Lecture Note 20 — Education, Human Capital, and Labor Market Signaling

David Autor, MIT and NBER

Introduction

It's often said that information is a public good: easy access to weather forecasts, traffic information, stock prices, etc. helps us out every day. This would suggest that (a) the more information the better—after all, in the worst case, we can ignore it; (b) *too little* information will be provided in a free market equilibrium, since public goods are undersupplied by the market. The 1973 paper by Michael Spence "Job Market Signaling" demonstrates that, in some cases, there can actually be *too much* information in a free market equilibrium. Why? It's not that disclosing information is per se harmful. But the social value of the information disclosed may not be worth the cost of conveying it. And yet, people may have an incentive to produce and disclose information that generates non-zero private benefits and yet generates zero (or negative) net social value. In the signaling model in Spence's 1973 paper, the incentives for disclosure or non-disclosure are purely private, and these private incentives may or may not generate desirable outcomes measured in terms of social efficiency.

A bit of background on the market for information

- Economists had historically conjectured that markets for information were well-behaved, just like markets for other goods and services. One could optimally decide how much information to buy, and hence equate the marginal returns to information purchases with the marginal returns to all other goods.

- In the 1970s, economists were given cause to reevaluate this belief by a series of papers by George Akerlof, Micheal Rothschild, Joseph Stiglitz, and Michael Spence. Many (not all) of these economists went on to share the 2001 Nobel for their work on the economics of information.

- Information is not a standard market good:

  - It is non-rivalrous, since there is no marginal cost to an additional person having that information.

- It is extremely durable, since it doesn't vanish once it's consumed.
- It not a typical experience good that you can 'try before you buy.' A seller cannot readily allow you to 'sample' information without actually giving you the information.
- Unlike other goods (or their attributes), information is extremely difficult to measure, observe, and verify.

- Another important property of information is that is may be *asymmetric*. Some agents in a market can be better informed than others about the attributes of a product or transaction. This feature of information is distinct from most other properties of goods and services. In conventional exchanges, there is no uncertainty about the nature of the good or service in question; the only market-relevant characteristics are quantity and price. Where information is involved, however, different parties to a transaction may not have the same information about the attributes of the good. This asymmetry of information (and knowledge of this asymmetry) will often affect the price and quantity at which parties are willing to trade.
- (Note: Economists will often use the terms *asymmetric information* and *private information* interchangeably. Both terms mean that one actor has different—generally *more*—information than another actor, that both actors typically recognize this asymmetry, and that they may behave strategically as a consequence.)
- The most natural (and surely ubiquitous) way in which this occurs is that buyers may have general information about the average characteristics of a product that they wish to purchase, whereas sellers will have *specific* information about the individual products that they are selling.
- When buyers and sellers have asymmetric information about market transactions, the trades that actually occur are likely to be a subset of the feasible, welfare-improving trades. Sellers will want to sell certain items whose attributes they know, and buyers may be cautious if they understand the incentives of the seller. Many trades that would voluntarily occur *if all parties had full information* will not take place.
- Economic models of information focus on the information environment—that is, who knows what when. Specifying these features carefully in the model is critical to understanding what follows.
- The next few lectures will dive into multiple phenomenon that arise when information is asymmetric. This lecture note discusses our first key insight from the literature on asymmetric information: that is possible for agents to engage in socially unproductive signaling in a free market equilibrium.

1 Context: Educational investment

- Education is perhaps the most significant investment decision you will make.

- Most citizens of developed countries spend $12 - 20$ years of their lives in school. The provision of schooling involves two types of costs:
  - Direct costs: Buildings, teachers, textbooks, etc. (In 2013, the United States spent 4.8 percent of its gross domestic product on education, see link.)
  - Indirect costs: Opportunity costs of attending school instead of working or having fun. These costs surely swamp the direct costs of schooling.
- Is this enormous investment a socially efficient use of resources? Or is it the equilibrium outcome of a process that doesn't necessarily maximize social welfare?
- Economists has historically used the Human Capital Model created by Gary S. Becker (1964) and extended by Jacob Mincer. This model views education as an up-front investment that increases future productivity, and suggests that the investment is well worth the costs to society.
- Spence suggested a second model: the signaling model. This model suggests that individuals who spend decades in school may be signaling that they are productive rather than actually becoming more productive. This model suggests very different conclusions about the optimal size and form of our educational institutions.
- We'll compare and contrast these models formally.

2 A simplified human capital investment model: The 'equalizing differences' model of Jacob Mincer

- Define $w(s)$ as the wage of someone with s years of schooling.
- Assume $w'(s) > 0$: productivity rises with schooling. We will also model the labor market as competitive, so that earnings reflect productivity (and also rise with schooling).
- Assume that the direct costs of schooling, c, are zero for now. (Of course they aren't zero in the real world, but we won't need them to illustrate the main lessons of this model.)
- Define $r > 0$ as the interest rate. If I decided not to put \$1 into schooling and instead put that \$1 in the bank, I would get $\$(1 + r)$ next period.
- For simplicity, assume people live forever. (a lifespan of40 years and a lifespan of infinity give very similar results in models with time discounting.)
- In this model, what is the present value of the lifetime earnings of someone with 1 year of schooling? It is the Discounted Present Value (DPV) of receiving $w(1)$ annually in each

subsequent year:

$$DPV\left[w(1)\right] = w(1) + \frac{w(1)}{1+r} + \frac{w(1)}{(1+r)^2} + ... + \frac{w(1)}{(1+r)^{\infty}},$$

which can be simplified as follows:

$$\begin{aligned} DPV\left[w(1)\right] \cdot \left(\frac{1}{1+r}\right) &= \frac{w(1)}{1+r} + \frac{w(1)}{(1+r)^2} + \frac{w(1)}{(1+r)^3} + ... + \frac{w(1)}{(1+r)^{\infty}}. \\ DPV\left[w(1)\right] - DPV\left[w(1)\right] \cdot \left(\frac{1}{1+r}\right) &= w\left(1\right), \\ DPV\left[w(1)\right] &= w(1)\left[\frac{1}{1-\frac{1}{1+r}}\right], \\ DPV\left[w(1)\right] &= w(1)\left(\frac{1+r}{r}\right) \end{aligned}$$

- Note however that if you attend school for one year to obtain $w\left(1\right)$, you do not receive the first payment of $w(1)$ until after your first year of schooling is complete. So the DPV of one year of education is $w(1)\left(\frac{1+r}{r}\right)\left(\frac{1}{1+r}\right) = \frac{w(1)}{r}$.
- Conversely, a person who does not attend an additional year of schooling receives:

$$DPV\left[w(0)\right] = w(0) + \frac{w(0)}{1+r} + \frac{w(0)}{(1+r)^2} + ... + \frac{w(0)}{(1+r)^{\infty}} = w(0)\left(\frac{1+r}{r}\right).$$

- So, the DPV net benefit of obtaining one additional year of schooling is:

$$DPV[\text{Going to School for One Year}] = w(1)\frac{1}{r} - w(0)\left(\frac{1+r}{r}\right)$$

- Now take as given:
  - A competitive market for labor.
  - Perfect capital markets (can always borrow the full cost of schooling).
  - Rational, identical individuals, each with same earnings potential.
- In equilibrium, it must be the case that the costs and benefits of an additional year of schooling are equated. (If the costs were lower than the benefits, no one would get schooling. If the costs were greater, everyone would get schooling. So, the equilibrium must have everyone indifferent.)

- This implies that

$$\begin{aligned} w(1)\frac{1}{r} &= w(0)\left(\frac{1+r}{r}\right), \\ \frac{w(1)}{w(0)} &= (1+r), \\ \ln w(1) - \ln w(0) &= \ln(1+r) \approx r. \end{aligned}$$

- In words, the wage increment for one more year of schooling must be approximately equal to the interest rate! If this were not true, then people would change their investments in schooling versus other opportunities to equalize these marginal returns. (Note, this log approximation holds for small values of r, say $r < 0.2$. Above that, the approximation gets very close.)

- Simple as this model is, it does a pretty good job at capturing a remarkable empirical regularity. Over the last 101 years (which is as long as we can measure for the U.S.), the estimated rate of return to a year of schooling has been about 5 to 10 percent—approximately equal to the real rate of interest plus inflation.

2.1 Mincer's equalizing differences model of human capital investment has four testable implications:

1. People who attend additional years of schooling are more productive.

2. People who attend additional years of schooling receive higher wages.

3. People will attend school while they are young, i.e., before they enter the workforce. [Why? Because the costs of school are the same whenever you attend it, but the benefits do not begin to accrue until you have completed it. You should therefore get your education before you start working.]

4. The rate of return to schooling should be roughly equal to the rate of interest.

3 The Spence signaling model of educational investment

- The Mincer model presupposes that education increases productivity. But if education were unproductive, would any of the above still be true?

- Prior to Spence's 1973 paper, most economists would have said "no" reflexively. If education is unproductive, why would people spend time acquiring it? And why would employers pay higher wages to educated workers?

- The surprise of the Spence model is that even if education is unproductive, there may be employee and employer demand for it *in equilibrium*. The main lesson of Spence's model is that people may use education to signal that they are productive, rather than to make themselves more productive.

3.1 Setup

- Consider the following stylized model:

  1. People are of heterogeneous "high" or "low" ability: H, L.
  2. High ability people are inherently more productive than low ability people. There is no way for a high ability person to become a low ability person, or vice versa.
  3. An individual's ability is known to him or her, but not to potential employers.
  4. Education does not affect ability/productivity.
  5. High ability people have lower cost of attending school than others. (Why would this be so? Lower psychic costs to studying topics you're naturally good at, subsidies to education are greater for high ability people via merit scholarships, etc.)

- Let's use these parameter values to illustrate this model:

Group	Productivity	Population Share	Cost of Education
L	$Y_L = 1$	λ	S
H	$Y_H = 2$	$1 - \lambda$	$\frac{1}{2}S$

- So average productivity of the population is $2(1 - \lambda) + 1\lambda = 2 - \lambda$.
- Notice that a worker's productivity does *not* depend on how much school she obtains.
- What are the possible equilibria of this model—specifically, what wages should employers offer to workers with different levels of schooling S, and how much schooling should H and L workers obtain?

3.2 Equilibrium in models with asymmetric information

One feature that makes models with asymmetric information somewhat different from models we've studied previously is that the equilibrium is *not* primarily determined by a set of marginal conditions (e.g., marginal profit is zero, or the marginal rate of substitution equals the price ratio). Instead, equilibrium depends on finding a set of compatible strategies, in which each player's behavior makes sense given the other players' behaviors. We think of parties on the different sides of the market (e.g., buyers v. sellers) as choosing strategies (feasible actions) that maximize their payoffs given the chosen strategies of the players on the other side of the market. But of course, the players on the other side of the market are likewise choosing strategies to maximize their payoffs given the actions (or anticipated actions) of the other players.

An equilibrium in this setting is a set of complementary strategies such that neither side wants to unilaterally change its strategy given the strategy of the other side. This notion is what is called a Nash Equilibrium after John Forbes Nash, who developed the idea and proved its existence in a 28 page Princeton doctoral dissertation in 1950. This dissertation eventually won him the Nobel prize in Economics in 1994 and, more importantly, led to the hit Hollywood biographical movie *A Beautiful Mind* in 2002 in which Russell Crowe played John Nash. As far as we know, this has been Russell Crowe's most important contribution to economic theory.

To further solidify your understanding: here's an informal definition of the Nash Equilibrium (paraphrased from Wikipedia): *A Nash Equilibrium is a solution concept for a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only his or her own strategy unilaterally. If each player has chosen a strategy and no player can benefit by changing his or her strategy while the other players keep theirs unchanged, then the current set of strategy choices and the corresponding payoffs constitute a Nash equilibrium.*[1]

We will invoke this idea below.

3.3 Separating equilibrium

- In a model with two 'types' of students (here H and L), there are two potential classes of pure strategy equilibria: a *separating equilibrium* where the two types follow different strategies (i.e., pursue different levels of education) and realize different outcomes; and a *pooling equilibrium* where both types follow the same strategy (i.e., pursue the same level of education) and realize the same outcomes. (There may also be *mixed strategy equilibria* where agents randomize among actions so as to make other players play compatible strategies. We will not consider mixed strategy equilibria here.) We will first consider the possibility of separating equilibria.

- Assume that firms offer the following wage schedule:

$$w(S) = 1 + I[S \geq S^*], \tag{1}$$

  where $I[\cdot]$ is the indicator function. A worker with $S_i \geq S^*$ years of education is paid 2 and otherwise 1. This might make sense from the firms' perspective if they believe that only H type workers are willing to go through S^* years of schooling.

- How much education will workers obtain? The worker's problem is

$$\max_S w(S) - c(S).$$

- For a type H worker, the cost of S^* years of education is $0.5S^*$ and the wage benefit is 1. So type H workers will attend school for S^* years if: $w\left(S \geq S^*\right) - 0.5S^* > w\left(S < S^*\right) \Rightarrow 2 - 0.5S^* > 1$.

- For a type L worker, the cost of S^* years of education is S^* and the wage benefit is 1, so L type workers will attend school if: $w\left(S \geq S^*\right) - S^* > w\left(S < S^*\right) \Rightarrow 2 - 1 > 1$.

[1] And here's a plot summary of the movie from IMDB.com, "After a brilliant but asocial mathematician accepts secret work in cryptography, his life takes a turn for the nightmarish."

- Consider if the employer sets $S^* = 1 + \varepsilon$, where ε is a very small, positive number. Give this wage policy, H workers will strictly prefer obtaining S^* years of education (since $2 - 0.5(1+\epsilon) > 1$) whereas L workers will not find it worthwhile to obtain education S^* (since $2 - (1+\epsilon) < 1$).

- Now let's turn to the firm side. Is the employer's wage schedule, represented by (1), an equilibrium wage schedule? Or would employers want to construct a different wage schedule given workers' behavior?

- For firms to offer this wage schedule, it must be the case that

$$E\left[Y\left(S\right)|w\left(S\right)\right] \geq w\left(S\right),$$

that is, the expected productivity of workers qualifying for a wage level based on their schooling must be at least equal to the wage. Here, the function $Y\left(S\right)$ gives the productivity of workers supplying labor with schooling level S. We are conditioning $Y\left(S\right)$ on the *function* $w\left(S\right)$ because employers offer a wage *schedule* rather than a single wage. The worker's choice of S_i therefore depends on the wage schedule $w\left(S\right)$, and so productivity $Y(S)$ changes with the wage schedule $w(S)$ as different types of workers select into different amounts of schooling S.

- In our example, workers of type L are not willing to go through $1 + \epsilon$ years of schooling, so they might as well pick $S = 0$. These workers have productivity 1 and receive wage 1, so the employer's wage schedule is rational for these workers ($E\left[Y\left(0\right)|w\left(S\right)\right] = w\left(0\right) = 1$).

- Similarly, in this example, workers of type H would pick schooling $S = 1 + \epsilon$. They have productivity 2 and wage 2, and so the employer's wage schedule is also rational for these workers: $E\left[Y\left(1\right)|w\left(S\right)\right] = w\left(1\right) = 2$

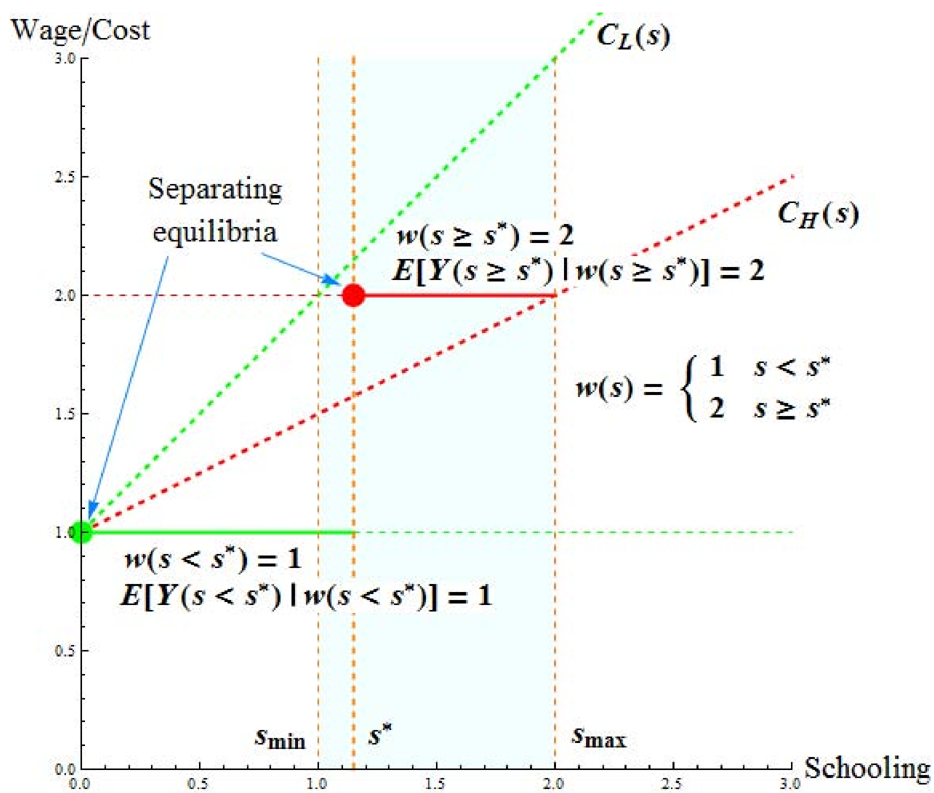
- So, this is an equilibrium: high ability workers will obtain $S = 1 + \epsilon$ education, low worker will obtain $S = 0$ education. Employers can perfectly determine the type of each worker based on his or her level of schooling and set the wage schedule accordingly, so the firm is fine with this arrangement. (Note that the labor market is assumed to be perfectly competitive, so there's no deviation that can give the firm strictly positive profits.) Neither H workers, L workers, nor employers will have incentive to deviate from the pay scheme.

- You may find it helpful to draw the indifference curves of both worker types in (w, S) space, as in Figures 1, 2 and 3 below.[2] Observe that in these figures, the cost curves $C_L\left(S\right)$ and $C_H\left(S\right)$ serve as indifference curves: workers are willing to go through a certain amount of schooling only if they are compensated with a higher wage. The indifference curves originate from the initial wage offer for $w\left(S = 0\right)$, and they slope upward with the worker's cost of education. Along these cost curves, workers of each type are indifferent among all bundles on their respective cost curves offering higher wages and higher schooling relative to $w\left(S = 0\right)$. Workers strictly prefer to be *above* (northwest of) these curves relative to their default bundle

[2] I thank Sergey Naumov for constructing these illuminating figures.

of $w(S=0)$. Workers strictly prefer to *not* be below (southeast of) these curves, i.e., worse off than at their default bundle of $w(S=0)$.

Figure 1: Potential Separating Equilibria with $\lambda = 0.5$



- Notice that the separating equilibrium requires that the wage schedule induce self-selection: high-productivity workers choose to obtain two years of schooling and low-productivity workers choose to obtain only one. In equilibrium, employers are happy to pay workers with S^* years of schooling a wage of 2 and workers with less than S^* year of schooling a wage of 1, and neither workers nor employers have an incentive to deviate from this equilibrium.

- The unfortunate aspect of this model is that education is completely unproductive, so these investments are socially wasteful. By obtaining education, H type workers 'signal' that they deserve a high wage—but this is a pure private benefit. From a social perspective, this signaling does nothing useful since it does not increase total output.

- Does it matter for this model whether employers believe that education is productive? Actually, it does not. So long as people who have schooling $S \geq S^*$ have productivity 2 and those who have schooling $S < S^*$ have productivity 1, employers have no incentive to deviate from the wage schedule. [Consider an experiment where employers were told that education is unproductive. Would they want to change their wage schedules?]

3.4 Pooling equilibria with positive education

- The example above is a separating equilibrium: L and H types obtain different levels of education in equilibrium. There are also a multiplicity of possible *pooling* equilibria, that is, cases where L and H types receive identical education.

- Imagine that employers offered a wage schedule of

$$w(S) = 1 + I[S \geq S^*] \cdot (1 - \lambda),$$

so workers with education less than S^* receive a wage of 1 and those with education greater than or equal to S^* receive wage of $2 - \lambda$. Who would invest in education?

- H types will acquire $S = S^*$ at cost $0.5S^*$ if

$$2 - \lambda - 0.5S^* > 1 \Rightarrow S^* < 2\left(1 - \lambda\right)$$

- And L types will acquire $S = S^*$ at cost S if

$$2 - \lambda - S^* > 1 \Rightarrow S^* < 1 - \lambda$$

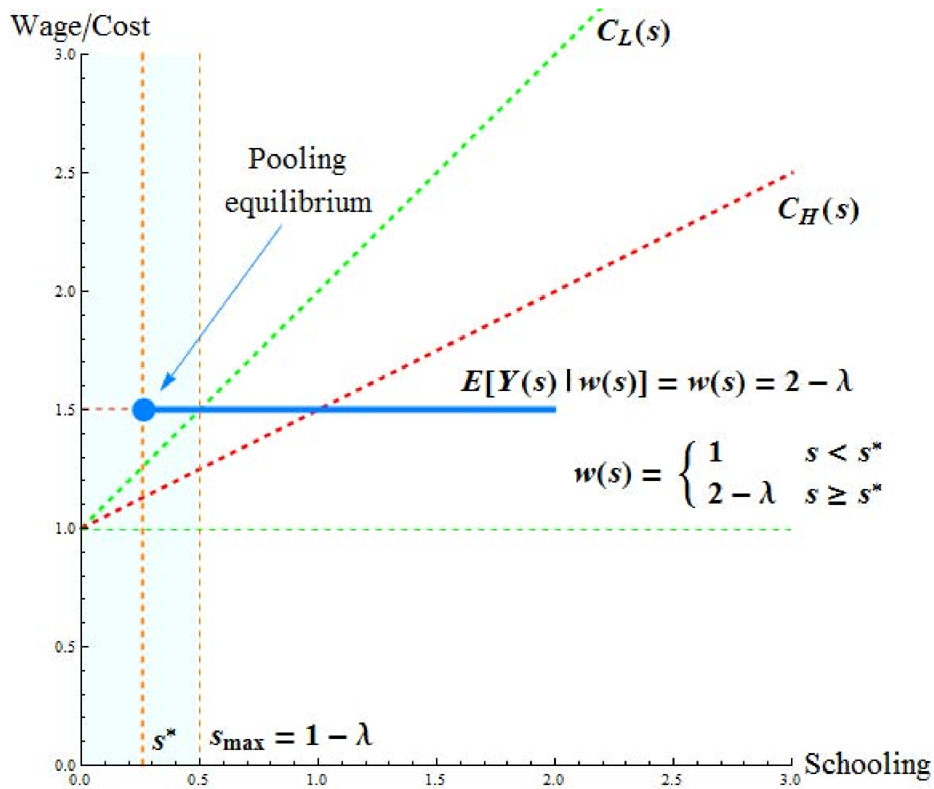
- So, if $S^* < 1 - \lambda$, all workers acquire education S^*.

- Are employers' wages rational given this fact? Yes. Because expected productivity of the working population is

$$E\left[Y\left(S \geq S^*\right)|w\left(S\right)\right] = 1 + (1 - \lambda) = 2 - \lambda = w\left(S > S^*\right)$$

- So, this is a feasible pooling equilibrium, as depicted below

Figure 2: Potential Pooling Equilibria with $\lambda = 0.5$



- [Note: this equilibrium is slightly strange because it does not specify what would happen if employers ever met a group of workers with $s = 0$ and found their productivity was also $2 - \lambda$. The Spence model was written when game theory was still in its infancy, and it does not do a tidy job of considering how 'off equilibrium' beliefs affect the model.]

- [Observe that $S^* > 1 - \lambda$ would not be a feasible equilibrium wage schedule. Under this schedule, high but not low productivity workers would acquire education, yet the wage paid to high productivity workers would only be $2 - \lambda$ whereas their productivity would be 2. Employers would have an incentive to deviate from this wage schedule to bid up wages of high productivity workers. This could theoretically happen, but it would create an entirely different possible equilibrium.]

3.5 Pooling equilibrium with no education

- It's crucial to remember that we haven't specified with equilibrium actually will happen here. We have only shown examples of many different equilibria which *could* possible occur. Now consider a different pooling equilibrium in which employers offer the wage schedule:

$$w(S) = (2 - \lambda) + I[S^* \geq 3].$$

- Who will obtain schooling in this case? The answer is no one, since the cost of obtaining 3 units of schooling for both H and L exceeds the wage benefit of 1.

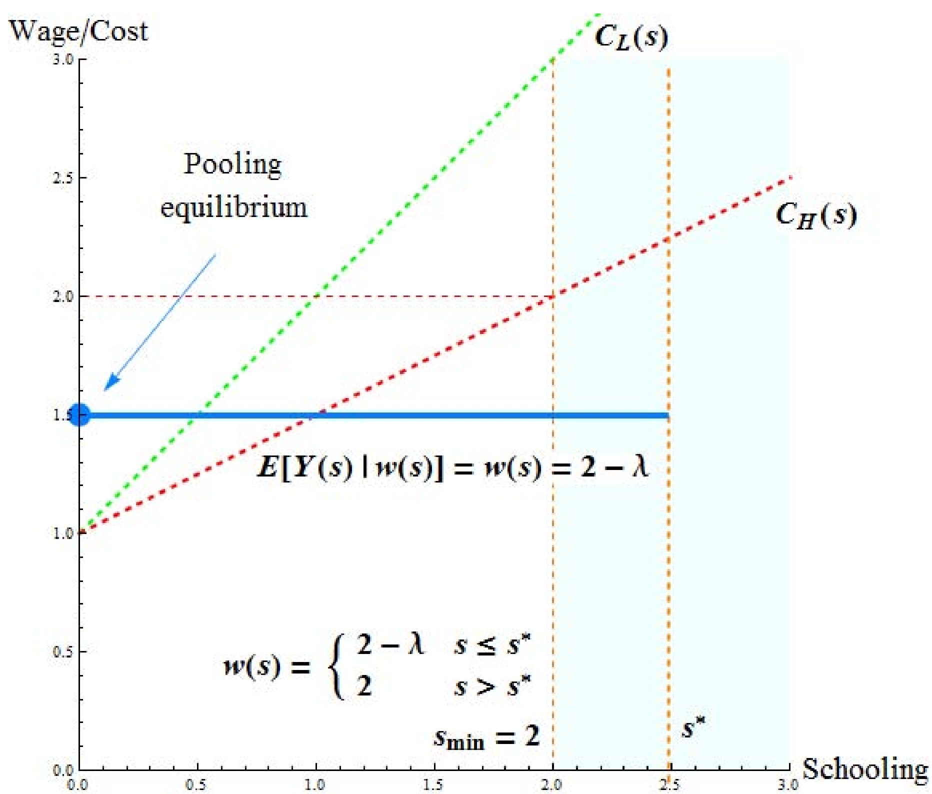
- But employer's beliefs are self-confirming: the pool of entirely uneducated workers *does* have average productivity $2-\lambda$, which is equal to the wage:

$$E\left[Y\left(0\right)|w\left(S\right)\right]=2-\lambda=w\left(0\right)$$

  So this is another feasible equilibrium.

- Can these three equilibria (two pooling, one separating) be ranked in terms of total welfare? Yes. In a world where schooling is only used for signaling productivity (and does not generate any new productivity), the pooling equilibrium with zero schooling is preferable to either of the equilibria that involve non-zero schooling. Productivity and aggregate wages are identical in all cases, but any equilibria involving non-zero levels of schooling includes wasteful expenditures on schooling. These schooling investments are pure deadweight losses from a social perspective, since they do not raise output.

Figure 3: Another Set of Potential Pooling Equilibria with $\lambda=0.5$



3.6 A slightly more ambitious example [For self-study]

- Consider a model with a continuous distribution of productivity and a single type of education: the "diploma."
  - Productivity η is distributed uniformly between 0 and 100.
  - The cost of obtaining a diploma is $80 - 0.50\eta$. That is, the cost of the diploma (in study time, perhaps) is lower for more productive workers.
  - Obtaining a diploma does not affect productivity.
  - Employers cannot distinguish worker productivity and so pay expected productivity. Without further information, the wage would be $w = E[\eta] = 50$.
- What are equilibrium education and wages in this model?
- The solution again relies on the Nash equilibrium concept.
  - We first solve for workers' optimal education choice taking wages as given.
  - We then solve for the employer wages given workers' education choices.
  - We finally find the equilibrium wages that satisfy both choices simultaneously (so that they are mutually consistent): $E\left[\eta\left(w\right)\right] = \eta$.
- Define w_1 as the wage of someone with a diploma and w_0 as the wage of someone without.
- A worker will get a diploma if the wage gain $(w_1 - w_0)$ exceeds the cost:

$$\begin{aligned} w_1 - w_0 &\geq (80 - 0.50\eta), \\ \eta^* &= 2 \cdot (w_0 - w_1 + 80). \end{aligned}$$

  A worker with $\eta \geq \eta^*$, will obtain a diploma, otherwise not.

- Now, we solve for wages given η^*. Employers will pay wages equal to expected productivity given diploma/no diploma. Using the uniform distribution of η, this gives:

$$\begin{aligned} w_1 &= E(\eta|\eta \geq \eta^*) = \frac{\eta^* + 100}{2}, \\ w_0 &= E(\eta|\eta < \eta^*) = \frac{\eta^*}{2} \end{aligned}$$

- So plugging employers' wage policies into workers' schooling strategies yields a solution for η^*:

$$\begin{aligned}
\eta^* &= 2 \cdot (w_0 - w_1 + 80) \\
&= 2 \cdot \left(\frac{\eta^*}{2} - \frac{\eta^* + 100}{2} + 80 \right) \\
&= 2 \cdot 30 \\
\eta^* &= 60, \\
&\Rightarrow w_0 = 30, w_1 = 80.
\end{aligned}$$

- Let's check this solution. At the equilibrium value of η^*, a worker with $\eta = \eta^* = 60$ must be indifferent between getting a diploma or not. Without a diploma, she gets a wage of 30. With a diploma, her net wage is $80 - (80 - 0.5\eta) = 30$. So she is indifferent. Clearly for $\eta > 60$, workers will get a diploma, otherwise not.

- As above, obtaining a diploma is privately optimal but socially unproductive. One way to see this is to check average wages in the economy:

$$E(w) = 0.6 \cdot 30 + 0.4 \cdot 80 = 50,$$

  which is exactly the wage that would prevail if *no one* got a diploma. Diplomas do not affect total societal output.

- But in the separating equilibrium, 40 percent of workers have bought an education at average cost of $80 - 0.5 \cdot 80 = 40$. And this is pure deadweight loss: total output and the sum of wages paid are identical whether or not workers obtain education. It is privately beneficial, however, for more productive workers to obtain an education to raise their wages (in the process, lowering the wages of less productive workers).

3.7 Empirical implications of signaling

Does the signaling model share any implications with the Becker Human Capital model?

1. People who attend additional years of schooling are more productive. YES.
2. People who attend additional years of schooling receive higher wages. YES.
3. People will attend school while they are young, i.e., before they enter the workforce. YES.
4. The rate of return to schooling should be roughly equal to the rate of interest. NO PREDICTION.

Because the empirical implications of the Human Capital and Signaling models appear so similar, many economists had concluded that these models could not be empirically distinguished. The paper published in the *Quarterly Journal of Economics* in 2000 by Tyler, Murnane and Willett (below) demonstrates that this conclusion was premature.

4 Testing signaling versus human capital models of education

Now that we've got the theory under our belts, let's move onto testing the role of learning versus signaling in education. Does it seem plausible that education serves (in whole or part) as a signal of ability rather than simply a means to enhance productivity?

- You obviously learn some valuable skills in school (e.g., engineering, computer science, signaling models).
- Many MIT students will be hired by finance and consulting firms that have no direct use for these skills.
- Why do these firms recruit at MIT rather than at Hampshire College, which produces many students with no engineering or computer science skills (let alone, knowledge of signaling models)?
- Why did you choose MIT over your state university that probably costs one-third as much? Is this entirely due to MIT's vaunted educational quality, or is some of it due to credentialism?

Harder question: How do you go about empirically distinguishing the human capital from the signaling model?

1. Measure whether more educated people are more productive? (Would be true for either model.)
2. Test whether higher ability obtain more education? (Could be true in either case—certainly true in the signaling case.)
3. Find people of identical ability and randomly assign some of them to go to college. Check if the college educated ones earn more? (Both models say they would.)
4. Measure people's productivity before and after they receive education—see if it improves. (Conceptually okay, difficult to do.)
5. Find people of identical ability and randomly assign them a diploma. See if the ones with diplomas earn more. (*This* is a pure test of signaling.)

5 The Tyler, Murnane and Willett study

- TMW are interested in knowing whether the General Educational Development certificate (GED) raises the subsequent earnings of recipients.
- This question is quite important for educational policy:
  - By 1996, 9.8% of those ages $18-24$ had completed High School via the the GED versus 76.5% via a HS diploma.

- See Table I. Notice that between 1990 and 1996, HS Diploma rates actually fell dramatically for Black, Non-Hispanics. The rise in the GED just offset this. We ought to hope that these GED-holders are doing somewhat better than HS dropouts.

- In 1996, $759,000$ HS Dropouts attempted the GED and some $500,000$ passed.
- The monetary cost of taking the GED is \$50 and the exam lasts a full day.
- The average person spends 20 hours studying for the GED (though some spend much more and some spend zero).
- See Table II. GED holders earn substantially less than HS graduates, but somewhat more than HS Dropouts.
- Why can't we simply compare wages of GED versus non-GED holders to measure the signaling effect of the GED?Self-selection (endogenous choice):
  - GED holders probably would have earned less than HS Diploma holders regardless. These are not typically the cream of your HS class, or else they would probably graduate HS and go onto college.
  - GED holders chose to take the GED, and probably would have earned more than other HS dropouts regardless. Relative to other dropouts, GED holders have:
    - More years of schooling prior to dropout.
    - Higher measured levels of cognitive skills.
    - Their parents have more education.
- So, simple comparisons of earnings among dropouts/ GED holders/ HS diploma holders tell us little about the causal effect of a GED for a person who obtains it.

5.1 The TMW strategy

- GED passing standards differ by U.S. state. Some test takers who would receive a GED in Texas with a passing score of $40-44$ would not receive a GED in New York, Florida, Oregon or Connecticut with the identical scores.
- But if GED score is a good measure of a person's ability/productivity, then people with same 'ability' $(40-44)$ are assigned a GED in Texas but not in New York.
- This quasi-experiment effectively randomly assigns the GED 'signal' to people with the same GED scores across different U.S. states.
- If we could determine who these marginal people are, we could identify the pure signaling effect of the GED, holding ability constant.

5.2 What does the signaling model predict in this case?

- Since some dropouts obtain the GED and some do not, it's plausible that the market is at some type of 'separating' equilibrium (i.e., not everyone gets the signal).
- For the GED to perform as a signal, it needs to be the case that the cost of obtaining it is lower for more productive workers (otherwise everyone or no one would get it). This seems quite plausible: you cannot pass the GED without some education and study.
- In equilibrium, the following must be true for individuals:

$$\begin{aligned} w_{GED} - w_{NO-GED} &\geq C_{GED} \Rightarrow \text{obtain}, \\ w_{GED} - w_{NO-GED} &< C_{GED} \Rightarrow \text{don't obtain}, \end{aligned}$$

  where C_{GED} is the direct and indirect costs of obtaining the GED.

- And the following must be true for employers:

$$\begin{aligned} w_{GED} &= E(\text{Productivity}|C_{GED} \leq w_{GED} - w_{NO-GED}), \\ w_{NO-GED} &= E(\text{Productivity}|C_{GED} > w_{GED} - w_{NO-GED}). \end{aligned}$$

- If these conditions are satisfied, firms will be willing to pay the wages (w_{GED}, w_{NO-GED}) to GED and non-GED holders respectively, and workers will self-select to obtain the GED accordingly.
- Notice an additional hidden assumption: firms cannot perfectly observe worker ability independent of the GED. If they could, the GED would not have any intrinsic signaling value since employers could judge productivity without needing this signal. It seems quite reasonable to assume that firms cannot observe ability perfectly.
- Given the quasi-experimental setup, the signaling model predicts that workers with GED scores of $40 - 44$ will earn more if they receive the GED certificate than if they do not.
- By contrast, the Human Capital model implies that since ability is comparable among these groups, their wages will also be comparable.

5.3 Estimation

- The econometric strategy should be quite familiar now. We want to estimate:

$$T = E\left[Y_1 - Y_0 | GED = 1\right]$$

  where Y_1, Y_0 are earnings with and without the GED *for the people who obtained the GED*—that is, we want to estimate the effect of 'treatment on the treated.′

- The variable that randomizes assignment of the GED is location: Texas vs. New York. So, we need to assume the following *for those in the relevant score range,* S, *(where* $S \in [40, 44]$*)*:

$$\begin{aligned} E[Y_1|NY,S] &= E[Y_1|TX,S], \\ E[Y_0|NY,S] &= E[Y_0|TX,S] \end{aligned}$$

- If these assumptions are correct, a valid estimate of the treatment effect is:

$$\hat{T} = E[Y_1|TX,S] - E[Y_0|NY,S].$$

  That is, we would compare GED holders from Texas (in score range $40 - 44$) to GED non-holders from NY in score range $40 - 44$ to get an estimate of T.

- However, we might be concerned that there is also a direct effect of being in NY vs. TX that operates independently of the GED at any level of ability. For example

$$E[Y_1|NY,S] - E[Y_1|TX,S] = E[Y_0|NY,S] - E[Y_0|TX,S] = \delta.$$

  In this case, $\hat{T}$ from our previous equation would estimate $T + \delta$, i.e., the treatment effect plus the location effect.

- To surmount this problem, TMW select a control group of GED test takers with scores just about the cutoff for both groups of states. Hence, the GED treatment works as follows:

	Low Passing Standard	High Passing Standard
Low Score (treatment group)	GED	NO GED
High Score (control group)	GED	GED

- The outcome variable will be earnings for each of these four groups:

	Low Passing Standard	High Passing Standard
Low Score (treatment group)	$E[Y_1|TX,S = Low]$	$E[Y_0|NY,S = Low]$
High Score (control group)	$E[Y_1|TX,S = High]$	$E[Y_1|NY,S = High]$

- Hence, the Diff-in-Diff estimate is:

$$\begin{aligned} E\left[\hat{T}\right] &= E[Y_0|TX,S = Low] - E[Y_1|NY,S = Low] \\ &\quad - E[Y_1|TX,S = High] - E[Y_1 NY,S = High] \\ &= T + \delta - \delta \\ &= T \end{aligned}$$

- Results:
  - See Table V.
  - See Figures I-III.

5.4 Conclusions from TMW study

- Large signaling effects for whites, estimated at 20% earnings gain after 5 years.
- Does this prove that GED holders are *not* more productive than non-GED holders?
  - No. Just the opposite.
  - For there to be a signaling equilibrium, it must be the case that GED holders *are* on average more productive than otherwise similar HS dropouts who do not hold a GED.
- Do these results prove that education is unproductive?
  - No, they also have nothing to say on this question because education/skill is effectively held constant by this quasi-experiment.
- What the study shows unambiguously is that the GED is taken as a positive signal by employers. And this can only be true if:
  1. GED holders are on average more productive than non-GED holders.
  2. The GED is in some sense more expensive for less productive than more productive workers to obtain. This probably has to do with maturity, intellect, etc.
  3. Employers are unable to perfectly distinguish productivity directly and hence use GED status as one signal of expected productivity.