Labor Economics, 14.661. Lectures 1 and 2: Labor Market Externalities

Daron Acemoglu

MIT

Oct. 27 & Nov. 1, 2022

Introduction

- General interest over the recent two decades or so on the modeling and estimation of various aspects of labor market externalities.
- Two different aspects of externalities:
    1. Externalities in (local) labor markets due to production, matching or other market interactions.
    2. Externalities in social environments, including schools, friendships, neighborhoods, and networks as well as their implications for social mobility, inequality
- Both types of externalities may be important in practice and have major welfare consequences.
- Both types of externalities present a range of challenges in estimation.
- These two lectures on labor market externalities, and then we will turn to social mobility and peer effects.

Labor Market Externalities

- *What happens to my wages/earnings if other workers in the same labor market/economy increase their schooling by one year?*
- The Becker-Mincer-Ben Porath framework silent on this (it answers the question of how much my earnings increased when I increase my schooling by one year).
- Useful to distinguish between non-pecuniary externalities (where the effect of others' education on me as "technological") vs. pecuniary externalities (where the effect of others' education on me works through changes in equilibrium prices).

What Is an Externality?

- Imagine your coworkers' human capital makes you more productive
  - e.g., academics would like to be together with other high-quality academics
- Imagine your production function is

$$y = f\left(h, \bar{H}\right)$$

where $\bar{H}$ is the average human capital of your coworkers.

- This is *a technological* spillover of productivity.
- Is it an externality?

What Is an Externality? (continued)

- Not necessarily.
- If all of the spillover is within the firm, the firm will internalize it in its hiring decisions and in its compensation of different workers with different amounts of human capital.
- In that case, there is a technological spillover, but no labor market externality.
- Externalities require
    1. either that productivity spillovers are beyond firm boundaries
    2. or that firms are unable to compensate workers appropriately for their contribution to their coworkers' productivity (why would this be the case?)
- Let us now focus on externalities.

Nonpecuniary Externalities

- Nonpecuniary externalities$\approx$ technological spillovers of productivity that are not internalized by prices.
- Canonical example due to Jane Jacobs *The Economy of Cities*: managers from different companies exchange ideas.
- Very popular in economics (e.g., Lucas's famous 1986 endogenous growth model)
- What other contexts would this be important in?

Nonpecuniary Positive Externalities

- A simple model of nonpecuniary externalities:
- Suppose that the output (or marginal product) of a worker, i, is

$$y_i = Ah_i^{\nu},$$

where h_i is the human capital (schooling) of the worker, and A is aggregate productivity.

- Assume that labor markets are competitive. So individual earnings are $W_i = Ah_i^{\nu}$.
- Key idea: the exchange of ideas among workers raises productivity.
- This can be modeled by allowing A to depend on aggregate human capital. In particular, suppose that

$$A = BH^{\delta} \equiv \mathbb{E}\left[h_i\right]^{\delta}, \tag{1}$$

where H is a measure of aggregate human capital, $\mathbb{E}$ is the expectation operator, B is a constant, and $\delta > 0$ (the case where < 0 would have negative externalities).

Nonpecuniary Positive Externalities (continued)

- Individual earnings can then be written as

$$W_i = Ah_i^{\nu} = BH^{\delta} h_i^{\nu}.$$

- Taking logs:

$$\ln W_i = \ln B + \delta \ln H + \nu \ln h_i. \tag{2}$$

- If external effects are stronger within a geographical area, as seems likely in a world where human interaction and the exchange of ideas are the main forces behind the externalities, then equation (2) should be estimated using measures of H at the local level.
- This is a theory of non-pecuniary externalities, since the external returns arise from the technological nature of equation (1).
- Nonpecuniary externalities unattractive for a number of reasons:
  1. Very reduced form.
  2. Do we really expect workers in chemical factories to have a direct productivity effect on retail workers?

Nonpecuniary Positive Externalities (continued)

- Instead, more compelling (at least from one perspective) sources of spillovers:
  - Interactions in the labor market mediated by prices, but externalities might still be at present; *pecuniary externalities.*
  - Interactions in the product market; when computer users become more productive, they can supply cheaper computers to retail companies, again pecuniary externalities.
  - Interactions via R&D and innovation; the semiconductor or the combustion engine have increased the productivity of many workers in many different sectors of the economy.
- The last one may or may not be a pecuniary externality.
- However, except those working in the labor market, the remaining externalities would be economy-wide (sometimes even world-wide), thus difficult to estimate with cross-sectional or panel data variation.
- Thus, let us focus on labor market interactions.

Pecuniary Externalities

- *Do pecuniary externalities matter?*
- Not in Arrow-Debreu.
- Why not?
- Could they matter in other environments?
- The answer is "perpaps yes"—if we are away from the complete markets benchmark.

An Aside on Pecuniary Externalities

- Here is a slightly more formal discussion. Consider an exchange economy with $N+1$ agents with quasi-linear utility given by

$$y_i + u_i(x_i),$$

where y_i is money and x_i is the focal good. The price of this good is determined from market clearing:

$$\sum_{i=0}^{N} x_i = 0,$$

since this is an exchange economy (some agents are sellers and some are buyers). Market clearing determines the price as

$$p = p(x_0, \ldots, x_N).$$

Suppose all of these functions are differentiable.

An Aside on Pecuniary Externalities (continued)

- Now suppose that we change the behavior of agent 0 and consider the welfare impact of this on all agents.
- This is given by:

$$U = \sum_{i=0}^{N} \frac{du_i(x_i|p)}{dp} \frac{\partial p}{\partial x_0}$$

  where $u_i(x_i|p)$ is the "reduced-form" utility function of agent i (meaning that after we substitute in the budget constraint).
- We also have

$$\frac{du_i(x_i|p)}{dp} = \frac{\partial u_i(x_i|p)}{\partial p} + \frac{\partial u_i(x_i|p)}{\partial x_i} \frac{dx_i}{dp}$$

- First suppose that all agents are optimizing. Then the second term in this expression is equal to zero because $\frac{\partial u_i(x_i|p)}{\partial x_i} = 0$ by the envelope theorem, and the first term is equal to

$$\frac{\partial u_i(x_i|p)}{\partial p} = -x_i.$$

An Aside on Pecuniary Externalities (continued)

- Thus
$$U = -\sum_{i=0}^{N} x_i = 0,$$
where the last equality follows from market clearing.
- Thus there is no first-order welfare effect from the change in the action of agent 0—no first-order pecuniary externalities.
- Now if instead we have that for a subset of the agents $\frac{\partial u_i(x_i|p)}{\partial x_i} > 0$ (and equal to zero for the rest), then we would have a first-order negative welfare effect from the increase in price, reducing overall welfare. In other words,
$$U = \sum_{i=0}^{N} \frac{\partial u_i(x_i|p)}{\partial x_i} \frac{dx_i}{dp} \frac{\partial p}{\partial x_0} - \sum_{i=0}^{N} x_i < 0$$
because $\frac{dx_i}{dp} < 0$ and $\frac{\partial p}{\partial x_0} > 0$—thus now there are first-order pecuniary externalities.

Pecuniary Labor Market Externalities

- First suggested in Alfred Marshall's *Principles of Economics* in the context of benefits of geographic concentration of industry.
- A complementary story with labor market imperfections, innovation investment by firms and training by workers developed in Acemoglu (1997).
    - Firms find it profitable to invest in new technologies only when there is a sufficient supply of trained workers to replace employees who quit.
- This is a *pecuniary* externality, since it is not built in in the form of technological spillovers, but works through market interactions and results from the fact that prices at which labor is transacted is not equal to its marginal product.
- A related model developed in Acemoglu (1996). Here is simplified version of this model.

A Search Model of Pecuniary Externalities

- Consider an economy lasting two periods, with production only in the second period, and a continuum of workers normalized to 1.
- Take human capital of each worker i, h_i, as given.
- A continuum of risk-neutral firms.
- In period 1, firms make an irreversible investment decision, k, at cost Rk.
- Workers and firms come together in the second period.
- The labor market is not competitive; instead, firms and workers are matched randomly, and each firm meets a worker.
- The only decision workers and firms make after matching is whether to produce together or not to produce at all (since there are no further periods).

Pecuniary Externalities (continued)

- If firm f and worker i produce together, their output is

$$k_f^{\alpha} h_i^{\nu}, \tag{3}$$

where $\alpha < 1$, $\nu \leq 1 - \alpha$.

- Since it is costly for the worker-firm pair to separate and find new partners in this economy, employment relationships generate quasi-rents.
- Wages will therefore be determined by rent-sharing. Here, simply assume that the worker receives a share β of this output as a result of bargaining, while the firm receives the remaining $1 - \beta$ share (a simplified version of Nash bargaining).
- An equilibrium in this economy is a set of schooling choices for workers and a set of physical capital investments for firms.

Pecuniary Externalities (continued)

- Firm f maximizes the following expected profit function:

$$(1-\beta)k_f^{\alpha}\mathbb{E}[h_i^{\nu}] - Rk_f, \tag{4}$$

  with respect to k_f.
- Since firms do not know which worker they will be matched with, their expected profit is an average of profits from different skill levels.
- The function (4) is strictly concave, so all firms choose the same level of capital investment, $k_f = k$, given by

$$k = \left(\frac{(1-\beta)\alpha H}{R}\right)^{1/(1-\alpha)}, \tag{5}$$

  where

$$H \equiv \mathbb{E}[h_i^{\nu}]$$

  is the measure of aggregate human capital.

Pecuniary Externalities (continued)

- Now the equilibrium is straightforward to characterize.
- Substituting (5) into (3), and using the fact that wages are equal to a fraction β of output, the wage income of individual i is given by

$$W_i = \beta \left((1-\beta)\alpha H \right)^{\alpha/(1-\alpha)} R^{-\alpha/(1-\alpha)} (h_i)^{\nu}.$$

- Taking logs, this is:

$$\ln W_i = c + \frac{\alpha}{1-\alpha} \ln H + \nu \ln h_i, \tag{6}$$

where c is a constant and $\alpha / (1-\alpha)$ and ν are positive coefficients.
- The presence of $\ln H$ on the right hand side corresponds to *positive pecuniary externalities* (in the local labor market).
- What about education choices? This actually is a trickier issue and I will return to it later.

Pecuniary Externalities (continued)

- For now the important thing is that human capital externalities arise here because firms choose their physical capital in anticipation of the average human capital of the workers they will employ in the future.
- Since physical and human capital are complements in this setup, a more educated labor force encourages greater investment in physical capital and to higher wages.
- In the absence of the need for search and matching, firms would immediately hire workers with skills appropriate to their investments, and there would be no human capital externalities.
- Nonpecuniary and pecuniary theories of human capital externalities lead to similar empirical relationships since equation (6) is identical to equation (2), with $c = \ln B$ and $\delta = \alpha / (1 - \alpha)$.
- Again presuming that these interactions exist in local labor markets, we can estimate a version of (2) using differences in schooling across labor markets (cities, states, or even countries).

Ranking and Negative Externalities

- The above models focused on positive externalities to education.
- In contrast, in a world where education plays a "ranking" role, we might also expect *negative externalities*.
- Suppose that there are different types of jobs with different quality/attributes and higher-quality jobs pay higher wages (this may be due to bargaining reasons).
- Suppose also that higher education workers get the higher quality jobs (a sort of assignment rule).
- What will happen if all of the other workers in my labor market become more educated?

Signaling and Negative Externalities

- Consider now signaling. Take the most extreme world in which education is only a signal—it does not have any productive role.
- What happens if all of the other workers in my labor market become more educated?
- Perhaps the same thing as ranking?
- To answer this question, let us consider a simple signaling model (and the evidence for signaling in the labor market.

Baseline Signaling Model

- Consider the following simple model to illustrate the issues.
- There are two types of workers, high ability and low ability.
- The fraction of high ability workers in the population is λ.
- Workers know their own ability, but employers do not observe this directly. They only observe schooling.
- High ability workers always produce y_H, low ability workers produce y_L. (No productive role of schooling for simplicity now).

Baseline Signaling Model (continued)

- Workers can invest in education, $e \in \{0, 1\}$.
- The cost of obtaining education is c_H for high ability workers and c_L for low ability workers.
- **Crucial assumption** ("single crossing")

$$c_L > c_H$$

- That is, education is more costly for low ability workers. This is often referred to as the "single-crossing" assumption, since it makes sure that in the space of education and wages, the indifference curves of high and low types intersect only once. For future reference, I denote the decision to obtain education by $e = 1$.
- To start with, suppose that education does not increase the productivity of either type of worker.
- Once workers obtain their education, there is competition among a large number of risk-neutral firms, so workers will be paid their *expected productivity*.

Baseline Signaling Model (continued)

- Game of incomplete information $\longrightarrow$ Perfect Bayesian Equilibrium
- Two (extreme) types of equilibria in this game.

1. *Separating*, where high and low ability workers choose different levels of schooling.
2. *Pooling*, where high and low ability workers choose the same level of education.

Separating Equilibrium

- Suppose that we have

$$y_H - c_H > \ y_L > y_H - c_L \tag{7}$$

- This is clearly possible since $c_H < c_L$.
- Then the following is an equilibrium: all high ability workers obtain education, and all low ability workers choose no education.
- Wages (conditional on education) are:

$$w\left(e=1\right) = y_H \text{ and } w\left(e=0\right) = y_L$$

- Notice that these wages are conditioned on education, and *not directly on ability*, since ability is not observed by employers.

Separating Equilibrium (continued)

- Let us now check that all parties are playing best responses.
- Given the strategies of workers, a worker with education has productivity y_H while a worker with no education has productivity y_L. So no firm can change its behavior and increase its profits.
- What about workers?
- If a high ability worker deviates to no education, he will obtain $w(e=0) = y_L$, but

$$w(e=1) - c_H = y_H - c_H > y_L.$$

Separating Equilibrium (continued)

- If a low ability worker deviates to obtaining education, the market will perceive him as a high ability worker, and pay him the higher wage $w(e=1)=y_H$. But from (7), we have that

$$y_H - c_L < y_L.$$

- Therefore, we have indeed an equilibrium.
- In this equilibrium, education is valued simply because it is a signal about ability.
- Is "single crossing important"?

Pooling Equilibrium

- The separating equilibrium is not the only one.
- Consider the following allocation: both low and high ability workers do not obtain education, and the wage structure is

$$w\left(e=1\right)=\left(1-\lambda\right)y_L+\lambda y_H \text{ and } w\left(e=0\right)=\left(1-\lambda\right)y_L+\lambda y_H$$

  (Does this wage structure make sense?)
- Again no incentive to deviate by either workers or firms.
- Is this Perfect Bayesian Equilibrium reasonable?

Pooling Equilibrium (continued)

- Provided that
$$y_H - c_H > (1 - \lambda)\, y_L + \lambda y_H, \tag{8}$$
The answer is no.
- This equilibrium is being supported by the belief that the worker who gets education is no better than a worker who doesn't.
- But education is more costly for low ability workers, so they should be less likely to deviate to obtaining education.
- This can be ruled out by various different refinements of equilibria.

Pooling Equilibrium (continued)

- Simplest refinement: *Intuitive Criterion* by Cho and Kreps.
- The underlying idea: if there exists a type who will never benefit from taking a particular deviation, then the uninformed parties (here the firms) should deduce that this deviation is very unlikely to come from this type.
- This falls within the category of "forward induction" where rather than solving the game simply backwards, we think about what type of inferences will others derive from a deviation.

Pooling Equilibrium (continued)

- Take the pooling equilibrium above. Consider a deviation to $e = 1$.
- There is no circumstance under which the low type would benefit from this deviation, since

$$y_L > y_H - c_L,$$

and the low ability worker is now getting

$$(1-\lambda)\, y_L + \lambda y_H.$$

- Therefore, firms can deduce that the deviation to $e = 1$ must be coming from the high type, and offer him a wage of y_H.
- Then (8) ensures that this deviation is profitable for the high types, breaking the pooling equilibrium.

Pooling Equilibrium (continued)

- The reason why this refinement is called The Intuitive Criterion is that it can be supported by a relatively intuitive "speech" by the deviator along the following lines:

  you have to deduce that I must be the high type deviating to $e = 1$, since low types would never ever consider such a deviation, whereas I would find it profitable if I could convince you that I am indeed the high type). Of course, this is only very loose, since such speeches are not part of the game, but it gives the basic idea.

- The overall conclusion: separating equilibria, where education is a valuable signal, may be more likely than pooling equilibria.
- When would this not be the case?

Generalizations

- The two most important insights generalize: (weak) overeducation by high types and signaling value of education.
- Suppose education is continuous $e \in [0, \infty)$.
- Cost functions for the high and low types are $c_H(e)$ and $c_L(e)$, which are both strictly increasing and convex, with $c_H(0) = c_L(0) = 0$.
- The single crossing property is that

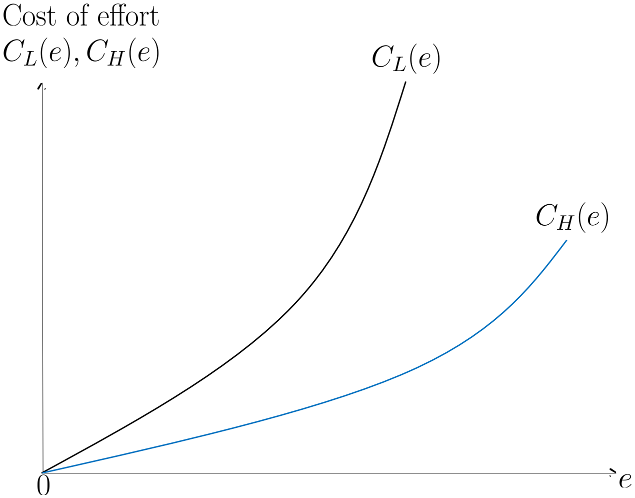
$$c'_H(e) < c'_L(e) \text{ for all } e \in [0, \infty),$$

  that is, the marginal cost of investing in a given unit of education is always higher for the low type (why is this the right condition?).
- Suppose that the output of the two types as a function of their educations are $y_H(e)$ and $y_L(e)$, with

$$y_H(e) > y_L(e) \text{ for all } e.$$

Generalizations (continued)

- The single crossing property:



Generalizations (continued)

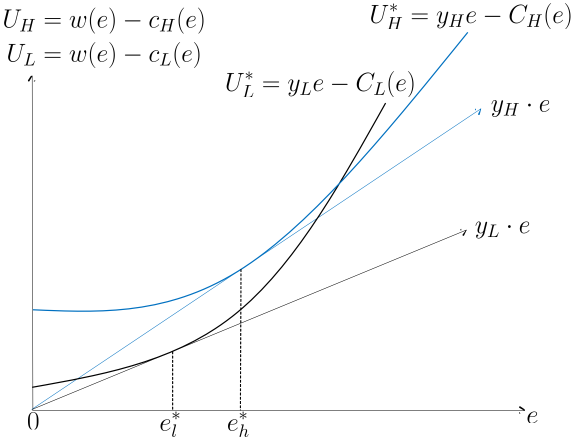
- Again there are many Perfect Bayesian Equilibria, some separating, some pooling and some semi-separating.
- But applying a stronger form of the Intuitive Criterion reasoning, we will pick the *Riley equilibrium* of this game, which is a particular separating equilibrium.

Riley equilibrium: first find the most preferred (*first-best*) education level for the low type in the perfect information case

$$y_L'(e_l^*) = c_L'(e_l^*).$$

Generalizations (continued)

- First best diagrammatically:



Generalizations (continued)

- Let us write the *incentive compatibility constraint* for the low type, such that when the market expects low types to obtain education e_l^*, the low type does not try to mimic the high type who is expected to do e_h:

$$y_L\left(e_l^*\right) - c_L\left(e_l^*\right) \geq w\left(e_h\right) - c_L\left(e_h\right). \tag{9}$$

- **Assumption**: the first-best is not "incentive compatible".
- Under this assumption, the Riley equilibrium involves a choice of e_h such that (9) holds as equality:

$$y_L\left(e_l^*\right) - c_L\left(e_l^*\right) = y_H\left(e_h\right) - c_L\left(e_h\right).$$

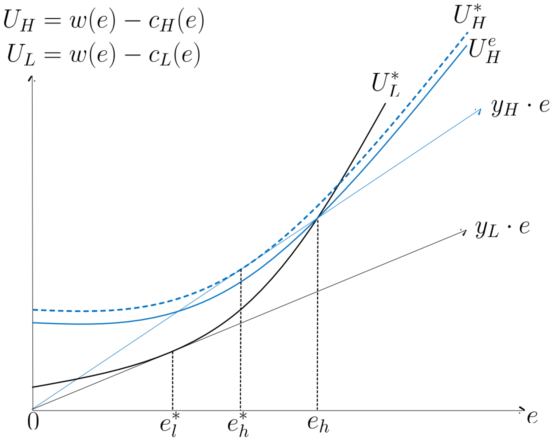
  (Question: why did we write $w\left(e_h\right) = y_H\left(e_h\right)$?)
- Then in the Riley equilibrium, we have "signaling value of education":

$$w\left(e_l^*\right) = y_L\left(e_l^*\right),$$

$$w\left(e_h\right) = y_H\left(e_h\right).$$

Generalizations (continued)

- *Riley equilibrium* diagrammatically:



Generalizations (continued)

- "Overeducation": high ability workers are investing in schooling more than they would have done in the perfect information case, in the sense that e_h characterized here is greater than the education level that high ability individuals chosen with perfect information, given by $y'_H(e_h^*) = c'_H(e_h^*)$.
- Why are high types are happy to do this? From the single-crossing property:

$$\begin{aligned} y_H(e_h) - c_H(e_h) &= y_H(e_h) - c_L(e_h) - (c_H(e_h) - c_L(e_h)) \\ &> y_H(e_h) - c_L(e_h) - (c_H(e_l^*) - c_L(e_l^*)) \\ &= y_L(e_l^*) - c_L(e_l^*) - (c_H(e_l^*) - c_L(e_l^*)) \\ &= y_L(e_l^*) - c_H(e_l^*). \end{aligned}$$

This is because of the "signaling value of education".

Evidence on Labor Market Signaling

- For different types of evidence:
  1. Do degrees matter?
  2. Do compulsory schooling laws affect schooling levels for higher grades?
  3. Returns to GED?
  4. Investigation of *negative externalities*
- Why are these informative about signaling?
- Which ones are more convincing?

Equilibrium Responses in Signaling Models

- The second approach was pioneered by Lang and Kropp (1986), who observed that in the presence of signaling, binding compulsory schooling laws should increase the education investments of those not directly affected—as an equilibrium response to the behavior of others whose education is being increased by the laws because they want to signal themselves apart.
  - E.g., you could send a signal distinguishing yourself from certain workers by graduating from high school, but now everybody is forced to graduate from high school, so you have to go to college.
- A related idea is tested in Bedard (2001): the introduction of access to university should lead to an increase in high school dropout rate. Why?
- Bedard tests this by comparing high school dropout rates in labor markets with and without two-year or four-year colleges

Evidence of Equilibrium Responses?

PREDICTED EDUCATIONAL GROUP SIZES

	MEN		WOMEN	
EDUCATION GROUP	Regions without Access	Regions with Access	Regions without Access	Regions with Access
	Access Is Defined as a Public 2- or 4-Year Degree-Granting Institution			
High school dropouts	18.5**	19.8	15.8***	18.5
High school graduates	34.2	28.3	51.5	45.0
University enrollees	47.3	51.9	32.7	36.5
	Access Is Defined as a Public 4-Year Degree-Granting Institution			
High school dropouts	18.9**	19.7	17.0***	17.8
High school graduates	33.8	27.5	50.3	44.9
University enrollees	47.3	52.8	32.7	37.3
	Access Is Defined as a 2- or 4-Year Degree-Granting Institution			
High school dropouts	18.0	19.6	14.0***	18.4
High school graduates	36.5	29.2	54.6	45.7
University enrollees	45.5	51.2	31.4	35.9
	Access Is Defined as a 4-Year Degree-Granting Institution			
High school dropouts	18.6*	19.6	17.2	17.5
High school graduates	35.9	28.3	51.6	45.8
University enrollees	45.5	52.1	31.2	36.7

Direct Signals

- Third approach: Tyler, Murnane and Willett.
- Passing grades in the Graduate Equivalent Degree (GED) differ by state.
- So an individual with the same grade in the GED exam will get a GED in one state, but not in another.
- If the score in the exam is an unbiased measure of human capital, and there is no signaling, these two individuals should get the same wages.
- If the GED is a signal, and employers do not know where the individual took the GED exam, these two individuals should get different wages.
- Using this methodology, the authors estimate that there is a 10-19 percent return to a GED signal.

Direct Signals (continued)

TABLE V
DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE IMPACT OF THE GED ON 1995 EARNINGS OF DROPOUTS WHO TESTED IN 1990 (STANDARD ERRORS ARE IN PARENTHESES)

	Experiment 4			Experiment 3			Experiment 3*		
	State passing standard is		Low-High standard contrast	State passing standard is		Low-High standard contrast	State passing standard is		Low-High standard contrast
	Low	High		Low	High		Low	High	
Panel A: Whites									
Test score is									
Low	9628	7849	1779	9362	7843	1509	9362	8616	746
	(361)	(565)	(670)	(400)	(312)	(507)	(400)	(219)	(456)
High	9981	9676	305	9143	9165	−23	9143	9304	−162
	(80)	(65)	(103)	(135)	(63)	(149)	(135)	(135)	(150)
Difference-in-differences			1473*			1531**			907~
for whites			(678)			(529)			(481)
Panel B: Minorities									
Test score is									
Low	6436	8687	−2252	7005	7367	−363	7005	6858	147
	(549)	(690)	(882)	(347)	(347)	(495)	(347)	(290)	(452)
High	7560	8454	−894	7782	8375	−593	7782	7568	214
	(184)	(96)	(207)	(214)	(93)	(233)	(214)	(133)	(252)
Difference-in-differences			−1357			231			−67
for minorities			(906)			(548)			(518)

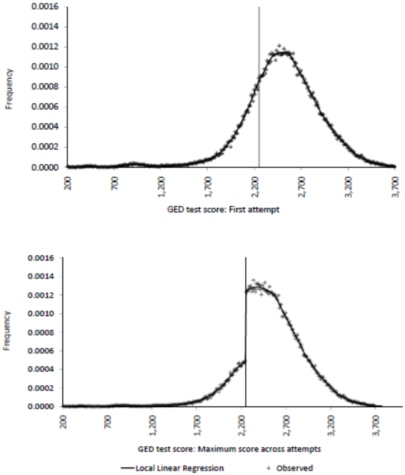
- Experiment 3 an 3*: treatment only Texas, Mississippi, Louisiana and Nebraska; Experiment 4:only New York as control

Direct Signals (continued)

- An interesting result we can see from the above table is that there are no GED returns to minorities.
- This is also consistent with the signaling view, since it turns out that many minorities prepare for and take the GED exam in prison. Therefore, GED would be not only a positive signal, but also likely a signal that the individual was at some point incarcerated. Hence not a good signal at all.

But Potential Problem

- Endogeneity of test taking behavior (from Heckman et al., 2010) from the state of Missouri:



Back to Negative Externalities

- Now suppose there is signaling, and suppose there are no direct productivity benefits from education.
- Contrast two situations: in the first, all individuals have 12 years of schooling and in the second all individuals have 16 years of schooling.
- Since education has no productive role, and all individuals have the same level of schooling, in both allocations they will earn exactly the same wage (equal to average productivity).
- Therefore, here the increase in aggregate schooling does not translate into aggregate increases in wages.
- But in the same world, if one individual obtains more education than the rest, there will be a private return to him, because he would signal that he is of higher ability.
- Is this a pecuniary or a nonpecuniary externality?

Back to Negative Externalities (continued)

- Now returning to regression models, in a world where signaling or ranking is important, we might also want to estimate an equation of the form (2), but when signaling issues are important, we would expect δ to be negative.
- When others invest more in their education, a given individual's rank in the distribution declines, hence others are creating a negative externality on this individual via their human capital investment.

Econometric Framework for Externalities/Peer Effects

- The simplest econometric model would be

$$y_{ij} = \beta_{\text{own}} x_{ij} + \beta_{\text{spillover}} \bar{X}_j + \varepsilon_{ij} \tag{10}$$

  where $\bar{X}$ is average characteristic (e.g., average schooling) and y_{ij} is the outcome of the ith individual in group j.
- This is the model we discussed in the context of human capital externalities.
- Manski (1993) calls this type of influences *contextual effects* — they come from the context in which the individual is situated.
- As we have already discussed, identification here will require some structural assumptions or preferably exogenous variation in both x_{ij} and $\bar{X}$.

Econometric Issues

- The alternative is what Manski refers to as *endogenous effects* — because they are created by endogenous variables.
- The simplest form would be

$$y_{ij} = \beta_{\text{own}} x_{ij} + \alpha_{\text{spillover}} \bar{Y}_j + \varepsilon_{ij} \tag{11}$$

  where $\bar{Y}$ is the average of the outcomes.
- The identification of such endogenous effects is even more difficult (though this hasn't stopped people estimating such models).

Econometric Issues (continued)

- An obvious problem is that because $\bar{Y}_j$ does not vary by individual, this regression amounts to one of $\bar{Y}_j$ on itself at the group level.
- This is a serious econometric problem.
- One imperfect way to solve this problem is to replace $\bar{Y}_j$ on the right hand side by $\bar{Y}_j^{-i}$ which is the average excluding individual i. (Why doesn't is solved the problem?)
- Another approach is to impose some timing structure.
- For example:

$$y_{ijt} = \beta_{\text{own}} x_{ijt} + \alpha_{\text{spillover}} \bar{Y}_{j,t-1} + \varepsilon_{ijt}$$

- There are still some serious problems irrespective of the approach taken;
  1. the timing structure is arbitrary, and
  2. there is no way of distinguishing peer group effects from *"common shocks"*.

Evidence on Labor Market Externalities

- Let us now return to our equation of interest

$$\ln W_i = \ln B + \delta \ln H + \nu \ln h_i .$$

- This is an example of contextual effects, but still quite challenging to estimate.
- Ordinary Least Squares (OLS) estimation of this type of equation using city or state-level data yield very significant and positive estimates of δ, indicating substantial positive human capital externalities; e.g., Jim Rauch's paper in the *Journal of Urban Economics* 1993.
- Similar results later found in many other papers.
- The next table shows it for US states.

OLS Evidence

Table 2 OLS ESTIMATES OF PRIVATE AND EXTERNAL RETURNS TO SCHOOLING

	1960–1980 (1)	1950–1980 (2)	1950–1990 (3)	1950 (4)	1960 (5)	1970 (6)	1980 (7)
(a) Private Returns							
Private return to schooling	0.073 (0.0003)	0.068 (0.0003)	0.075 (0.0003)	0.055 (0.002)	0.069 (0.001)	0.076 (0.001)	0.075 (0.001)
State of residence main effects?	Yes	Yes	Yes	No	No	No	No
(b) Private and External Returns							
Private return to schooling	0.073 (0.000)	0.068 (0.000)	0.074 (0.000)	0.055 (0.002)	0.068 (0.001)	0.075 (0.001)	0.074 (0.000)
External return to schooling	0.073 (0.016)	0.061 (0.004)	0.072 (0.003)	0.136 (0.017)	0.136 (0.016)	0.128 (0.021)	0.160 (0.027)
State of residence main effects?	Yes	Yes	Yes	No	No	No	No
N	609,852	626,511	729,695	16,659	72,344	161,029	376,479

Notes: Standard errors corrected for state–year clustering are shown in parentheses. The data are from the Census IPUMS for 1950 through 1990,

Interpretation

- There are at least two problems with this type OLS estimates.
- **First problem:** high-wage cities or states may attract a large number of high education workers or give strong support to education.
  - Rauch uses a cross-section of cities.
  - Including city or state fixed affects ameliorates this problem, but does not solve it, since states' attitudes towards education and the demand for labor may comove. The ideal approach would be to find a source of quasi-exogenous variation in average schooling across labor markets.
- Acemoglu and Angrist (2000): exploit exogenous sources of variation due to cross-state differences in compulsory schooling laws. The advantage is that these laws not only affect individual schooling but average schooling in a given area.

Interpretation (continued)

- **Second problem:** even if we have an instrument for average schooling in the aggregate, estimates of labor market externalities might be spurious.
- In particular, if individual schooling is measured with error (or for some other reason OLS returns to individual schooling are not the causal effect), some of this discrepancy between the OLS returns and the causal return may load on average schooling, even when average schooling is instrumented.
- This suggests that we may need to instrument for individual schooling as well (so as to get to the correct return to individual schooling).

A Little Bit of Econometrics

- To elaborate on the second problem, let Y_{ijt} be the log weekly wage, than the estimating equation is

$$Y_{ijt} = X_i' \mu + \delta_j + \delta_t + \gamma_1 \overline{S}_{jt} + \gamma_{2i} s_i + u_{jt} + \varepsilon_i, \tag{12}$$

- To illustrate the main issues, ignore time dependence, and consider the population regression of Y_{ijt} on s_i:

$$Y_{ijt} = \mu_0 + \delta_j + \delta_t + \rho_0 s_i + \varepsilon_{0it};\ \text{where } \mathbb{E}[\varepsilon_{0it} s_i] \equiv 0. \tag{13}$$

- Next consider the IV population regression using a full set of state dummies. This is equivalent to

$$Y_{ijt} = \mu_1 + \delta_j + \delta_t + \rho_1 \overline{S}_j + \varepsilon_{1it};\ \text{where } \mathbb{E}[\varepsilon_{1it} \overline{S}_j] \equiv 0, \tag{14}$$

since the projection of individual schooling on a set of state dummies is simply average schooling in each state.

A Little Bit of Econometrics (continued)

- Now consider the estimation of the empirical analogue of equation (2):

$$Y_{ijt} = \mu^* + \delta_j + \delta_t + \pi_0 s_i + \pi_1 \overline{S}_j + \xi_{it}; \text{ where } \mathbb{E}[\xi_{it} s_i] = \mathbb{E}[\xi_{it} \overline{S}_j] \equiv 0. \tag{15}$$

- Then, we have

$$\begin{aligned} \pi_0 &= \rho_1 + \phi(\rho_0 - \rho_1) \\ \pi_1 &= \phi(\rho_1 - \rho_0) \end{aligned} \tag{16}$$

  where

$$\phi = 1/(1 - R^2) > 1,$$

  and R^2 is the first-stage R-squared for the 2SLS estimates in (14).

- Therefore, when $\rho_1 > \rho_0$, for example because there is measurement error in individual schooling, we may find positive external returns even when there are none.

What Can Be Done?

- Instrument for both individual and average schooling, we would solve this problem.
- But what type of instrument?
- Consider the relationship of interest, and for simplicity ignore the time dimension:

$$Y_{ij} = \mu + \gamma_1 \overline{S}_j + \gamma_{2i} s_i + u_j + \varepsilon_i, \tag{17}$$

  which could be estimated by OLS or instrumental variables, to obtain an estimate of γ_1 as well as an average estimate of γ_{2i}, say γ_2^*.
- An alternative way of expressing this relationship is to adjust for the effect of individual schooling by directly rewriting (17):

$$\begin{aligned} Y_{ij} - \gamma_2^* s_i &\equiv \widetilde{Y}_{ij} \\ &= \mu + \gamma_1 \overline{S}_j + [u_j + \varepsilon_i + (\gamma_{2i} - \gamma_2^*) s_i]. \end{aligned} \tag{18}$$

Interpretation

- In this case, instrumental variables estimate of external returns is equivalent to the Wald formula

$$
\begin{aligned}
\gamma_1^{IV} &= \frac{\mathbb{E}[\widetilde{Y}_{ij}|z_i=1]-\mathbb{E}[\widetilde{Y}_{ij}|z_i=0]}{\mathbb{E}[\overline{S}_j|z_i=1]-\mathbb{E}[\overline{S}_j|z_i=0]} \\
&= \gamma_1+\left[\frac{\mathbb{E}[\gamma_{2i}s_i|z_i=1]-\mathbb{E}[\gamma_{2i}s_i|z_i=0]}{\mathbb{E}[s_i|z_i=1]-\mathbb{E}[s_i|z_i=0]}-\gamma_2^*\right] \\
&\qquad \times\left[\frac{\mathbb{E}[s_i|z_i=1]-\mathbb{E}[s_i|z_i=0]}{\mathbb{E}[\overline{S}_j|z_i=1]-\mathbb{E}[\overline{S}_j|z_i=0]}\right].
\end{aligned}
$$

Interpretation (continued)

- This shows that to obtain consistent estimates of external returns to schooling we should set

$$\begin{aligned} \gamma_2^* &= \frac{\mathbb{E}[\gamma_{2i} s_i | z_i = 1] - \mathbb{E}[\gamma_{2i} s_i | z_i = 0]}{\mathbb{E}[s_i | z_i = 1] - \mathbb{E}[s_i | z_i = 0]} \\ &= \frac{\mathbb{E}[(Y_{ij} - \gamma_1 \overline{S}_j) | z_i = 1] - \mathbb{E}[(Y_{ij} - \gamma_1 \overline{S}_j) | z_i = 0]}{\mathbb{E}[s_i | z_i = 1] - \mathbb{E}[s_i | z_i = 0]} \end{aligned} \tag{19}$$

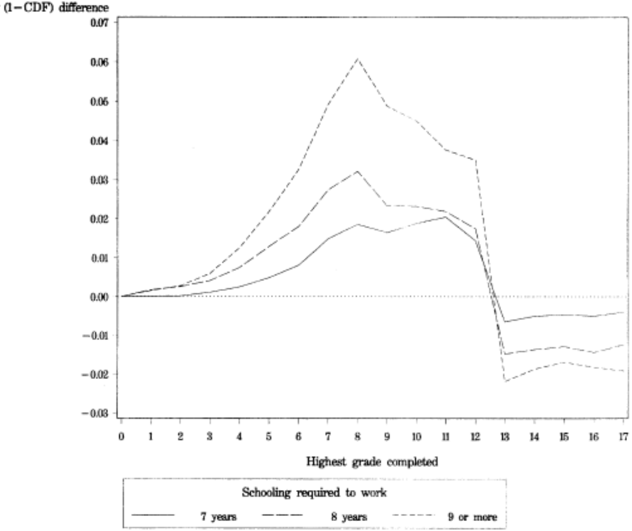
- This is typically not the OLS estimator of the private return, and we should be using some instrument to simultaneously estimate the private return to schooling. The ideal instrument would be one affecting exactly the same people as the compulsory schooling laws.

A Feasible Strategy?

- Quarter of birth instruments might come close to this.
- Since quarter of birth instruments are likely to affect the same people as compulsory schooling laws, adjusting with the quarter of birth estimate, or using quarter of birth dummies as instrument for individual schooling, is the right strategy.
- So the strategy is to estimate an equation similar to (2) or (15) using compulsory schooling laws for average schooling and quarter of birth dummies for individual schooling.

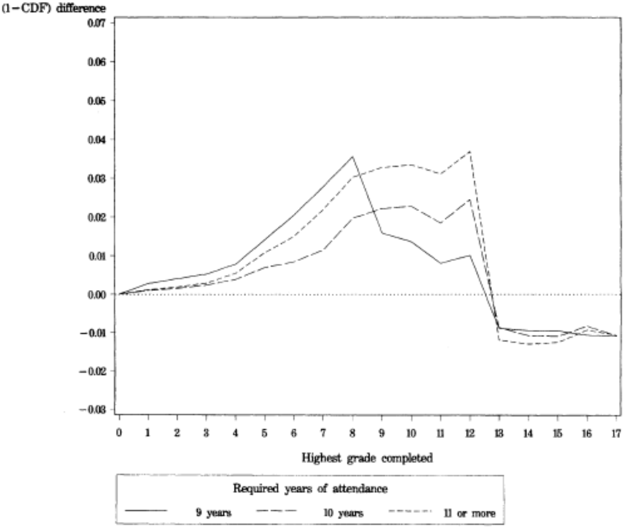
Evidence on Child Labor Laws

Figure 2 CDF DIFFERENCE BY SEVERITY OF CHILD LABOR LAWS



Evidence on Compulsory Schooling Laws

Figure 3 CDF DIFFERENCE BY SEVERITY OF COMPULSORY ATTENDANCE LAWS



Estimates

Table 7 2SLS ESTIMATES OF PRIVATE AND EXTERNAL RETURNS INSTRUMENTS, 1960–1980 AND MEN AGED 40–49

	Individual Schooling Endogenous			
	(1)	*(2)*	*(3)*	*(4)*
Second-Stage Estimates				
Instrument set	QOB & CL	QOB & CA	QOB, CA & CL	QOB, CA & CL, interactions
Private return to schooling	0.074 (0.012)	0.074 (0.012)	0.075 (0.012)	0.060 (0.013)
External return to schooling	0.003 (0.040)	0.017 (0.043)	0.004 (0.035)	0.005 (0.033)

Discussion

- The estimation results from using this strategy in Acemoglu and Angrist (2000) suggest that there are no significant external returns.
- The estimates are typically around 1 percent or less, and statistically not different from zero.
- They also suggest that in the aggregate signaling considerations are unlikely to be very important (at the very least, they do not dominate positive externalities).
- But *caveat*: the first stage and thus the results are significantly weaker is a full set of state times linear trends are included.

An Application to Job Placement Assistance

- Of course, even better if you can get random assignment via some sort of experiment. This is not possible for general externalities, but may be feasible for more specific interactions.
- A nice example is provided by recent work by Crepon et al. (2014).
- A randomized job placement assistance offers across young, educated job-seekers in France, using both randomization across individuals within a labor market and also across labor markets.
- This enables them to estimate both the own effect of job placement assistance and the spillover effect of other workers in the labor market receiving such assistance.

An Application: The Reduced Form

Table 5: Reduced form: Impact of the program, accounting for externalities

	By job type: share of job seekers who are eligible for program					
	Not employed			Not employed, above third quartile		
	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Long term fixed contract					
Assigned to program (β)	0.023^{+++}	0.043^{+++}	0.013	0.040^{++}	0.072^{++}	0.021
	(0.008)	(0.013)	(0.010)	(0.016)	(0.029)	(0.022)
In a Program area (δ)	-0.013	-0.036^{+++}	-0.001	-0.040^{+}	-0.086^{++}	-0.013
	(0.009)	(0.013)	(0.012)	(0.021)	(0.035)	(0.027)
Net effect	0.010	0.007	0.012	0.000	-0.014	0.008
of program assignment ($\beta+\delta$)	(0.008)	(0.011)	(0.011)	(0.019)	(0.031)	(0.024)
Control Mean	0.16	0.131	0.177	0.19	0.161	0.204
	Panel B: Long term employment					
Assigned to program (β)	0.025^{++}	0.037^{++}	0.019	0.019	0.059	0.000
	(0.012)	(0.018)	(0.014)	(0.021)	(0.039)	(0.028)
In a Program area (δ)	-0.021^{+}	-0.043^{++}	-0.010	-0.005	-0.081^{+}	0.033
	(0.013)	(0.020)	(0.018)	(0.023)	(0.047)	(0.032)
Net effect	0.003	-0.006	0.009	0.014	-0.022	0.033
of program assignment ($\beta+\delta$)	(0.011)	(0.018)	(0.016)	(0.019)	(0.037)	(0.026)
Control Mean	0.365	0.372	0.36	0.403	0.408	0.401
Observations	11,806	4,387	7,419	3,066	1,016	2,050

An Application: Structura/Causal Estimates

Table 10: Effect of the treatment, accounting for externalities

	By job type: share of job seekers who are eligible for program					
	Not employed			Not employed, above third quartile		
	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Long term fixed contract					
Program participation (β)	0.054^{+++}	0.095^{+++}	0.030	0.094^{++}	0.176^{++}	0.048
	(0.018)	(0.030)	(0.023)	(0.039)	(0.074)	(0.052)
In a Program area (δ)	-0.014	-0.036^{+++}	-0.001	-0.041^{+}	-0.088^{++}	-0.014
	(0.009)	(0.014)	(0.012)	(0.021)	(0.036)	(0.028)
Net effect	0.040^{+++}	0.060^{+++}	0.029	0.053	0.088	0.035
of program participation ($\beta+\delta$)	(0.014)	(0.023)	(0.019)	(0.033)	(0.062)	(0.042)
Control Mean	0.16	0.131	0.177	0.19	0.161	0.204
	Panel B: Long term employment					
Program participation (β)	0.057^{++}	0.083^{++}	0.044	0.044	0.145	0.000
	(0.027)	(0.041)	(0.034)	(0.049)	(0.096)	(0.064)
In a Program area (δ)	-0.021^{+}	-0.043^{++}	-0.010	-0.005	-0.083^{+}	0.033
	(0.013)	(0.020)	(0.018)	(0.024)	(0.048)	(0.033)
Net effect	0.036	0.040	0.034	0.039	0.063	0.033
of program participation ($\beta+\delta$)	(0.022)	(0.034)	(0.029)	(0.038)	(0.075)	(0.049)
Control Mean	0.365	0.372	0.36	0.403	0.408	0.401
Observations	11,806	4,387	7,419	3,066	1,016	2,050

- Why are the reduced-form and structural estimates different?
- What is the right structural model?