) among others have noted, inability to know in advance a person's marginal product need not invalidate theories which assume that in equilibrium, a person's real wage will equal his marginal product. Reder, for example, suggests that piecework, percentage commissions, and other institutional arrangements can be used to reveal a person's marginal product (MP) before payment is made. Yet there are many occupations and firms where the workers are hired for some relatively lengthy period at a fixed hourly or weekly wage and where a person's MP is not known in advance though perhaps known ex post.

5. Given that the wage structure is open, the worker need not stay with one firm. For some purposes, it is interesting to view each job change as a change in firms. In this sense, the sorting model encompasses external mobility, and it would be a simple matter to append a job-training model that covers internal mobility within the firm. For a discussion of internal labor markets see Williamson, Wachter, and Harris (1975).

Comment John G. Riley

Ross, Taubman, and Wachter (RTW) have provided us with a useful framework within which to analyze on-the-job sorting. They demonstrate convincingly that the stylized facts linking the variance of earnings and time in the work force can be explained purely as a sorting phenomenon. However, RTW also make it clear that their model can explain almost

John G. Riley is Professor of Economics, University of California, Los Angeles.

any earnings-experience profile, so it is hard to visualize how either cross-section or panel data might be used to distinguish their sorting story from the Mincerian hypothesis of different rates of on-the-job investment.

My own feeling is that it would be interesting to combine sorting with aspects of on-the-job training in an attempt to explain observed differences in earnings growth paths sometimes ascribed to "dual labor markets." To illustrate this point, consider figure C11.1, indicating the productivity of different workers in different jobs. On-the-job training is introduced by making productivity in job 3 (J_3) dependent upon whether or not a worker spends an earlier period in job 2.

Suppose a group with identifiable characteristics α is known to be eighty percent type x and twenty percent type y. In a two-period model it is easy to check that members of this group will be placed first in J_1 and then either held in J_1 or advanced to $J_{3\cdot 1}$. Similarly a group with characteristics β which is twenty percent type x and eighty percent type y is optimally placed first in J_2 and then either in J_1 or $J_{3\cdot 2}$. So far this is very much the RTW story. However, suppose in addition that the per capita cost of monitoring performance on the job satisfies $.6 < c < 1 \cdot 2$. For such values of c the expected gains to sorting out the twenty percent of type y in group α are outweighed by the monitoring costs. Then monitoring of this group will not take place, and type y will presumably become "discouraged workers" and end up performing at the same rate as type x in J_1. *Opportunities for advancement are then open only to those groups with sufficiently favorable initial characteristics.*

It is natural, therefore, to ask what characteristics firms will use to identify different groups. Educational achievement is an obvious candidate, so RTW are surely incorrect in describing their sorting hypothesis as an alternative to the signaling hypothesis. Instead, the two hypotheses are complementary.

This brings me to the discussion of signaling in labor markets by Spence. The paper, essentially a minisurvey, provides a nice summary of many of the issues. Particularly interesting is the discussion of imperfect signaling and its distributional implications. However there are two issues whose omission is somewhat surprising.

First, various authors (Rothschild and Stiglitz (1976), Wilson (1977), Riley (1979), and indeed Spence himself) have raised doubts about the viability of signaling or "informational" equilibria. To clarify the issues

$$\begin{array}{cc} & \begin{array}{cccc} J_1 & J_2 & J_{3\cdot 1} & J_{3\cdot 2} \end{array} \\ \begin{array}{c} x \\ y \end{array} & \begin{bmatrix} 3 & 1 & -3 & -3 \\ 4 & 2 & 5 & 8 \end{bmatrix} \end{array}$$

Figure C11.1

involved, I shall consider a simple version of the model described in section 10.3 of Spence's paper. Let $t_n(y)$ be the time required for an individual of type n to achieve educational level y, and let $M_n(y)$ be the lifetime marginal productivity of type n discounted to the time of exit from the educational system. Higher values of n are associated with higher productivity.

The present value of lifetime productivity is then

$$V_n = e^{-rt_n(y)} M_n(y)$$

Taking logarithms we have

$$v_n = \log V_n = \log M_n(y) - rt_n(y)$$

Under the signaling hypothesis, firms offer workers a discounted lifetime income $W(y)$ which is a function of educational achievement y. An individual of type n then chooses y to maximize

$$\bar{v}_n = \ln W(y) - rt_n(y)$$

This is depicted in figure C11.2 for types α and β ($\alpha < \beta$). For simplicity, time in school is shown as a linear function of educational achievement. Given the assumption central to signaling that the marginal (time) cost of education is smaller for the more productive workers, there is an earnings function $W(y)$ such that type β chooses a higher level of y and both end up being paid discounted lifetime earnings equal to lifetime productivity. Spence's original work suggested that such an equilibrium was not unique and indeed that there existed a whole family of these earnings functions. It followed that two subsets of the population with identical unobservable characteristics, but differing in an observable characteristic such as race or sex, might be in quite different equilibria. For example, lifetime earnings at every educational level might be strictly higher for one subset. This possibility generated a rich set of policy implications for affirmative action programs, etc.

However, more recent work suggests strongly that the critical equilibrium issue is not whether there are *many* equilibria but whether there are *any*! Consider again figure C11.2. Firms are initially offering lifetime earnings profiles of $W(y)$. Suppose one firm then makes the alternative depicted offer $<\bar{y}, \bar{W}>$. Both type α and type β are just indifferent between their old best offer and the new alternative. Moreover all those types n with $\alpha < n < \beta$ strictly prefer the new offer. Whether or not such an offer is profitable therefore depends upon whether or not the average lifetime productivity of the types in this interval exceeds $\bar{W}$. Recently it has been shown that under relatively weak conditions expected profits will be positive (see Riley 1979). Therefore the signaling profile $W(y)$

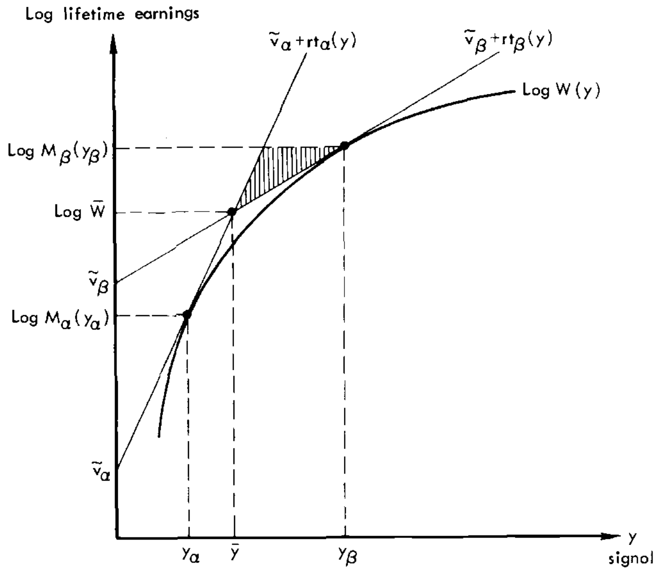

Figure C11.2 Educational Signaling

does not have the stability properties that one would wish of an equilibrium.

However, further reflection suggests that the potential instability described above may not be very damaging. Note that prior to the new offer there is some type $\hat{n}$ for whom

$$M_{\hat{n}}(\bar{y}) = W(\bar{y})$$

Since $\bar{W} > W(\bar{y})$, the new offer loses money on type $\hat{n}$ and hence on all those types n such that $\alpha \leq n \leq \hat{n}$. Therefore the new offer is profitable only on average. It can be shown that there is always a second alternative offer (in the shaded region of figure C11.2) which generates profits to a reacting firm and losses to the firm offering $<\bar{y}, \bar{W}>$. Essentially the reactive offer succeeds in skimming off all the better workers from the pool attracted by $<\bar{y}, \bar{W}>$. Recognition of such an undesirable outcome will then tend to deter firms from making offers above the signaling profile $W(y)$.

While space constraints preclude discussion of the subtleties (see Riley 1979; Wilson 1977), it can be shown that of the family of signaling profiles described by Spence, only one, the Pareto dominating profile is a "reactive equilibrium."

I do not wish to argue that the reactive equilibrium concept provides an entirely satisfactory resolution of the instability problems. However, at the very least it indicates that the signaling hypothesis is not easily rejected on purely theoretical grounds. On the other hand, the elimination of multiple equilibria does eliminate a major difference between the implications of the traditional human capital model and its screening variant.

This brings me finally to the second omission from Spence's paper: the absence of any discussion of the different policy implications associated with the basic signaling model. (As I have already noted, there is an examination of the welfare implications of imperfect signaling.) Accepting the unique reactive signaling equilibrium, I believe that the differences are still very important, especially in the design of programs aimed at improving the education of lower-income groups (see Stiglitz 1975).

The issues are dramatized by considering the simplest case in which there are only two types of workers, type α and type β. With perfect information about productivity, each type chooses a level of education to maximize

$$\ln M_n(y) - rt_n(y)$$

This is depicted in figure C11.3a with type α choosing y_α^* and type β choosing y_β^*. Note that type α would prefer the education-earnings contract of type β. Therefore if, as assumed in the signaling model, productivity is not observable, type β must increase its education level to $\tilde{y}_\beta$ in order to be separated out from the less able. The logarithms of the present value of lifetime income of the two types are then v_α^* and $\tilde{v}_\beta$ with signaling, rather than v_α^* and v_β^*.

Now suppose funds are allocated for research into the improvement of educational achievement for the less able. The broken lines in figures C11.3a illustrate the effect of an educational innovation which increases value added by the less able. In the traditional human capital model, the gains go to this group alone. However, with signaling, the increase in productivity of type α reduces the amount of signaling needed by type β and hence raises lifetime income of the latter group as well. Adoption of such a policy is therefore enlightened self-interest!

A quite different result follows from the adoption of an innovation which increases the rate of educational advancement of the less able. This is depicted in figure C11.3b. The higher rate of educational advancement implies a reduction in the marginal time costs of education $[t'_\alpha(y)]$ and hence an increase in the education of type α. If productivity is directly observed, workers of type β remain at y_β^* and the gains again go only to type α. However, if there is signaling, the flatter cost curve of type α implies that workers of type β must *increase* their education beyond $\tilde{y}_\beta$ in order to be differentiated. This *reduces* their present value of lifetime earnings. To summarize, educational signaling magnifies the potential payoff to increasing value added by the less able and diminishes the payoff to reducing the educational costs of this same group. In both cases the difference is due to first-order spillover effects which alter the income of more able workers.

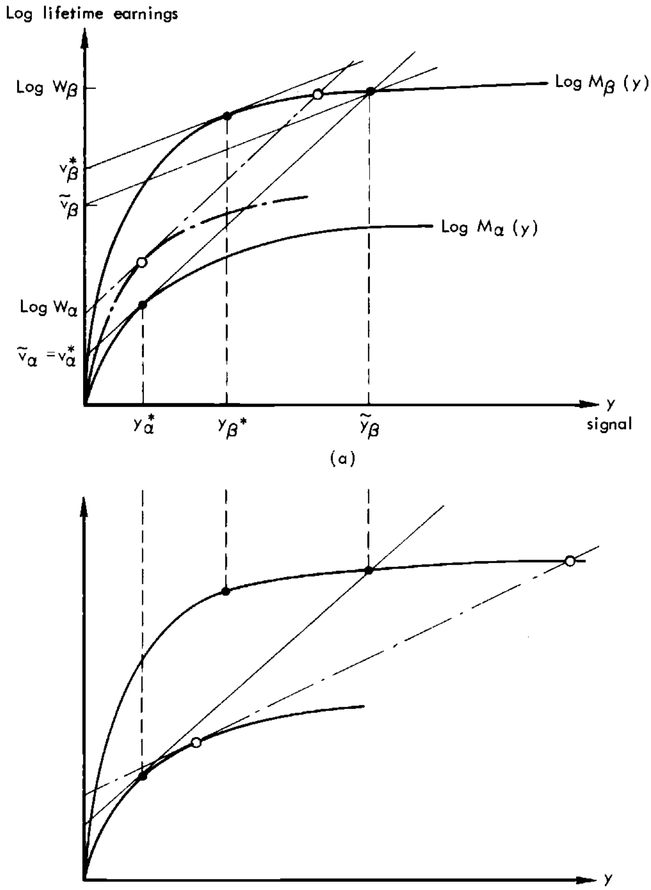

Figure C11.3 Human capital and Signaling Equilibria with Two Types of Workers

Comment Charles Wilson

Among the several topics treated by Spence in his paper on educational signaling is the role of contingent contracts as an alternative to education for screening workers. I will confine my attention to a closer examination of this issue. My general thesis is that the effect of contingent contracts may be very sensitive to the opportunities of the worker to borrow. In a world with perfect capital markets, contingent contracts are in principle

Charles Wilson is Professor of Economics, University of Wisconsin.

efficient substitutes for educational screening. When workers face an imperfect capital market, however, not only may contingent contracts be inefficient, but the same problems with the existence of an equilibrium associated with any self-selection model also appear.

Contingent Contracts with Adverse Selection

Suppose there are two types of workers both of whom work for two periods. Type 1 workers are least productive and generate a marginal value product of s^1 in each period. Type 2 workers are more productive with a marginal value product of $s^2 > s^1$ in each period. Each worker knows his own productivity at the beginning of the first period. However, firms are unable to determine the productivity of a worker until the beginning of the second period. A contingent contract is a first-period wage w_1 and second-period wage w_2^i for $i = 1, 2$ which depends on the productivity of the worker. Therefore any contingent contract can be represented by a three-dimensional vector (w_1, w_2^1, w_2^2).

Consider first the case where contingent contracts are binding on both firms and workers in the second period. Assuming that firms may borrow and lend at a fixed rate of interest r, they will be indifferent to hiring a worker if the present value of his productivity over the two periods equals the present value of his wage payments. Therefore, a firm just breaks even on a type i worker if $w_1 + (1 + r)w_2^i = s^i + (1 + r)s^i$. The firm's "break-even" lines for each type worker, labeled B^iB^i, are illustrated in figure C11.4. Assume for simplicity that $r = 0$; then both have slopes equal to -1 and pass through their respective marginal productivity points (s^i, s^i).

The workers' preferences across different combinations of first and second-period wage rates depends critically on their access to capital markets. Suppose that workers are able to borrow and lend at the same rate of interest as firms. Then independent of their preferences between first and second-period consumption, they are indifferent between any two income streams with the same present value. In this case, therefore, any indifference curve for each worker has the same slope as the break-even line of the firm.

It should be apparent that under these conditions any contract (w_1, w_2^1, w_2^2) for which (w_1, w_2^1) lies on the B_1B_1 line and (w_1, w_2^2) lies on the B_2B_2 line is consistent with equilibrium in a competitive market. Firms are willing to pay a wage in excess of the marginal value product in the first period if the worker will accept a correspondingly lower wage in the second. Likewise, workers will accept a lower wage in the first period if they are appropriately compensated in the second period.

Some of this indeterminacy disappears if we relax the assumption that workers may borrow at the same fixed rate of interest as firms. Suppose a worker may lend at the fixed rate r but faces a marginal interest rate schedule which increases with the amount he borrows. In this case, his

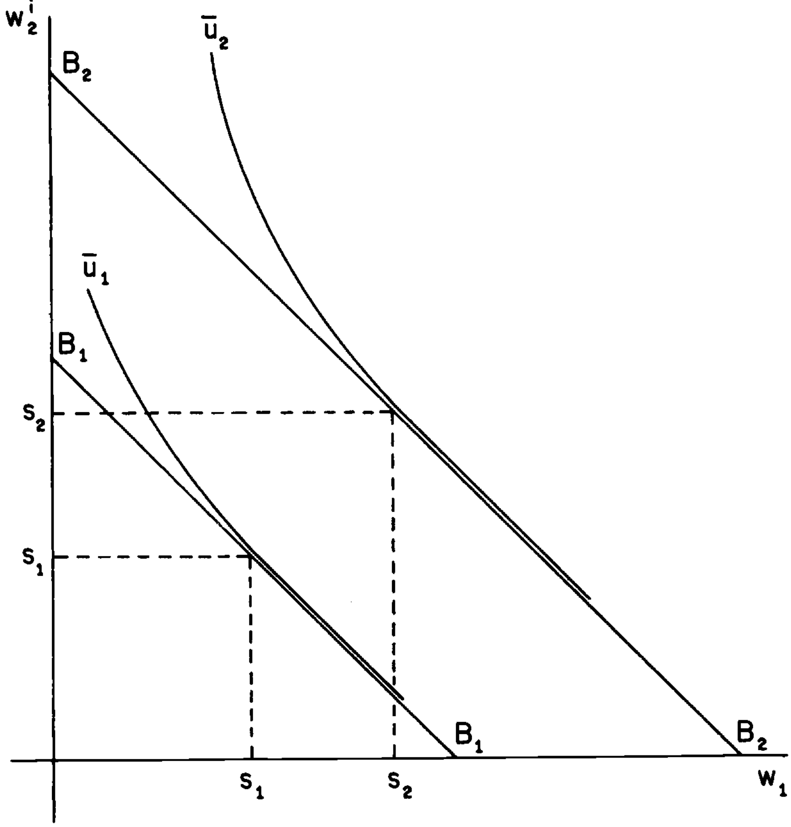

Figure C11.4

feasible bundles of first and second-period consumption depend on more than just the present value of his wages evaluated at interest rate r. Assuming his marginal rate of substitution between first and second-period consumption is strictly decreasing, his marginal rate of substitution between first and second-period *wage rates* will also be strictly decreasing at any combination of wage rates at which he chooses to borrow in the first period. Typical indifference curves are illustrated in figure C11.4. As we increase w_1, the slope becomes increasingly flatter (reflecting a lower marginal interest rate), until a combination of wage rates is reached at which the worker no longer chooses to borrow. Thereafter, the curve becomes a straight line parallel to the firm's break-even line.

Under these conditions, it is no longer true that an equilibrium can be attained given any first-period wage. Because firms may borrow at a lower interest rate than workers, any contract which induces workers to borrow in the first period presents obvious arbitrage opportunities to firms. Competition then forces firms to "lend" to workers at the market rate of interest by increasing the first-period wage to a point at least as large as the workers' desired level of first-period consumption. The equilibrium wage contracts can be illustrated in figure C11.4. For type 1 workers, any contract on their break-even line to the right of s^1 is an equilibrium; for type 2 workers, contracts to the right of s^2 are equilibrium contracts.

In general, it is difficult to enforce the terms of a contingent contract in the second period if the worker can command a higher wage elsewhere. Therefore, let us assume henceforth that the terms of the contract in the second period are binding only on firms. In the second period, workers are free to change employers in order to obtain a higher wage.

Consider first how this affects the worker's preferences among different contracts. As long as the second-period wage is greater than the worker's marginal value product, the worker has no incentive to leave the firm. For these contracts, therefore, the worker's indifference map remains unchanged. However, once the second-period wage falls below the worker's marginal value product, its level becomes irrelevant. The worker can guarantee himself a higher second-period wage by changing (or threatening to change) employers. Therefore, the typical indifference curve for a type i-worker becomes truncated at $w_2^i = s^i$. Once the second-period wage falls below s^i, the worker prefers any contract with a higher first-period wage.

The break-even lines for the firms are also affected. Since the firm must pay a type i worker at least s^i in the second period or lose him to another employer, it can never break even on a type i worker if it pays him more than s^i in the first period. In particular, firms must lose money on type 1 workers if $w_1 > s^1$. Nevertheless, the firm may still break even on average when type 1 workers choose a contract with $w_1 > s^1$, if type 2 workers also accept the contract with a second-period wage low enough to compensate the firm for its loss on the less-productive workers. The only constraint is that the second-period wage for type 2 workers exceed s^2; otherwise, they too will leave the firm in the second period. Let a_i be the proportion of type i workers. Then if the firm is to break even on average when $w_1 > s^1$, w_2^2 must satisfy: (a) $w_2^2 > s^2$; (b) $(a_1 s^1 + a_2 s^2 - w_1) + a_2(s^2 - w_2) = 0$.

This line is labeled $B_a B_a$ in figure C11.5. It starts at $(s^1, 2s^2 - s^1)$ and declines with slope $-1/a_2$ until $w_2^2 = s^2$, at which point the line becomes vertical. The lines labeled $B_1' B_1'$ and $B_2' B_2'$ are the break-even lines for each type individually. Each is identical to the corresponding $B_i B_i$ line up to $w_i = s^i$, at which point it also becomes vertical.

Now consider the equilibrium for this market under the assumption that workers have access to perfect capital markets. In this case, any

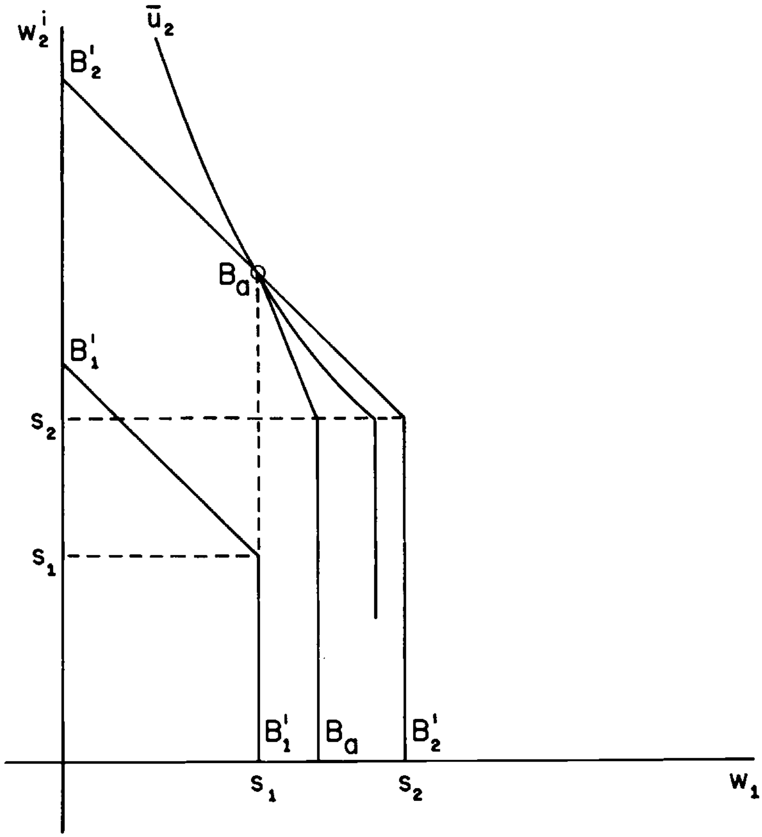

Figure C11.5

first-period wage less than or equal to s^1 is consistent with equilibrium. Both w_2^1 and w_2^2 can be adjusted so that each worker obtains a contract on his break-even line. Furthermore, because the indifference curve of each worker through these points is coincident with his corresponding break-even line, there is no other contract which is profitable for the firm and preferred by this type of worker. Note that because the firm breaks even on both types individually, it is not even necessary for both types to earn the same first-period wage.

The requirement that w_1 be less than s^1 is essential, however. Otherwise, type 1 workers will choose that contract with the highest first-period wage. But in order for such a contract to break even, it must also attract type 2 workers to the corresponding contract on the B_aB_a line. Since any such contract is less preferred than contracts with $w_1 \leqq s_1$ on the $B_2'B_2'$ line, no type 2 work will accept it. Therefore $w_1 > s^1$ cannot be an equilibrium.

In short, the restriction on feasible first-period wages resulting from the inability to enforce contingent contracts on workers does not present any serious problems if workers have access to the same capital markets as firms. They are willing to accept any first-period wage if the second-period wage is high enough to generate an income stream with a present value equal to the present value of their marginal product.

This conclusion changes in a fundamental way, however, when we reintroduce the possibility that workers face an upward-sloping marginal interest rate schedule. Suppose that type 2 workers will choose to borrow at any contract on the $B_2'B_2'$ line with $w_1 \leqq s^1$. In this case the indifference curve for the worker will have a slope which is steeper than the $B_1'B_1'$ line at that point. However, in order to obtain a contract with a higher first-period wage, the worker must be willing to subsidize the type 1 workers who will also choose the new contract.

If the slope of the type 2 indifference curve is less in absolute value than $1/a_2$, type 2 workers will prefer to remain at a contract with $w_1 = s^1$, as illustrated in figure C11.5. Consequently, the equilibrium looks no different than when workers could borrow at interest rate r; however, it will be less efficient. There are two distinct problems. First, if there were no type 1 workers, the free rider problem associated with higher first-period wages would disappear and type 2 workers could obtain any contract on the $B_2'B_2'$ line yielding a higher level of satisfaction. Second, if either type workers' preferred consumption point on their B_iB_i line requires a first-period wage greater than s^i, another source of inefficiency results because the firm no longer is able to make loans to its workers in the first period by increasing w_1. Any higher first-period wage becomes essentially a transfer payment when the worker leaves the firm in the following period.

If the slope of the type 2 indifference curve is greater in absolute value than $1/a_2$ where $w_1 = s^1$, then type 2 workers will strictly prefer a contract on the B_aB_a line with $w_1 > s^1$, such as point c in figure C11.6. Firms who offer this contract attract both types of workers. Type 1 workers leave the firm at the end of the first period, but the second-period wage to type 2 workers is sufficiently low so that the firm breaks even on the average worker.

Is this contract then an equilibrium? It is *not* a Nash equilibrium. Suppose some firms are offering contract c and attracting both types of workers. Another firm could offer a contract such as (d) with a lower first-period wage and a type 2 second-period wage sufficiently higher to attract the type 2 workers but still low enough more than to break even on such workers. Note that this contract will not attract type 1 workers because their second-period wage would be no higher than s^1. They would be sacrificing a higher first-period wage without receiving a higher second-period wage in return.

This is precisely the same problem with the existence of equilibrium as was discovered by Rothschild and Stiglitz (1970) and myself (Wilson

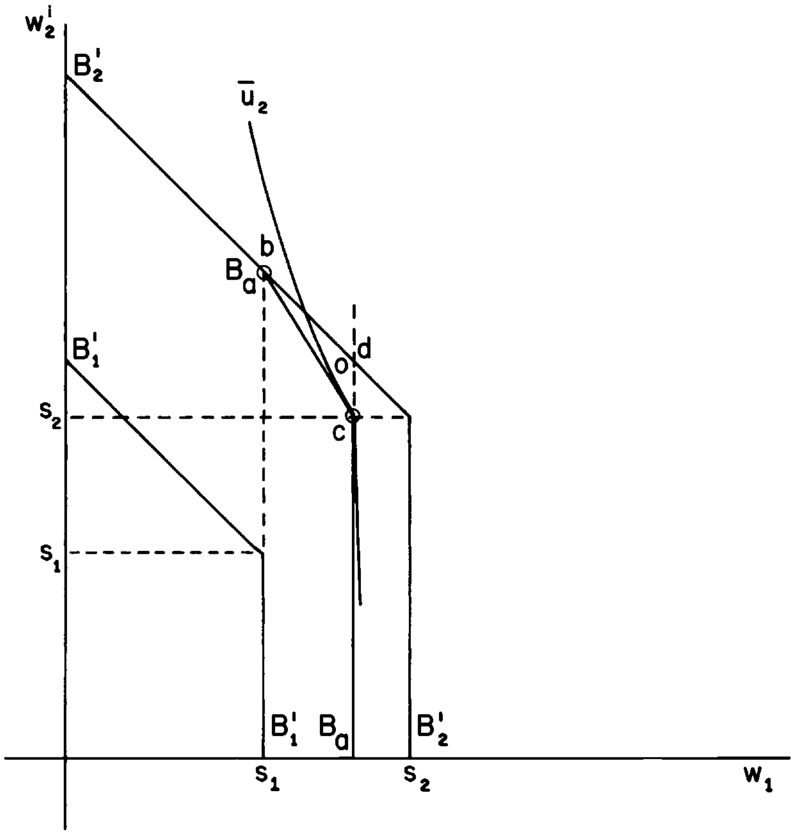

Figure C11.6

1977) in the context of an insurance market. It can appear in any model with signaling or self-selection. If one adopts the equilibrium concept that is employed in Wilson (1977), then point *c* does become an equilibrium. Firms do not offer a contract like (*d*) because they anticipate that firms offering (*c*) will be left with only type 1 workers and consequently will drop the contract. Type 1 workers will then move to (*d*) and it will lose money as well. On the other hand, if one adopts the reactive equilibrium concept suggested by Riley (1979), then contract *b* is the equilibrium. Firms will not offer a contract like (c) because they fear retaliation by other firms who may offer contract *d*.

Little will be gained by discussing in any more detail what is the appropriate equilibrium concept for this market. The issue has already been examined at some length elsewhere. However, a few words about the feasibility of contingent contracts are in order. Recall that the original issue was whether or not contingent contracts can replace signaling as a screen for productive workers. Throughout the analysis the implicit assumption was that such contracts have essentially zero enforcement costs on firms. As a practical matter this may be a difficult assumption to justify.

As I see it, the problem is not so much that firms have an incentive to break the contract. The short-run benefits of breaking a contract will be more than offset by the long-run cost to the firm resulting from its loss of credibility. The problem is in *verifying* that the firm is in fact fulfilling the contract. It is not sufficient that the firm actually pay high-productivity workers a higher wage in the second period; they must be able to convince new workers that they are actually following this policy. An obvious solution to this problem is the use of credentials, either formal or informal. In order to receive a higher wage in the second period, the worker must satisfy certain public criteria. But this "solution" may not be without its own inefficiencies. In fact, we may have essentially reintroduced signaling into the second-period wage decision. Insofar as workers overinvest in credentials which certify their productivity (in the academic market, they may publish too many papers or attend too many professional meetings), the solution may be no more efficient than if education were used as a signal in the first place.

In a more complete model, I suspect that education before the first period would simply supplement other "credentials" the worker must acquire in order to receive a higher wage in the second period. Thus, we have come full circle. In searching for contracts which avoid the inefficiency of educational signaling, firms may require signaling in other forms and in fact may even require educational signaling to enforce the contracts.

The Ross-Taubman-Wachter (RTW) paper presents a convincing and elegant explanation of many of the properties of the typical age-earnings profile. They focus exclusively on the implications for the distribution of earnings when firms optimally assign their employees to jobs based on the workers' performances at earlier jobs. I will confine my comments to two points. The first is that the argument may be strengthened if one takes into account the incentives for intertemporal maximization of a worker's output. The second is that when contingent contracts cannot be introduced efficiently, the problem of adverse selection may tend to generate some inefficiency in the assignment of workers.

RTW argue that depending on the distribution of the worker's productivity and the types of jobs available, incomplete sorting may result. This has the effect of flattening the experience-earnings profile. Although I believe their point is essentially correct, the bias toward incomplete sorting is less severe if firms and workers consider the future benefits of less productive jobs at the beginning of worker careers, followed by a more effective sort later on.

Consider their example with two types of workers and two jobs given by the matrix in figure 11.3. If each worker has a .5 probability of being an X or y worker, the expected payoff from assigning a worker to job 1 is 2; for job 2 the expected payoff is 1.5. Therefore all workers will be assigned to job 1. From this example RTW conclude that complete sorting may not occur even though total output could be increased if type X workers could be identified and assigned to job 2.

This conclusion changes, however, if we consider the implications for intertemporal maximization of a worker's output. Suppose each worker works three periods and the discount rate is zero. If all workers are assigned to job 1 in each period, then total output is 6 per worker. But firms can do better than this. If they assign each worker to job 2 in period 1, the average return is 1.5; however, this permits them to identify each worker's type. In the next two periods, therefore, type X workers can be assigned to job 2 and type y workers assigned to job 1, yielding an average output of 5 which, added to 1.5 in the first period, gives a total of 6.5. Assuming firms pay workers their *expected* marginal product in each period, workers will choose to work for a wage of 1.5 in the first period for a chance to obtain a higher wage in later periods.

Now suppose that workers know their productivity before they take their first job. If contingent contracts can be enforced, then type X workers will immediately choose job 2 and type y workers job 1. In the absence of contingent contracts, however, some inefficiencies appear. In each period, the workers in each job are paid their expected marginal product. Type X workers will immediately go to job 2 in order to establish their productivity. Type y workers will go to the job which pays the highest wage. This will be job 2 unless some type y workers take that job lowering its expected marginal product to 2. Consequently one-third of the type y workers will also take job 2 in period 1. In the following period all workers are perfectly sorted. The firm does not achieve a first best optimum, but does do better than it would if workers had no information at all.

This result need not be obtained in general. Suppose that in job 1, type X workers produced 3 units and type y workers 2, but individual output could not be distinguished. Output would be maximized by leaving all workers in job 1. But, the equilibrium with adverse selection would remain unchanged with one-third of the type y workers assigned to job 2 in period 1.

References

Arrow, K. "Higher Education as a Filter." *Journal of Public Economics*, July 1973.

Diamond, P., et. al. Appendixes in Report to Congressional Research Service by Consultant Panel on Social Security, August 1976.

Fagerlind, I. *Formal Education and Adult Earnings*. Almqvist and Wicksell, 1975.

Fardoust, S. "Risk Taking Behavior, Socioeconomic Background, and Distribution of Income: A Theoretical and Empirical Analysis." Ph.D. Thesis, University of Pennsylvania, 1978.

Hauser, W., and Daymont, T. "Schooling, Ability, and Earnings: Cross Sectional Findings 8 to 14 Years After High School Graduation." Center for Demography and Ecology, University of Wisconsin, 1976.

Jencks, C. "Who Gets Ahead." Basic Books, 1979.

Levhari, D., and Weiss, Y. "The Effect of Risk on the Investment in Human Capital." *American Economic Review*, December 1974, pp. 950–63.

Lilliard, L., and Willis, R. "Dynamic Aspects of Earnings Mobility." *Econometrica*, forthcoming.

Mincer, J. *Schooling, Experience and Earnings*. Columbia University Press, 1974.

Reder, M. "A Partial Survey of the Theory of Income Size Distribution," in L. Soltow, ed., *Six Papers on Size Distribution of Income and Wealth*. Columbia University Press, 1969.

Riley, J. "Testing the Educational Screening Hypothesis." Rand Corporation, mimeo., 1978.

______. "Informational Equilibrium." *Econometrica* 47 (1979): 331–60.

Rothschild, M., and Stiglitz, J. "Equilibrium in Competitive Insurance Markets: The Economics of Imperfect Information." *Quarterly Journal of Economics* 90 (1976): 629–49.

Spence, M. *Market Signalling: Informational Transfer in Hiring and Related Processes*. Harvard University Press, 1974.

______. "Competition in Salaries and Signalling Prerequisites for Jobs." *Quarterly Journal of Economics* 90 (1976): 51–75.

Stiglitz, J. E. "The Theory of Screening, Education and the Distribution of Income." *American Economic Review* 65 (1975): 283–300.

Taubman, P. *Sources of Inequality of Earnings*. North-Holland, 1975.

______. "The Relative Influence of Inheritable and Environmental Factors and the Importance of Intelligence in Earnings Functions." IEA Conference Paper, Nordwijk, 1977.

Taubman, P., and Wales, T. "Higher Education, Mental Ability, and Screening." *Journal of Political Economy*, January/February 1973.

Thaler, R., and Rosen, S. "Estimating the Value of a Life," in Terleckyj, ed., *Household Production and Consumption*. Columbia University Press, 1976.

Weiss, Y. "The Risk Element in Occupational and Educational Choices." *Journal of Political Economy*, November/December 1972, pp. 1203–13.

Williamson, O.; Wachter, M. L.; and Harris, J. "Understanding the Employment Relation: An Analysis of Idiosyncratic Exchange." *Bell Journal of Economics*, Spring 1975, pp. 250–78.

Wilson, C. "A Model of Insurance Markets with Incomplete Information." *Journal of Economic Theory* 16 (1977): 167–207.