Labor Economics, 14.661. Lectures 3 and 4: Social Mobility, Peer Effects and Human Capital

Daron Acemoglu

MIT

November 2 and 7, 2017.

Evidence on Intergenerational Linkages

- Let us now turn to social mobility.
- Does parental income have an effect on schooling?
- A simple regression

$$\text{schooling} = \text{controls} + \alpha \cdot \log \text{ parental income}$$

- Result: often positive estimates of α.
- But what does positive α mean?

1. Credit constraints: rich parents invest more in schooling (why is this associated with credit constraints?)
2. Children's education may also be a consumption good, so rich parents will "consume" more of this good as well as other goods.
3. The distribution of costs and benefits of education differ across families, and are likely to be correlated with income.

Evidence (continued)

- Include other characteristics to proxy for the costs and benefits of education or for attitudes toward education.
- When parents' education is also included in the regression, the role of income is substantially reduced.
- Conclusion?
- Two considerations:

1. First, parents' income may affect more the quality of education, especially through the choice of the neighborhood in which the family lives.
2. Parental income is often measured with error, and has a significant transitory component, so parental education may be a much better proxy for permanent income than income observations in these data sets.

Social Mobility More Directly

- This motivates a simpler and at some level more interesting approach: measure intergenerational mobility and earnings.
- The typical regression here is

$$\text{log child income} = \text{controls} + \alpha \cdot \text{log parental income} \qquad (1)$$

- Regressions of this sort were first investigated by Becker and Tomes. They found relatively small coefficients, typically in the neighborhood of 0.2.
- This would be particularly striking since there is a significant amount of inheritability of various income-earning characteristics (including IQ). For example, the literature finds a correlation of IQ between parents and offspring between 0.42 and 0.72. (There is also similar evidence from twin studies.)
- Though some of this is because of better education and resources leading to higher IQ of the offspring, it suggests a significant "genetic" inheritance.

Social Mobility Evidence

- Now returning to the above equation, estimates of α around 0.2 mean that if your parents are twice as rich as my parents, you will typically be about 20 percent as rich as me. Your children will be only 4 percent richer than my children!
- With this degree of intergenerational dependence, differences in initial conditions will soon disappear$\longrightarrow$ converges to a relatively "egalitarian" society (does this mean inequality will disappear?)

Interpreting the Evidence

- To elaborate on this, consider the following simple model:

$$\ln y_{it} = \mu + \alpha \ln y_{it-1} + \varepsilon_{it},$$

  y_{it} is the income of t-th generation of dynasty i, and ε_{it} is serially independent disturbance term with variance σ_ε^2.
- Then the long-term (stationary distribution) variance of log income is:

$$\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\alpha^2} \tag{2}$$

  (To derive this, set $\sigma_{y,t-1}^2 = \sigma_{y,t}^2$; why is that the right thing to do?)
- Using the estimate of 0.2 for α, equation (2) implies that the long-term variance of log income will be only about 4 percent higher than σ_ε^2.
- Therefore, the long-run income distribution will largely reflect transitory shocks to dynasties' incomes and skills — not inherited differences. (But inequality could be very large if σ_ε^2 is large.)

Interpretation

- What does this say about credit market problems?
- Persistence of about 0.3 is not very different from what we might expect to result simply from the inheritance of IQ between parents and children, or from the children's adoption of cultural values favoring education from their parents.
- Therefore, relatively small effect of parents income on children's human capital.

Interpretation (continued)

- However, econometric problems biasing α toward zero.
- First, **measurement error**.
- Second, in typical panel data sets (most often the PSID), we observe children at an early stage of their life cycles, where differences in earnings may be less than at later stages.
- Third, income mobility may be very nonlinear, with a lot of mobility among middle income families, but very little at the tails.
- Solon and Zimmerman: dealing with the first two problems increases α to about 0.45 or even 0.55.
- If $\alpha = 0.55$, then $\sigma_y^2 \approx 1.45 \cdot \sigma_\varepsilon^2$ instead of $\sigma_y^2 \approx 1.04 \cdot \sigma_\varepsilon^2$ with the coefficient of 0.2—substantial difference.

Limitations

- One limitations about the functional forms. The linearity (or log linearity) rules out the possibility of "mobility traps" in some part of the distribution.
- Cooper, Durlauf and Johnson find that there are important heterogeneities.
- In particular, using the PSID, they find that while mobility estimate for the entire sample is 0.34, focusing on families living the poorest 33%, this coefficient increases to 0.46.
- But perhaps the most important limitation of the earlier work was data quality.

Some Recent Evidence

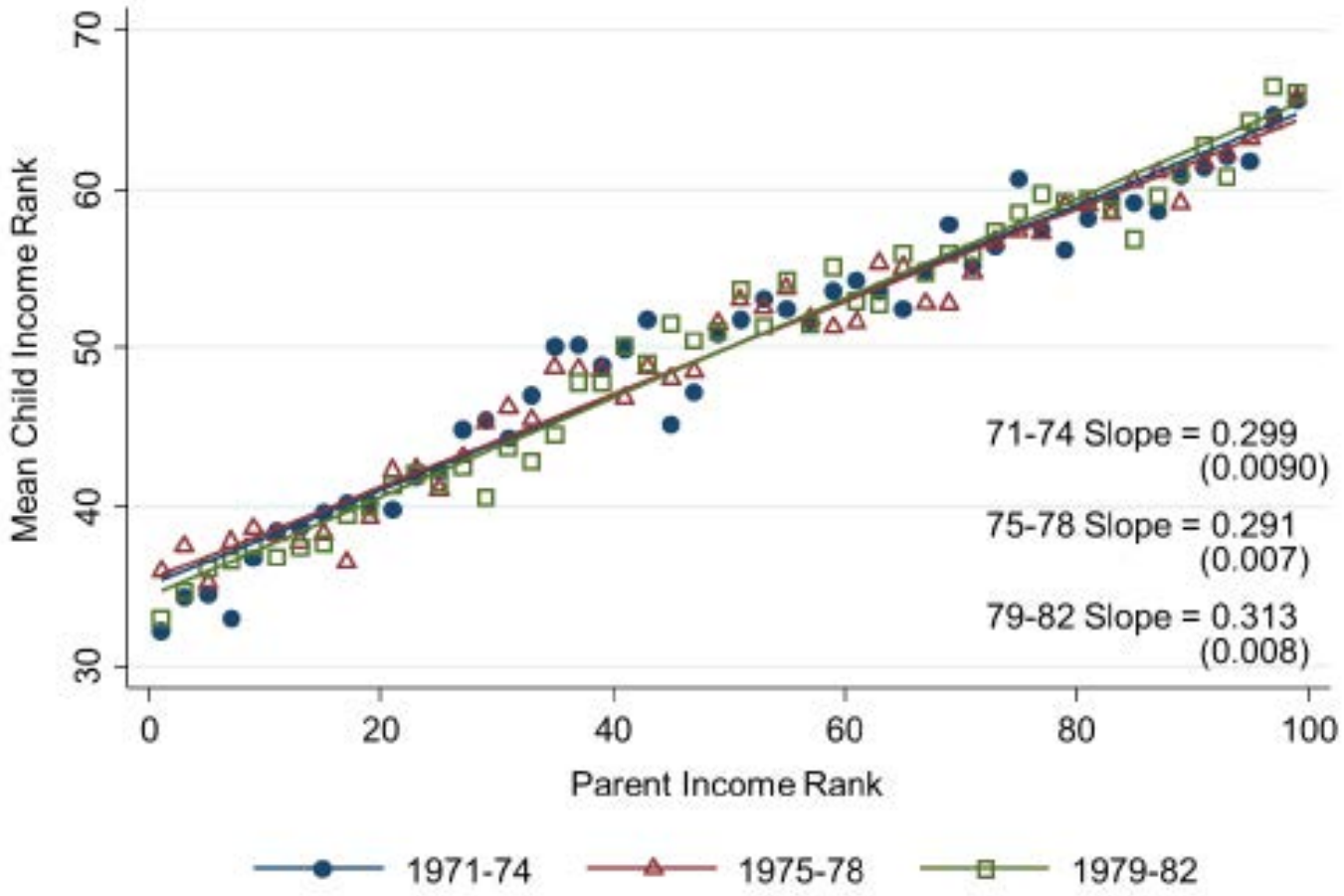
- Recent work by Chetty, Hendren, Klein and Saez uses matched Social Security records and tax data to have a better picture of social mobility for cohorts born since the 1970s.
- They report on the log-log correlation coefficient, but also two other measures:
  1. Estimates of the parameter α' from the regression

$$\text{rank}_{it} = \mu + \alpha' \times \text{rank}_{it-1} + \varepsilon_{it}$$

     where rank_{it} is the rank of family i the income distribution of cohort t
  2. Absolute mobility at the 25th percentile, y'_{25}, computed as the expected rank of children from families at or below the 25th percentile of the distribution.

Some Recent Evidence (continued)

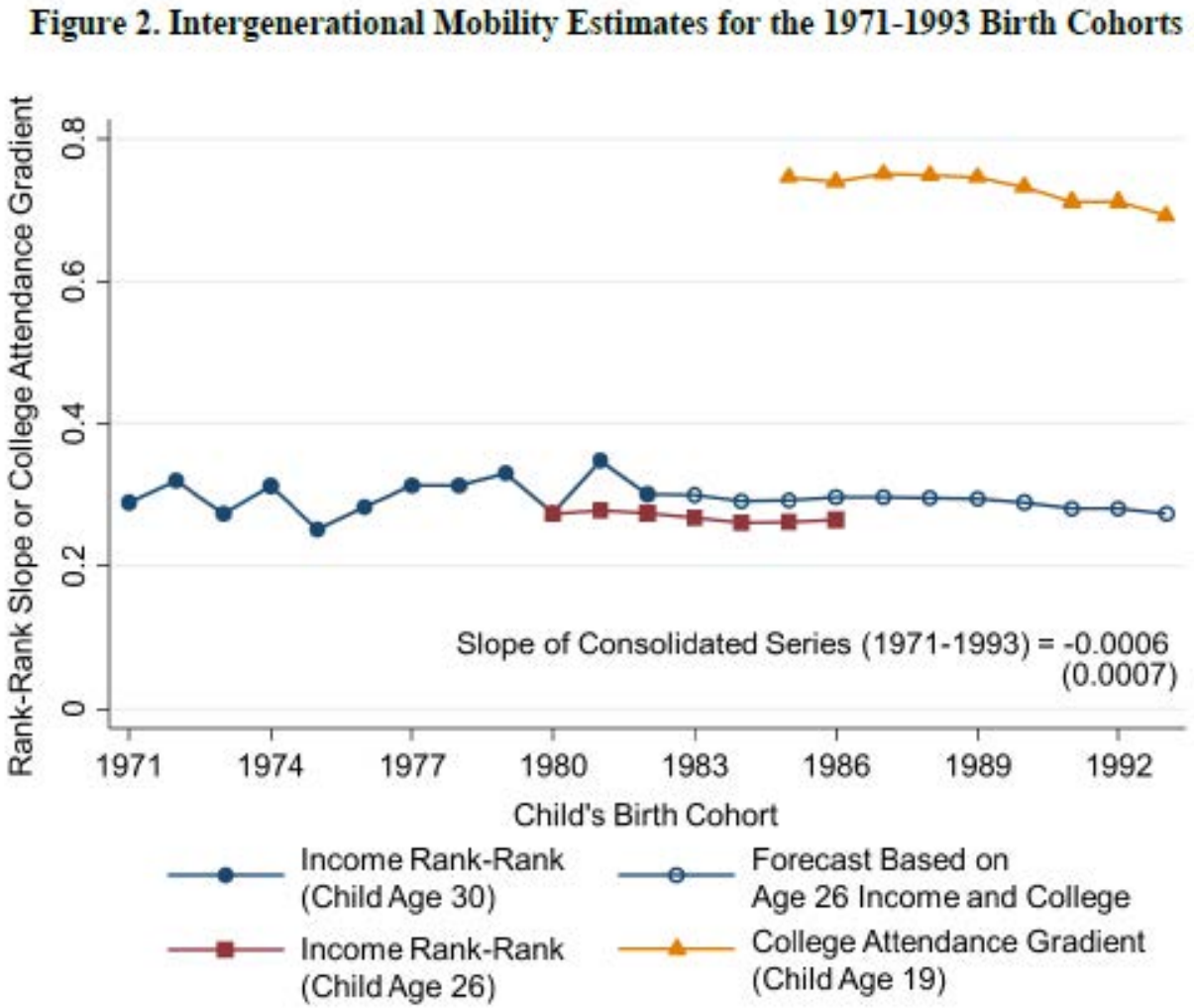
Figure 1. Child Income Rank vs. Parent Income Rank by Birth Cohort



Courtesy of Raj Chetty, Nathanial Hendren, Patrick Kline, Emmanual Saez, and Nicholas Turner. Used with permission.

Some Recent Evidence (continued)

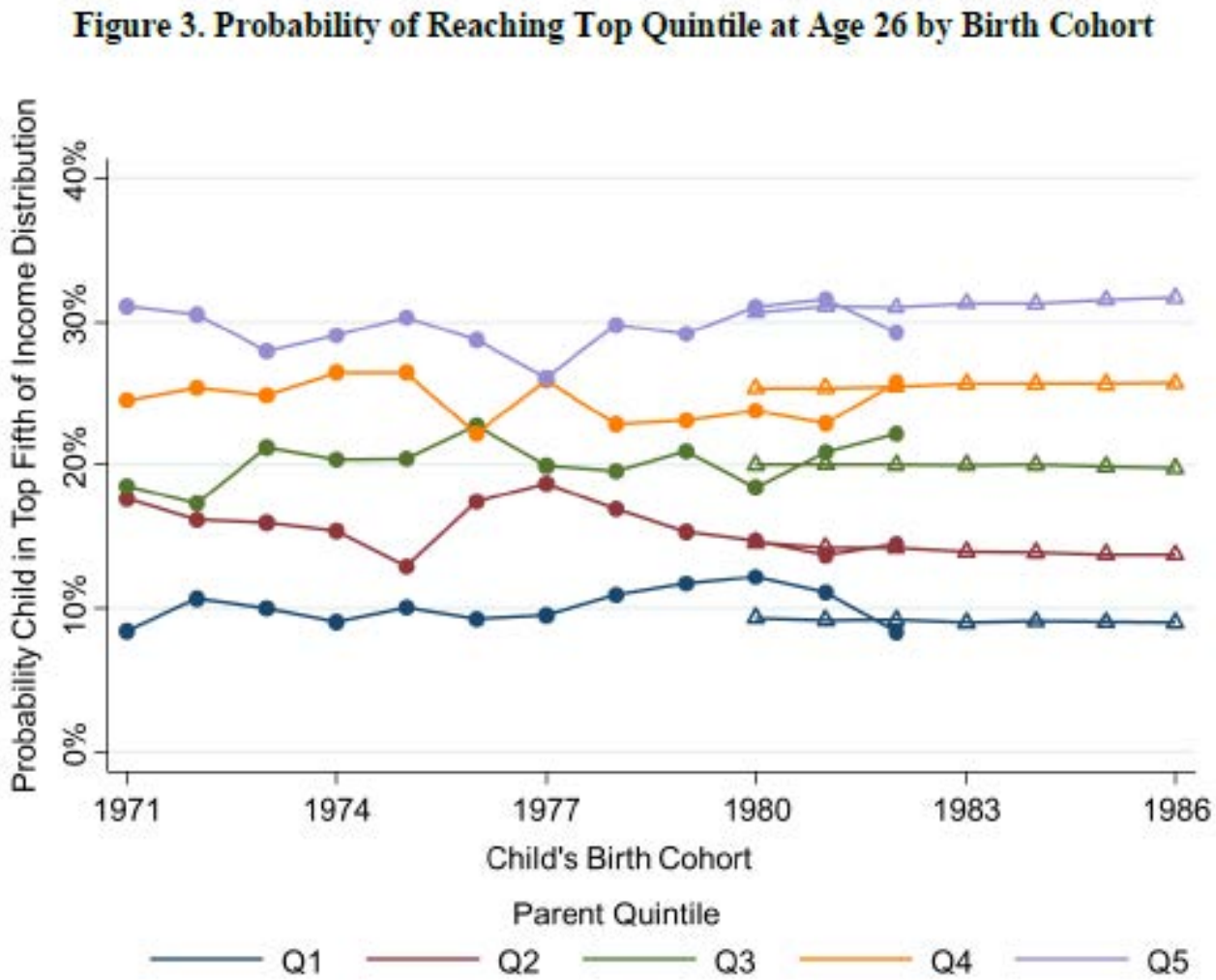
- In fact, little change in mobility:



Courtesy of Raj Chetty, Nathanial Hendren, Patrick Kline, Emmanual Saez, and Nicholas Turner. Used with permission.

Some Recent Evidence (continued)

- Allowing for nonlinearities



Courtesy of Raj Chetty, Nathanial Hendren, Patrick Kline, Emmanual Saez, and Nicholas Turner. Used with permission.

Determinants of Social Mobility

- What could determine social mobility?
  - Credit market constraints and inequality, the "Great Gatsby curve" (as we will see next).
  - Peer effects and residential/social sorting (as we will also see next).
  - Inequality and rewards to success (as we will also discuss later).
  - Policy.
  - Sociological factors.
- Before studying the theories, let's look at some of the data

Patterns of Social Mobility

TABLE II
Intergenerational Mobility Estimates at the National Level

Child's outcome	Parent's Income Def.	Sample: Core sample (1)	Male children (2)	Female children (3)	Married parents (4)	Single parents (5)	1980-1985 cohorts (6)
1. Log family income (excluding zeros)	Log family income	0.344 (0.0004)	0.349 (0.0006)	0.342 (0.0005)	0.303 (0.0005)	0.264 (0.0008)	0.316 (0.0003)
2. Log family income (recoding zeros to $1)	Log family income	0.618 (0.0009)	0.697 (0.0013)	0.540 (0.0011)	0.509 (0.0011)	0.528 (0.0020)	0.580 (0.0006)
3. Log family income (recoding zeros to $1000)	Log family income	0.413 (0.0004)	0.435 (0.0007)	0.392 (0.0006)	0.358 (0.0006)	0.322 (0.0009)	0.380 (0.0003)
4. Family income rank	Family income rank	0.341 (0.0003)	0.336 (0.0004)	0.346 (0.0004)	0.289 (0.0004)	0.311 (0.0007)	0.323 (0.0002)
5. Family income rank	Top parent income rank	0.312 (0.0003)	0.307 (0.0004)	0.317 (0.0004)	0.256 (0.0004)	0.253 (0.0006)	0.296 (0.0002)
6. Individual income rank	Family income rank	0.287 (0.0003)	0.317 (0.0004)	0.257 (0.0004)	0.265 (0.0004)	0.279 (0.0007)	0.286 (0.0002)
7. Individual earnings rank	Family income rank	0.282 (0.0003)	0.313 (0.0004)	0.249 (0.0004)	0.259 (0.0004)	0.273 (0.0007)	0.283 (0.0002)
8. College Attendance	Family income rank	0.675 (0.0005)	0.708 (0.0007)	0.644 (0.0007)	0.641 (0.0006)	0.663 (0.0013)	0.678 (0.0003)
9. Teenage birth (females only)	Family income rank	-0.298 (0.0006)			-0.231 (0.0007)	-0.322 (0.0016)	-0.285 (0.0004)
Number of observations		9,867,736	4,935,804	4,931,066	6,854,588	3,013,148	20,520,588

Courtesy of Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Used with permission.

Determinants of Social Mobility: The Great Gatsby Curve

TABLE VIII
Income Inequality and Intergenerational Mobility: The "Great Gatsby" Curve

Dep. Var.:	Across CZs within the U.S.					Across Countries		
	Upward mobility				Relative mobility	Log-Log Elasticity 1985 Inequality		Log-Log Elasticity 2005 Inequality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gini Coefficient	-0.578							
	(0.093)							
Gini Bottom 99%		-0.634	-0.624		0.476	0.72	0.62	0.78
		(0.090)	(0.113)		(0.088)	(0.21)	(0.27)	(0.27)
Top 1% Income Share		-0.123	0.029		-0.032		0.17	-0.11
		(0.035)	(0.039)		(0.032)		(0.27)	(0.26)
Fraction Between p25 and p75				0.679				
				(0.111)				
Urban Areas Only			x					
R-Squared	0.334	0.433	0.380	0.462	0.224	0.518	0.536	0.531
Observations	709	709	325	709	709	13	13	12

Courtesy of Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Used with permission.

Determinants of Social Mobility: Segregation

TABLE IX
Correlates of Intergenerational Mobility: Comparing Alternative Hypotheses

Dep. Var.:	Absolute Upward Mobility				Relative Mobility		Absolute Upward Mobility	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Racial Segregation	-0.086	-0.111	-0.042	-0.109	0.190	0.281	-0.166	
	(0.028)	(0.020)	(0.091)	(0.045)	(0.044)	(0.047)	(0.032)	
Gini Bottom 99%	-0.042	-0.021	0.069	0.012	0.038	0.338	-0.307	
	(0.062)	(0.038)	(0.093)	(0.113)	(0.076)	(0.093)	(0.063)	
High School Dropout Rate	-0.152	-0.132	-0.262	-0.155	0.019	0.166	-0.282	
	(0.052)	(0.028)	(0.100)	(0.086)	(0.059)	(0.051)	(0.059)	
Social Capital Index	0.291	0.109	0.120	0.269	0.044	0.028	0.304	
	(0.060)	(0.054)	(0.086)	(0.068)	(0.061)	(0.074)	(0.069)	
Fraction Single Mothers	-0.489	-0.444	-0.537	-0.510	0.553			-0.791
	(0.072)	(0.073)	(0.114)	(0.103)	(0.063)			(0.088)
Fraction Black								0.035
								(0.073)
State Fixed Effects		x						
Population Weighted			x					
Urban Areas Only				x				
R-Squared	0.698	0.847	0.441	0.605	0.459	0.329	0.596	0.584
Observations	709	709	709	325	709	709	709	709

Courtesy of Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Used with permission.

Modeling Social Mobility

- Let us next turn to modeling social mobility.
- I will present three sets of models emphasizing different aspects:
  1. The role of credit constraints and inequality in social mobility.
  2. The role of rewards to success in social mobility.
  3. The role of community structure/segregation/school structure in social mobility.
- The last topic will then act as a segue into the discussion of peer effects more generally.

Imperfect Credit Markets and Mobility

- Simplified version Galor and Zeira, 1993.
- Each individual still lives for two periods.
- In youth, he can either work or acquire education.
- Utility function of each individual is

$$(1-\delta)\log c_i(t)+\delta\log b_i(t),$$

- Budget constraint is

$$c_i(t)+b_i(t)\leq y_i(t),$$

- Preferences of the "warm glow" form, depending on monetary bequest rather than level of education expenditures.
- Logarithmic formulation ensures a constant saving rate δ.

Imperfect Credit Markets (continued)

- Education: binary outcome, and educated (skilled) workers earn wage w_s while uneducated workers earn w_u.
- Expenditure to become skilled is h, and not earn the unskilled wage w_u during the first period.
- Binary education: introduces *a nonconvexity*.
- Imperfect capital markets: some amount of monitoring required for loans to be paid back.
- Cost of monitoring: *wedge* between the borrowing and the lending rates.
- Linear savings technology, which fixes lending rate at some constant r, but borrowing rate is $i > r$.
- Also assume investment in skills is socially efficient:

$$w_s - (1 + r) h > w_u (2 + r) \tag{3}$$

Imperfect Credit Markets (continued)

- Implies investment in human capital is profitable when financed at the lending rate r.
- Consider an individual with wealth x.
  - If $x \geq h$, assumption (3) implies that individual will invest in education.
  - If $x < h$, then whether it is profitable to invest in education will depend on wealth of individual and borrowing interest rate, i.
- Utility of this agent (with $x < h$), when he invests in education:

$$\begin{aligned} U_s(x) &= \log\left(w_s + (1+i)(x-h)\right) + \log\left(1-\delta\right)^{1-\delta}\delta^{\delta} \\ b_s(x) &= \delta\left(w_s + (1+i)(x-h)\right), \end{aligned}$$

Imperfect Credit Markets (continued)

- When he chooses not to invest:

$$\begin{aligned} U_u(x) &= \log\left((1+r)(w_u+x)+w_u\right)+\log\left(1-\delta\right)^{1-\delta}\delta^{\delta} \\ b_u(x) &= \delta\left((1+r)(w_u+x)+w_u\right). \end{aligned}$$

- Individual likes to invest in education if and only if:

$$x \geq f \equiv \frac{(2+r)w_u+(1+i)h-w_s}{i-r}$$

- *Equilibrium correspondence* describing equilibrium dynamics is

$$x(t+1)=\begin{cases} b_u=\delta\left((1+r)(w_u+x(t))+w_u\right) & \text{if } x(t)<f \\ b_s=\delta\left(w_s+(1+i)(x(t)-h)\right) & \text{if } h>x(t)\geq f \\ b_n=\delta\left(w_s+(1+r)(x(t)-h)\right) & \text{if } x(t)\geq h \end{cases} \tag{4}$$

Imperfect Credit Markets (continued)

- Equilibrium dynamics: (4) describes both the behavior of the wealth of each individual and the behavior of the wealth distribution in the economy ("Markovian").
- Define x^* as the intersection of the equilibrium curve (4) with the 45 degree line, when the equilibrium correspondence is steeper than the 45 degree line.
- Such an intersection will exist when the borrowing interest rate, i, is large enough.

Imperfect Credit Markets (continued)

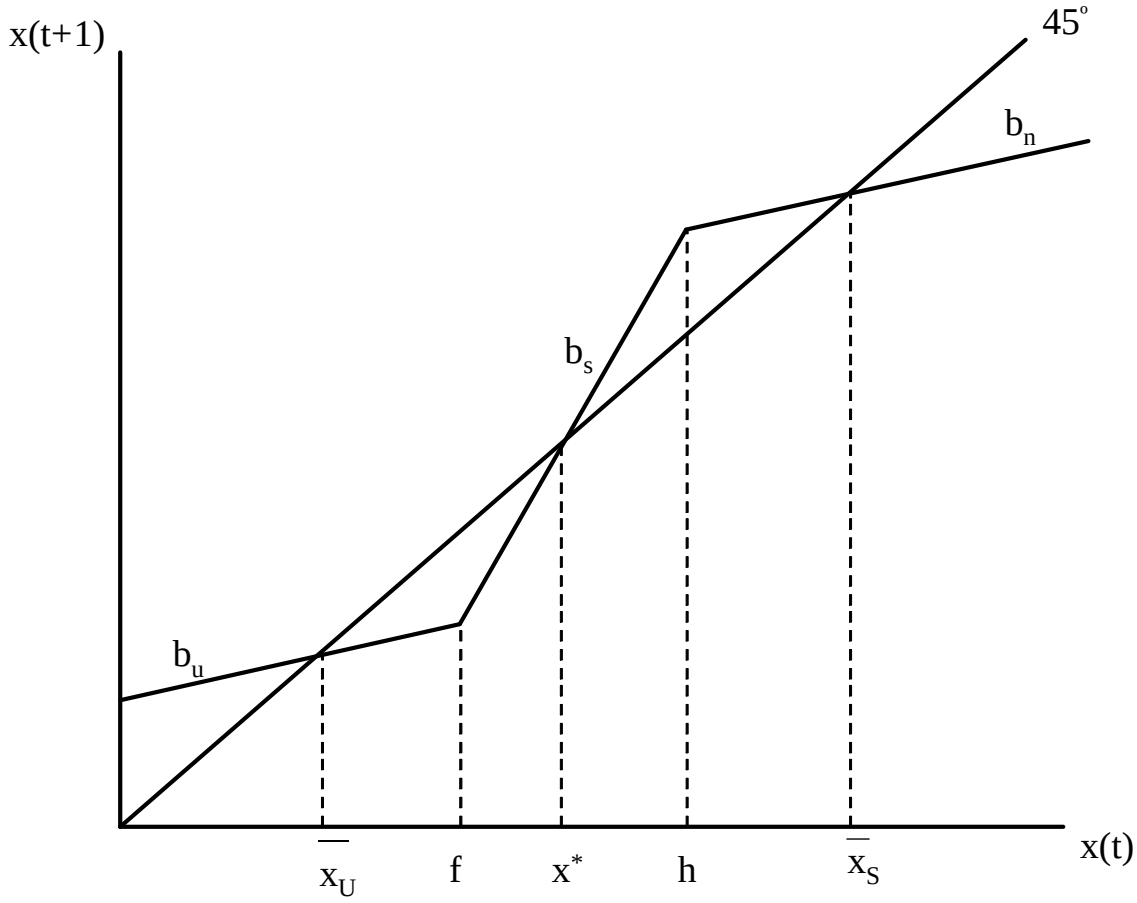


Figure: Multiple steady-state equilibria in the Galor and Zeira model.

Imperfect Credit Markets (continued)

- All individuals with $x(t) < x^*$ converge to the wealth level $\bar{x}_U$, while all those with $x(t) > x^*$ converge to the greater wealth level $\bar{x}_S$.
- "Poverty trap," attracts agents with low initial wealth.
- Distribution of income again has a potentially first-order effect, but it is straightforward to construct examples where an increase inequality can lead to either worse or better outcomes.
- Implications of financial development: i smaller given r.
  - More agents will escape the poverty trap, and poverty trap may not exist

Shortcomings of the Model

- No social mobility in the long run.
- Does inequality lead to lower or greater social mobility?
  - The answer is unclear: if by increasing inequality, you push more people above the threshold f, then you increase mobility (and efficiency), but if you push more people below the threshold, then the opposite happens.
- Theoretically, this is a partial equilibrium model:
  - Models in which prices determined in general equilibrium affect wealth (income) dynamics may be more relevant (and also may have some additional robust features as we describe next).

Shortcomings of the Model (continued)

- Multiple steady states here may not be robust to addition of noise in income dynamics—long-run equilibrium then corresponds to a stationary distribution of human capital levels.
- In particular, suppose ε is a random variable, and change the law of motion of wealth to:

$$x(t+1) = \begin{cases} b_u = \delta\left((1+r)\left(w_u + x(t)\right) + w_u\right) + \varepsilon & \text{if } x(t) < f \\ b_s = \delta\left(w_s + (1+i)\left(x(t) - h\right)\right) + \varepsilon & \text{if } h > x(t) \geq f \\ b_n = \delta\left(w_s + (1+r)\left(x(t) - h\right)\right) + \varepsilon & \text{if } x(t) \geq h \end{cases}$$

- What does the long-run (stationary) distribution of wealth and human capital look like in this case?

Generalizations

- Now suppose that with probability q, an individual does not have the ability to acquire skills (which we denote by $\sigma(t) = 0$).
- Then the equilibrium correspondence becomes a stochastic correspondence, taking the form

$$x(t+1) = \begin{cases} \delta\left(w_u + (1+r)\left(w_u + x(t)\right)\right) & \text{if } x(t) < f \\ \delta\left(w_u + (1+i)\, x(t)\right) & \text{if } h > x(t) \geq f \ \& \ \sigma(t) = 0 \\ \delta\left(w_s + (1+i)\left(x(t) - h\right)\right) & \text{if } h > x(t) \geq f \ \& \ \sigma(t) = 1 \\ \delta\left(w_u + (1+r)\left(x(t) + w_u\right)\right) & \text{if } x(t) \geq h \ \& \ \sigma(t) = 0 \\ \delta\left(w_s + (1+r)\left(x(t) - h\right)\right) & \text{if } x(t) \geq h \ \& \ \sigma(t) = 1 \end{cases}$$

- What does the limiting distribution look like in this case? Does it generate social mobility? Is that the right type of social mobility?

Stratification, Human Capital and Inequality

- More general (Benabou (1996)): study dynamics of inequality and its costs for efficiency of production resulting from its effect on human capital.
- Aggregate output in the economy at time t:

$$Y(t) = H(t),$$

- $H(t)$ is an aggregate of the human capital of all the individuals in the society.
- Normalizing total population to 1 and denoting the distribution of human capital at time t by $\mu_t(h)$:

$$H(t) \equiv \left(\int_0^{\infty} h^{\frac{\sigma-1}{\sigma}} d\mu_t(h) \right)^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

Stratification, Human Capital and Inequality (continued)

- σ=degree of complementarity or substitutability in the human capital of different individuals.
  - $\sigma \longrightarrow \infty$: perfect substitutes and $H(t)$ is simply equal to the mean of the distribution.
  - $\sigma \in (0, \infty)$: complementarity between the human capital levels of different individuals.
- Effect of heterogeneity of human capital on aggregate productivity, for given mean level, is most severe when σ is close to 0.
- But formulation is general enough to allow for the case in which greater inequality is productivity-enhancing.
  - Defined for $\sigma < 0$ as well: in this case, greater inequality for a given mean level increases $H(t)$ and productivity.
  - Extreme case $\sigma \longrightarrow -\infty$, $H(t) = \max_i \{h_i(t)\}$.
- Focus on potential costs of inequality on human capital: $\sigma \geq 0$.

Stratification, Human Capital and Inequality (continued)

- Then, mean preserving spread of the human capital distribution μ will lead to a lower level of $H(t)$
- Human capital of an individual from dynasty i at time $t+1$:

$$h_i(t+1) = \xi_i(t) B (h_i(t))^{\alpha} (N_i(t))^{\beta} (H(t))^{\gamma}, \tag{6}$$

- B is a positive constant, $h_i(t)$ human capital of parent, $\xi_i(t)$ random shock, and $N_i(t)$ "average" human capital in the neighborhood.
- Assume neighborhood human capital is also a constant elasticity of substitution aggregator,with an elasticity ε:

$$N_i(t) \equiv \left(\int_0^{\infty} h^{\frac{\varepsilon-1}{\varepsilon}} d\mu_t^i(h) \right)^{\frac{\varepsilon}{\varepsilon-1}},$$

- $\mu_t^i(h)$ denotes the distribution of human capital in the neighborhood of individual i at time t.

Stratification, Human Capital and Inequality (continued)

- $\varepsilon \in (0, \infty)$: mean preserving spread of neighborhood human capital will reduce the human capital of all the offsprings.
- Plausible if presence of some low human capital children will slow down learning by those with higher potential (one "bad apple" will spoil the pack)—We will discuss this in greater detail next.
- This suggests segregation of high and low human capital parents might be beneficial for human capital accumulation—though we will see why this may not follow.
- Multiplicative structure in (6): tractable evolution of human capital if initial distribution of human capital and the $\xi(t)$s are log normal.

Stratification, Human Capital and Inequality (continued)

- Assume:

$$\begin{aligned} \ln h_i(0) &\sim \mathcal{N}\left(m_0, \Delta_0^2\right) \\ \ln \xi_i(t) &\sim \mathcal{N}\left(-\frac{\omega^2}{2}, \omega^2\right), \end{aligned} \tag{7}$$

  where $\mathcal{N}$ denotes the normal distribution.
- The draws of $\xi_i(t)$ are independent across time and across individuals.
- Distribution of $\ln \xi$ is assumed to have mean $-\omega^2/2$ so that ξ has a mean equal to 1 (that is independent of its variance).

Stratification, Human Capital and Inequality (continued)

- Thus distribution of human capital within every generation will remain log normal:

$$\ln h_i(t) \sim \mathcal{N}\left(m_t, \Delta_t^2\right), \tag{8}$$

  for some endogenous mean m_t and variance Δ_t, which will depend on parameters and the organization of society.
- Analysis of output and inequality dynamics boils down to characterizing the law of motion of m_t and Δ_t.
- Two alternative organizations: *full segregation* and *full mixing*.
- Full segregation: each parent is in a neighborhood with identical parents.
  - Because the neighborhood human capital is the same as the parent's human capital, (6) becomes

$$h_i(t+1) = \xi_i(t) B \left(h_i(t)\right)^{\alpha+\beta} \left(H(t)\right)^{\gamma}, \tag{9}$$

Heterogeneity, Stratification and the Dynamics of Inequality VII

- Full mixing: each neighborhood is a mirror image of the entire society.
  - Thus for all neighborhoods $N^i(t) = N(t) \equiv \left(\int_0^\infty h^{\frac{\varepsilon-1}{\varepsilon}} d\mu_t(h) \right)^{\frac{\varepsilon}{\varepsilon-1}}$, where μ_t refers to the aggregate distribution.
  - Accumulation equation:

$$h_i(t+1) = \xi_i(t) B (h_i(t))^\alpha N(t)^\beta H(t)^\gamma . \tag{10}$$

- Intuition above: segregation might be preferable.
- But this may not be entirely accurate:
  - lack of segregation may reduce long-run inequality leading to better economic outcomes.

Stratification, Human Capital and Inequality (continued)

- With full segregation:

$$\begin{aligned} m_{t+1} &= \ln B - \frac{\omega^2}{2} + (\alpha+\beta+\gamma)\, m_t + \gamma\left(\frac{\sigma-1}{\sigma}\right)\frac{\Delta_t^2}{2} \quad (11) \\ \Delta_{t+1}^2 &= (\alpha+\beta)^2\, \Delta_t^2 + \omega^2 \end{aligned}$$

- With full integration:

$$\begin{aligned} \hat{m}_{t+1} &= \ln B - \frac{\omega^2}{2} + (\alpha+\beta+\gamma)\, \hat{m}_t + \left[\begin{array}{c} \gamma\left(\frac{\sigma-1}{\sigma}\right) \\ +\beta\left(\frac{\varepsilon-1}{\varepsilon}\right) \end{array}\right]\frac{\hat{\Delta}_t^2}{2} \quad (12) \\ \hat{\Delta}_{t+1}^2 &= \alpha^2 \hat{\Delta}_t^2 + \omega^2, \end{aligned}$$

- $\hat{m}_t$ and $\hat{\Delta}_t^2$ refer to the values of the mean and the variance of the distribution under full integration.

Stratification, Human Capital and Inequality (continued)

- Note there will be persistence in the distribution of human capital (autoregressive nature of the behavior of m_t):
  - human capital of offsprings reflects that of parents (either through direct effect or through neighborhood and aggregate spillovers).
- Dispersion of the parents' human capital affects the mean of the distribution.
  - when $\sigma < 1$ or when $\varepsilon < 1$, so degree of complementarity in the aggregate or the neighborhood spillovers is high, greater dispersion reduces the mean of the distribution of human capital.

Stratification, Human Capital and Inequality (continued)

- Behavior of the variance of the distribution:
  - With full segregation, costs of heterogeneity resulting from neighborhood spillovers are avoided.
  - But variance of log human capital is more persistent than under full integration.
  - In paticular, when $\varepsilon < 1$, starting with the same m_t and Δ_t:

$$\hat{m}_{t+1} < m_{t+1} \text{ and } \hat{\Delta}^2_{t+1} < \Delta^2_{t+1},$$

  - Thus human capital in the next period is higher under segregation.
  - But inequality is also higher and from (5) inequality has efficiency costs.
- To determine which effect dominates, first find the long-run level of inequality under segregation and integration.

Stratification, Human Capital and Inequality (continued)

- Equations (11) and (12) imply these variances are given by:

$$\Delta_{\infty}^{2} = \frac{\omega^{2}}{1-(\alpha+\beta)^{2}} > \hat{\Delta}_{\infty}^{2} = \frac{\omega^{2}}{1-\alpha^{2}},$$

- i.e., greater inequality of human capital and income with segregation of neighborhoods.
- Mean of the two distributions will also be different: suppose $\alpha+\beta+\gamma<1$, so steady state distribution exists under both full segregation and full integration.

Stratification, Human Capital and Inequality (continued)

- Then:

$$m_\infty = \frac{1}{1-(\alpha+\beta+\gamma)}\left[\ln B - \frac{\omega^2}{2} + \gamma\left(\frac{\sigma-1}{\sigma}\right)\frac{\omega^2}{2\left(1-(\alpha+\beta)^2\right)}\right],$$

and

$$\hat{m}_\infty = \frac{1}{1-(\alpha+\beta+\gamma)}\left[\begin{array}{c}\ln B - \frac{s^2}{2} \\ +\left[\gamma\left(\frac{\sigma-1}{\sigma}\right)+\beta\left(\frac{\varepsilon-1}{\varepsilon}\right)\right]\frac{s^2}{2(1-\alpha^2)}\end{array}\right].$$

- Mean level of human capital in the long run may be higher or lower under full integration or full segregation.

Stratification, Human Capital and Inequality (continued)

- Using the production function, taking logs on both sides of (5) and using log normality:

$$\ln Y(t) = \ln H(t) = m_t + \left(\frac{\sigma - 1}{\sigma}\right)\frac{\Delta_t^2}{2},$$

- Thus long-run income levels under full segregation and full integration are:

$$\begin{aligned} \ln Y(\infty) &= m_\infty + \left(\frac{\sigma - 1}{\sigma}\right)\frac{\Delta_\infty^2}{2} \\ \ln \hat{Y}(\infty) &= \hat{m}_\infty + \left(\frac{\sigma - 1}{\sigma}\right)\frac{\hat{\Delta}_\infty^2}{2}. \end{aligned}$$

Depending on parameters long-run income levels may be higher or lower under full segregation and full integration.

Stratification, Human Capital and Inequality (continued)

- This framework highlights various different costs arising from income inequality.
- But somewhat "reduced form": what are the micro interactions leading to segregation and also costs of inequality?
- I will not discuss costs of inequality given our focus, but these could be better micro-founded, though at the end we do not have great evidence that there are indeed productivity costs from greater inequality of human capital in the economy as a whole (as opposed to within a given firm etc.).

Peer Effects

- How segregation of families by socioeconomic status (and organization of schools) affects human capital accumulation is related to the literature on peer effects.
- There may be "technological" human capital externalities in the classroom (e.g., learning from peers, teamwork).
- There may also be other, more "sociological" effects: children growing up in different areas and with different peers may choose different role models.
- But important theoretical and empirical challenges in understanding and estimating peer effects.

Segregation and Mixing

- Let us return to the same question posed above: from the viewpoint of "human capital production efficiency" is it better to have segregation or mixing of students by different abilities/achievements/social economic backgrounds?
- The basic issue here is equivalent to an *assignment problem.*
- The general principle in assignment problems, such as Becker's famous model of marriage, is that if inputs from the two parties are "*complementary,*" there should be assortative matching, that is the highest quality individuals should be matched together.
- In the context of schooling, assortative matching implies that children with better characteristics will be segregated in their own schools, and children with worse characteristics should go to separate schools.
- This practically means segregation along income lines, since often children with "better characteristics" are those from better parental backgrounds, while children with worse characteristics are often from lower socioeconomic backgrounds

Segregation and Mixing (continued)

- Though exceptions exist, it is natural to assume that there are *positive externalities* in this context: higher human capital pupils create positive learning/teamwork/role model externalities on their classmates.
- Also, richer individuals live in more expensive neighborhoods, generating greater tax revenue and thus schools in such neighborhoods tend to have access to greater resources.
- But this does not answer the question of whether these inputs are complements or substitutes.

Segregation and Mixing (continued)

- A potential confusion in the literature (especially in the applied literature): deducing complementarity from the fact that in equilibrium we do observe segregation;
  - e.g., rich parents sending their children to private schools with other children from rich parents, or living in suburbs and sending their children to suburban schools, while poor parents live in ghettos and children from disadvantaged backgrounds go to school with other disadvantaged children in inner cities.
- This reasoning is often used in discussions of *Tiebout competition*, together with the argument that allowing parents with different characteristics/tastes to sort into different neighborhoods will often be efficient.

Segregation and Mixing (continued)

- The underlying idea can be illustrated using the following simple model.
- Suppose that schools consist of two kids, and denote the parental background (e.g., home education or parental expenditure on non-school inputs) of kids by e, and the resulting human capitals by h.
- Suppose

$$h_i = f(e_i, e_{-i}), \tag{13}$$

And this implies of course

$$h_1 = f(e_1, e_2) \text{ and } h_2 = f(e_2, e_1).$$

- Throughout, we will assume positive externalities:

$$\partial h_1/\partial e_2 = \partial f(e_1, e_2)/\partial e_2 > 0 \text{ and } \partial h_2/\partial e_1 = \partial f(e_2, e_1)/\partial e_1 > 0.$$

Segregation and Mixing (continued)

- More important than the positive first derivatives are the *cross-partial derivatives*—or whether the education production function exhibits *supermodularity*.
- Suppose first that cross-partial derivatives are positive: i.e.,

$$\frac{\partial^2 f(e_1, e_2)}{\partial e_1 \partial e_2} > 0.$$

Example

$$\begin{aligned} h_1 &= f(e_1, e_2) = e_1^{\alpha} e_2^{1-\alpha} \\ h_2 &= f(e_2, e_1) = e_1^{1-\alpha} e_2^{\alpha} \end{aligned}$$

where $\alpha > 1/2$ so that own characteristic matters more than a year's characteristic.

Segregation and Mixing (continued)

- This type of supermodularity implies that parental backgrounds are complementary, and each kid's human capital will depend mostly on his own parent's background, but also on that of the other kid in the school.
- For example, it may be easier to learn or be motivated when other children in the class are also motivated.
- We can think of this as the "bad apple" theory of classroom: one bad kid in the classroom brings down everybody (Lazear).

Digression

- Notice an important feature of the way we wrote (13) linking the outcome variables, h_1 and h_2, to predetermined characteristics of children e_1 and e_2, which creates a direct analogy with the human capital externalities discussed above.
- However, this may simply be the reduced form of that somewhat different model, for example,

$$\begin{aligned} h_1 &= H_1(e_1, h_2) \\ h_2 &= H_2(e_2, h_1) \end{aligned} \tag{14}$$

  whereby each individual's human capital depends on his own background and the human capital choice of the other individual.
- Although in reduced form (13) and (14) are very similar, they provide different interpretations of peer group effects, and econometrically they pose different challenges, which we will discuss below.

Segregation and Mixing (continued)

The complementarity in the human capital production function has two implications:

1. It is socially efficient, in the sense of maximizing the sum of human capitals, to have parents with good backgrounds to send their children to school with other parents with good backgrounds.
    - This follows simply from the definition of complementarity, positive cross-partial derivative, which is clearly verified by the production functions in (13).
2. It will also be an equilibrium outcome that parents will do so.

Segregation and Mixing (continued)

- To see that segregation is an equilibrium, suppose that we have a situation in which there are two sets of parents with background e_l and $e_h > e_l$.
- Suppose that there is mixing.
- Now the marginal willingness to pay of a parent with the high background to be in the same school with the child of another high-background parent, rather than a low-background student, is

$$f(e_h, e_h) - f(e_h, e_l).$$

Segregation and Mixing (continued)

- Instead, the marginal willingness to pay of a low-background parent to stay in the school with the high-background parents is

$$f(e_l, e_h) - f(e_l, e_l).$$

- The definition of supermodularity is that

$$f(e_h, e_h) + f(e_l, e_l) > f(e_h, e_l) + f(e_l, e_h),$$

  and this immediately follows from positive cross-partial derivatives, $\partial^2 f(e_1, e_2)/\partial e_1 \partial e_2 > 0$.
- Thus the willingness to pay of high-background parents always exceeds that of low-background parents.
- Therefore, the high-background parent can always outbid the low-background parent for the privilege of sending his children to school with other high-background parents.
- Thus with profit maximizing schools, segregation will arise as the outcome.

Segregation and Mixing (continued)

- The results are very different when the human capital production function features *negative cross-partial derivatives*, $\partial^2 f(e_1, e_2)/\partial e_1 \partial e_2 < 0$ or exhibits *submodularity*.
- For example, we might have

$$\begin{aligned} h_1 &= \phi e_1 + e_2 - \lambda e_1^{1/2} e_2^{1/2} \\ h_2 &= e_1 + \phi e_2 - \lambda e_1^{1/2} e_2^{1/2} \end{aligned}$$

  where $\phi > 1$ and $\lambda > 0$ but small, so that human capital is increasing in parental background.
- In this case, background characteristics or resources are "substitutes".
- This can be thought as corresponding to the "good apple" theory of the classroom, where the kids with the best characteristics and attitudes bring the rest of the class up.

Segregation and Mixing (continued)

- In this case, because the cross-partial derivative is negative, the marginal willingness to pay of low-background parents to have their kid together with high-background parents is higher than that of high-background parents.
- With perfect markets, we will observe mixing, and in equilibrium schools will consist of a mixture of children from high- and low-background parents.
- Now combining the outcomes of these two models, many people jump to the conclusion that since we do observe segregation of schooling in practice, parental backgrounds must be complementary, so segregation is in fact efficient, and that Tiebout competition and parental sorting will increase efficiency.
- However, this conclusion is not correct; even if the correct production function does have the substitute property, segregation would arise in the presence of credit market problems or under reasonable limitations on prices.

Segregation and Mixing (continued)

- The way that mixing is supposed to occur with substitutes is that low-background parents make a payment to high-background parents so that the latter send their children to a mixed school.
- To see why such payments are necessary, recall that we always have that the first derivatives are positive, that is

$$\frac{\partial h_1}{\partial e_2} > 0 \text{ and } \frac{\partial h_2}{\partial e_1} > 0.$$

- This means that everything else being equal all children benefit from being in the same class with other children with good backgrounds. However, children from better backgrounds benefit *less* than children from less good backgrounds. This implies that there has to be payments from parents of less good backgrounds to high-background parents.

Segregation and Mixing (continued)

- Payments from poor backgrounds families to better off families to ensure mixing are both difficult to implement in practice, and practically impossible taking into account the credit market problems facing parents from poor socioeconomic status.
- Therefore, if the true production function exhibits the substitute property (submodularity), but there are credit market problems, we will observe segregation in equilibrium, and the segregation will be inefficient.
- This implies that we cannot simply appeal to Tiebout competition, or deduce efficiency from the equilibrium patterns of sorting.
- Another implication of this analysis is that in the absence of credit market problems (and with complete markets), cross-partials determine the allocation of students to schools.
- With credit market problems, first there of it has become important.
- This is a general result, with a range of implications for empirical work.

Workings of the City

- We have seen how inefficient segregation can occur in the presence of submodularity or substitute property.
- A very interesting paper by Benabou shows that even with supermodularity inefficient segregation can occur because of other "missing markets".
- His model has competitive labor markets, and local externalities (externalities in schooling in the local area).
- All agents are assumed to be ex ante homogeneous, and will ultimately end up either low skill or high skill.
- Utility of agent i is assumed to be

$$U^i = w^i - c^i - r^i$$

where w is the wage, c is the cost of education, which is necessary to become both low skill or high skill, and r is rent.

Workings of the City (continued)

- The cost of education is assumed to depend on the fraction of the agents in the neighborhood, denoted by x, who become high skill. In particular, we have $c_H(x)$ and $c_L(x)$ as the costs of becoming high skill and low skill.
- Both costs are decreasing in x, meaning that when there are more individuals acquiring high skill, becoming high skill is cheaper (positive peer group effects).
- In addition,

$$c_H(x) > c_L(x)$$

so that becoming high skill is always more expensive.

Workings of the City (continued)

- More importantly, the effect of increase in the fraction of high skill individuals in the neighborhood is bigger on the cost of becoming high skill, i.e., $c'_H(x) < c'_L(x)$. Or

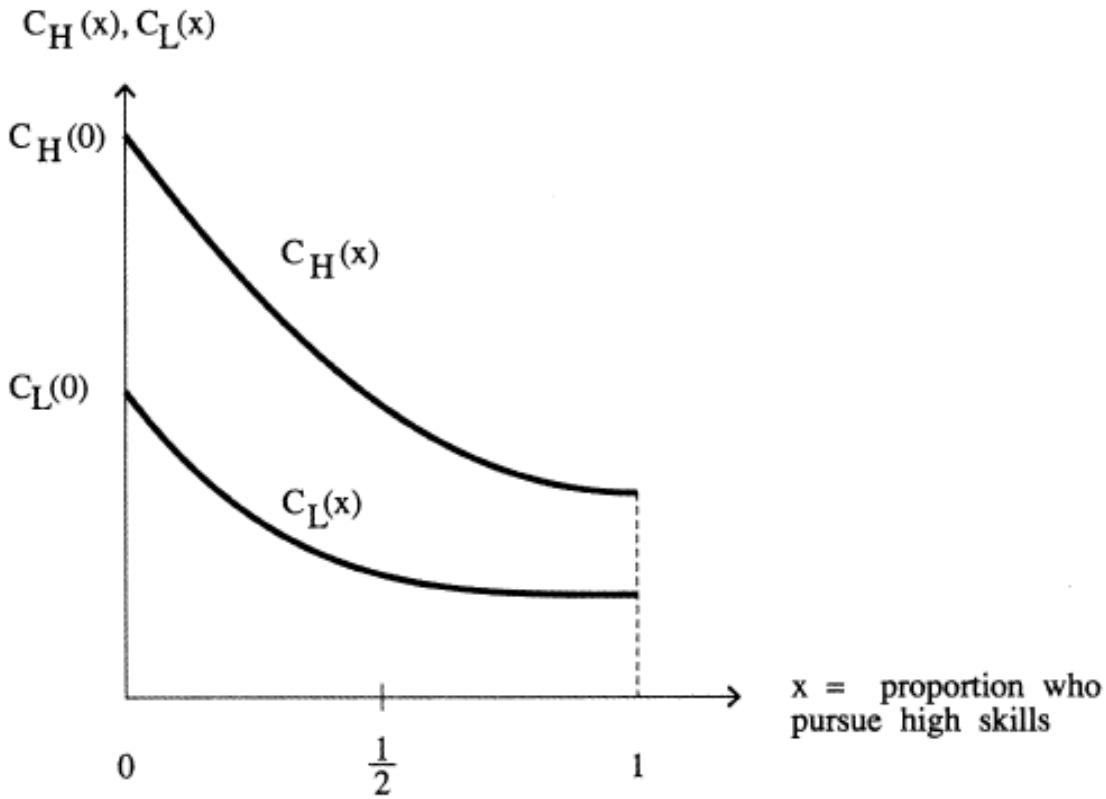


FIGURE I
The Costs of Skill Acquisition

Courtesy of Roland Benabou. Used with permission.

Workings of the City (continued)

- Since all agents are ex ante identical, in equilibrium we must have

$$U(L) = U(H)$$

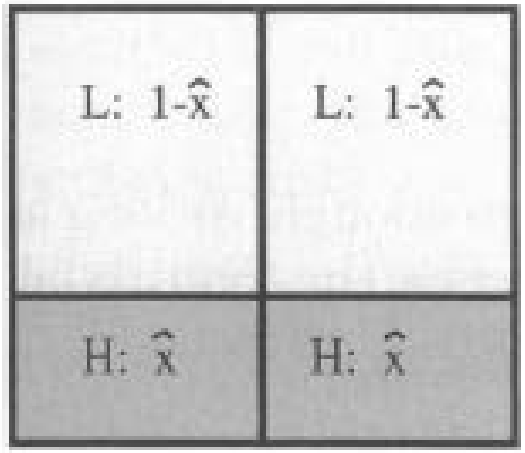
that is, the utility of becoming high skill and low skill must be the same.

- Assume that the labor market in the economy is global, and takes the constant returns to scale form $F(H, L)$.
- The important implication here is that irrespective of where the worker obtains his education, he will receive the same wage as a function of his skill level.
- Also assume that there are two neighborhoods of fixed size, and individuals will compete in the housing market to locate in one neighborhood or the other.

Workings of the City (continued)

- There can be two types of equilibria:

1. Integrated city equilibrium, where in both neighborhoods there is a fraction $\hat{x}$ of individual obtaining high education.

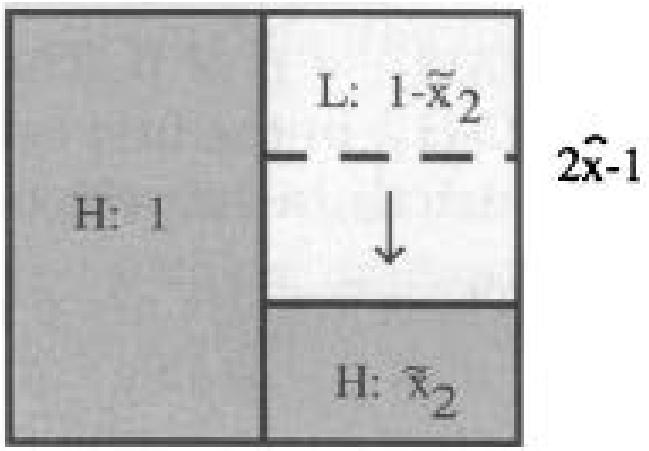


$$r_1 = r_2 = \hat{r}$$

Courtesy of Roland Benabou. Used with permission.

Workings of the City (continued)

2. Segregated city equilibrium, where one of the neighborhoods is homogeneous. For example, we could have a situation where one neighborhood has $x = 1$ and the other has $\tilde{x} < 1$, or one neighborhood has $x = 0$ and the other has $\bar{x} > 0$.



$$r_1 - r_2 = C_H(\tilde{x}_2) - C_H(1)$$

Courtesy of Roland Benabou. Used with permission.

Workings of the City (continued)

- The important observation here is that only segregated city equilibria are "stable".
- To see this consider an integrated city equilibrium, and imagine relocating a fraction ε of the high-skill individuals (that is individuals getting high skills) from neighborhood 1 to neighborhood 2.
- This will reduce the cost of education in neighborhood 2, both for high and low skill individuals.
- But by assumption, it reduces it more for high skill individuals, so all high skill individuals now will pay higher rents to be in that city, and they will outbid low-skill individuals, taking the economy toward the segregated city equilibrium.

Workings of the City (continued)

- In contrast, the segregated city equilibrium is always stable.
- Thus segregation arises as the equilibrium (stable equilibrium) outcome, because of "complementarities".
- As in the previous model with spillovers between students within the school, high-skill individuals can outbid the low-skill individuals because they benefit more from the peer group effects of high skill individuals.
- But crucially there are again *missing markets* in this economy.
- In particular, rather than paying high skill individuals for the positive externalities that they create, as would be the case in complete markets, agents transact simply through the housing market.
- In the housing market, there is only one rent level, which both high and low skill individuals pay.
- In contrast, with complete markets, housing prices would be such that high skill individuals pay a lower rent (to be compensated for the positive externality that they are creating on the other individuals).

Workings of the City (continued)

- This discussion implies that there are missing markets, and efficiency is not guaranteed.
- Is the allocation with segregation efficient?
- It turns out that it may or may not.
- To see this consider the problem of a utilitarian social planner maximizing total output minus costs of education for workers.
- This implies that the social planner will maximize

$$F(H, L) - H_1 c_H(x_1) - H_2 c_H(x_2) - L_1 c_L(x_1) - L_2 c_L(x_2)$$

where

$$x_1 = \frac{H_1}{L_1 + H_1} \text{ and } x_2 = \frac{H_2}{L_2 + H_2}$$

Workings of the City (continued)

- This problem can be broken into two parts:
  1. the planner will choose the aggregate amount of skilled individuals, and then she will choose how to actually allocate them between the two neighborhoods.
  2. then, there is simple cost minimization, and the solution depends on whether

$$\Phi(x) = xc_H(x) + (1-x)c_L(x)$$

  is concave or convex.
  This function is simply the cost of giving high skills to a fraction x of the population.
  When it is convex, it means that it is best to choose the same level of x in both neighborhoods, and when it is concave, the social planner minimizes costs by choosing two extreme values of x in the two neighborhoods.

Workings of the City (continued)

- It turns out that this function can be convex, i.e. $\Phi''(x) > 0$. More specifically, we have:

$$\Phi''(x) = 2\left(c_H'(x) - c_L'(x)\right) + x\left(c_H''(x) - c_L''(x)\right) + c_L''(x)$$

  We can have $\Phi''(x) > 0$ when the second and third terms are large. Intuitively, this can happen because although a high skill individual benefits more from being together with other high skill individuals, he is also creating a positive externality on low skill individuals when he mixes with them.
- This externality is not internalized, potentially leading to inefficiency.
- This model gives another example of why equilibrium segregation does not imply efficient segregation.

Evidence

- What does the evidence say?
- A lot of uncertainty (for reasons we will discuss in the next lecture).
- Positive externalities are probably present in several different settings.
  - School quality seems to matter, but in several instances, selective schools do not seem to have a positive effect on (marginally-admitted) students—either evidence that in this instance quality does not matter or strong supermodularity.
- Few papers look at the issue of supermodularity vs. submodularity, and the answer is like you to depend on the specific context also.
- Let us start with a discussion of the estimation of peer effects.

Evidence (continued)

- One thing is fairly clear: parents are willing to pay for their children to be together with high human capital peers (and with peers from good social economic background and in schools with various dimensions of higher quality).
- One nice illustration of this comes from Sandra Black's work focusing on parents' willingness to pay for housing is a function of school quality.
- She focuses on variation within school districts across hous es lying on different sides of attendance district boundaries, which determine which limit entry school child will go to.
- These households are subject to the same taxes and have the same access to other non-elementary school amenities (safety, public services etc.).
- But parents are willing to pay 2.5% more for houses for a 5% increase in test scores.

Evidence (continued)

Distance from boundary:	(1)	(2)	(3)	(4)	(5)
	All houses[d]	0.35 mile from boundary (616 yards)	0.20 mile from boundary (350 yards)	0.15 mile from boundary (260 yards)	0.15 mile from boundary (260 yards)
Elementary school test score[c]	.035	.016	.013	.015	.031
	(.004)	(.007)	(.0065)	(.007)	(.006)
Bedrooms	.033	.038	.037	.033	.035
	(.004)	(.005)	(.006)	(.007)	(.007)
Bathrooms	.147	.143	.135	.167	.193
	(.014)	(.018)	(.024)	(.027)	(.028)
Bathrooms squared	−.013	−.017	−.015	−.024	−.025
	(.003)	(.004)	(.005)	(.006)	(.007)
Lot size (1000s)	.003	.005	.005	.005	.004
	(.0003)	(.0005)	(.0005)	(.0007)	(.0006)
Internal square footage (1000s)	.207	.193	.191	.195	.191
	(.007)	(.01)	(.01)	(.02)	(.012)
Age of building	−.002	−.002	−.003	−.003	−.002
	(.0003)	(.0002)	(.0005)	(.0006)	(.0004)
Age squared	.000003	.000003	.00001	.000009	.000005
	(.000001)	(.0000006)	(.000002)	(.000003)	(.000002)
Boundary fixed effects	NO	YES	YES	YES	NO
Census variables	Yes	No	No	No	Yes

Evidence: Magnitudes

	(1) Basic hedonic regression[a]	(2) 0.35 sample boundary fixed effects	(3) 0.20 sample boundary fixed effects	(4) 0.15 sample boundary fixed effects
Coefficient on elementary school test score[b]	.035 (.004)	.016 (.007)	.013 (.0065)	.015 (.007)
Magnitude of effect (percent change in house price as a result of a 5% change in test scores)[c]	4.9%	2.3%	1.8%	2.1%
$ Value (at mean tax-adjusted house price of $188,000 in $1993)	$9212	$4324	$3384	$3948
$ Value (at median tax-adjusted house price of $158,000 in $1993)	$7742	$3634	$2844	$3318

Evidence: Interpretation

- Very clean result.
- But should she have stopped here?
- What other implications should one have checked?

Econometric Framework for Peer Effects

- The simplest econometric model for estimating peer effects would be

$$y_{ij} = \beta_{\text{own}} x_{ij} + \beta_{\text{spillover}} \bar{X}_j + \varepsilon_{ij} \tag{15}$$

  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 $\bar{X}$.

An Application to Job Placement Assistance

- In some contexts, contextual effects are natural. In such contexts, if we can also find (or design) random assignment, then credible estimates can be obtained.
- 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^{+++}	0.042^{+++}	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	0.265	0.372	0.26	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^{+++}	0.085^{+++}	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)	(0.026)
Net effect	0.040^{+++}	0.050^{+++}	0.029	0.053	0.088	0.035
of program participation ($\beta+\delta$)	(0.014)	(0.022)	(0.019)	(0.033)	(0.062)	(0.042)
Control Mean	0.16	0.121	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	0.272	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?

Econometric Issues

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

$$y_{ij} = \beta_{\text{own}} x_{ij} + \alpha_{\text{spillover}} \bar{Y}_j + \varepsilon_{ij} \tag{16}$$

  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).

Example: Ethnic Effects

- The correlation between the outcomes of people from the same ethnic group is well known, and is often interpreted in various different ways.
- Borjas (1994, 1995) suggested that these are related to the effect of "ethnic capital" — meaning that if in a group has low human capital, then the next generation will be at a disadvantage in human capital acquisition.
- What could be wrong with 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"*.

An Application

- As a concrete example of the estimation of an important set of peer effects, and also illustrating these problems, consider the paper by Sacerdote (2001), which uses random assignment of roommates in Dartmouth.
- He finds that the GPAs of randomly assigned roommates are correlated, and interprets this as evidence for peer group effects.

An Application (continued)

TABLE III

PEER EFFECTS IN ACADEMIC OUTCOMES

	(1) Fresh year GPA	(2) Fresh year GPA w/ dorm f.e.	(3) Senior year GPA	(4) Fresh year GPA	(5) Fresh year GPA	ye
Roommates' GPA	0.120**	0.068**	0.008			
	(0.039)	(0.029)	(0.026)			
HS academic score (self)	0.014**	0.015**	0.013**			
	(0.0008)	(0.0007)	(0.0009)			
HS academic score (roommates')	−0.001	−0.0003	0.0009			
	(0.001)	(0.0009)	(0.001)			
roommates' academic score bottom 25 percent				0.016	0.014	
				(0.028)	(0.025)	
roommates' academic score top 25 percent				0.060**	0.047*	
				(0.028)	(0.026)	
roommates' intention to graduate w/honors (1–4)						
own academic score bottom 25 percent					−0.284**	−
					(0.025)	

An Application (continued)

- Despite the very nice nature of the experiment, the conclusion is problematic, because Sacerdote attempts to identify (16) rather than (15).
- For example, to the extent that there are common shocks to both roommates;
  - e.g., they are in a noisier dorm), this may not reflect peer group effects.
- This identification problem would not have arisen if the right-hand side regressor was some predetermined characteristic of the roommate
  - in this case, we would be estimating something similar to (15) rather than (16).

An Application (continued)

- A paper using random assignment of cadets to companies (approximately consisting of 38 cadets) at West Point, David Lyle (2007) can look specifically into this issue.
- He finds that results similar to Sacerdote's are more likely to be due to common shocks than pure peer effects.

An Application (continued)

TABLE 3.—CONTEMPORANEOUS PEER EFFECTS WITH POTENTIAL COMMON SHOCKS
OUTCOME VARIABLE: INDIVIDUAL-LEVEL PLEBE ACADEMIC GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own total SAT/100	0.042	0.042	0.042	0.042	0.042	0.037	0.037
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)
Average peer total SAT/100		−0.002		−0.024	−0.018	−0.013	−0.011
		(0.035)		(0.027)	(0.030)	(0.033)	(0.038)
Average peer academic GPA			0.234	0.241	0.206	0.256	0.140
			(0.056)	(0.057)	(0.061)	(0.076)	(0.092)
CEER/100	0.398	0.398	0.399	0.398	0.397	0.352	0.352
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.030)	(0.030)
WCS/1,000	0.203	0.203	0.206	0.206	0.205	0.283	0.281
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.039)	(0.039)
Female	−0.071	−0.071	−0.071	−0.071	−0.071	−0.064	−0.065
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.023)	(0.023)
Black	−0.141	−0.141	−0.142	−0.141	−0.141	−0.115	−0.114
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.025)	(0.026)
Hispanic	−0.046	−0.046	−0.047	−0.047	−0.047	−0.048	−0.048
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.033)	(0.032)
Recruited football player	−0.031	−0.031	−0.032	−0.032	−0.033	−0.014	−0.013
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.030)	(0.030)
Other recruited athlete	−0.009	−0.009	−0.010	−0.010	−0.010	0.013	0.015
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.020)	(0.020)
Attended the West Point Prep School	−0.036	−0.036	−0.038	−0.038	−0.037	−0.037	−0.035
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.019)	(0.019)
R^2	0.43	0.43	0.43	0.43	0.43	0.40	0.41
Observations	7,527	7,527	7,527	7,527	7,527	4,048	4,048
Battalion and regiment controls	No	No	No	No	Yes	No	No
Average upperclassmen controls (shocks)	No	No	No	No	No	No	Yes

David S Lyle, 'Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point', The Review of Economics and Statistics, 89:2 (May, 2007), pp. 289-299. © 2007 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology, published by the MIT Press.
https://www.mitpressjournals.org/doi/abs/10.1162/rest.89.2.289/

An Example of Submodularity

- Carrell, Sacerdote and West (2011) and an earlier paper by Carrell, Fullerton and West (2009) exploits the random assignment of cadets to different squadrons in the U.S. Air Force Academy, they have convinced U.S. Air Force to change the composition of squadrons.
- The results in the earlier study, using random assignment resulting from the existing policy of the U.S. Air Force, show that "low ability" cadets (students) benefit most from high ability peers in their squadron.
- We will discuss this paper more in the next lecture, but for now it provides one piece of evidence of the substitute/submodularity effects.
- In particular, a large positive effects are on those at the bottom of the predicted GPA distribution (in terms of their pre-treatment covariates).

An Example of Submodularity (continued)

TABLE II

PEER EFFECTS IN THE PRE-TREATMENT GROUP[a]

Variables	(1) GPA	(2) GPA	(3) GPA	(4) GPA
Fraction of High SAT-V Peers	0.181[d]	0.190[c]		
	(0.094)	(0.096)		
Fraction of Low SAT-V Peers	−0.050	−0.061		
	(0.095)	(0.094)		
Fraction of High SAT-V Peers × High $\widehat{GPA}$			0.222	0.233
			(0.156)	(0.151)
Fraction of High SAT-V Peers × Middle $\widehat{GPA}$			−0.136	−0.119
			(0.136)	(0.137)
Fraction of High SAT-V Peers × Low $\widehat{GPA}$			0.464[b]	0.474[b]
			(0.150)	(0.152)
Fraction of Low SAT-V Peers × High $\widehat{GPA}$			0.026	0.009
			(0.144)	(0.147)
Fraction of Low SAT-V Peers × Middle $\widehat{GPA}$			−0.219	−0.230
			(0.145)	(0.142)
Fraction of Low SAT-V Peers × Low $\widehat{GPA}$			0.065	0.061
			(0.141)	(0.140)
Observations	14,024	14,024	14,024	14,024
R^2	0.345	0.345	0.346	0.345
F All Peer Variables	0.994		1.089	
p-value All	0.430		0.365	
F All Peer SAT Verbal Variables	2.304	2.627	2.412	2.464
p-value SAT Verbal	0.102	0.075	0.028	0.025
F High SAT Verbal Peers (High $\widehat{GPA}$ v Middle $\widehat{GPA}$)			3.068	3.135
p-value H v M			0.081	0.078
F High SAT Verbal Peers (High $\widehat{GPA}$ v Low $\widehat{GPA}$)			1.598	1.665
p-value H v L			0.208	0.198
F High SAT Verbal Peers (Low $\widehat{GPA}$ v Middle $\widehat{GPA}$)			9.850	9.441
p-value L v M			0.002	0.002

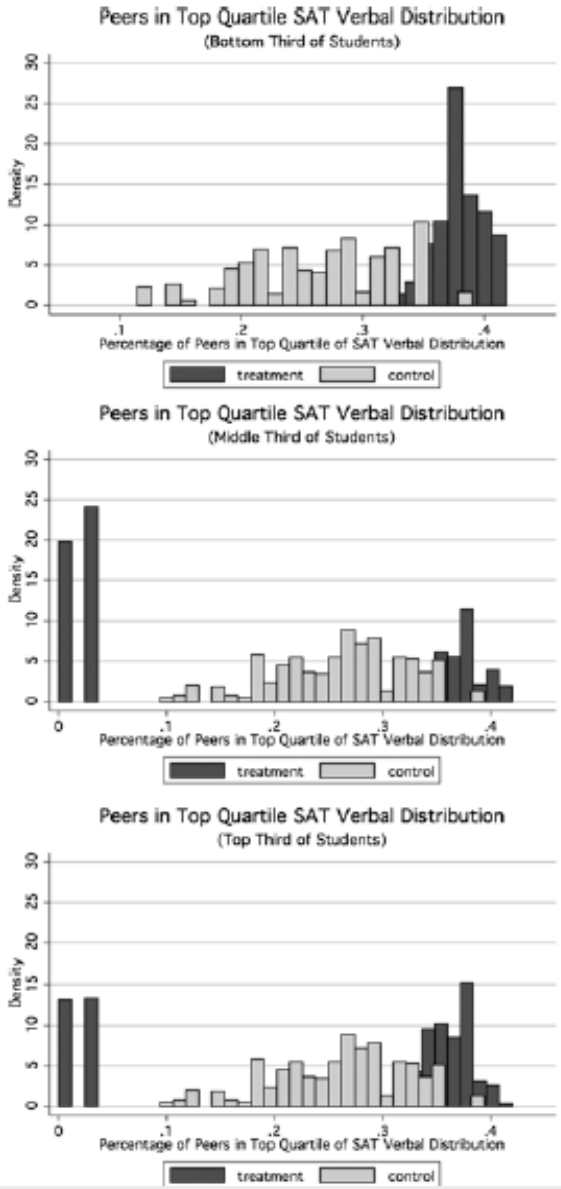
Endogenous Networks

- A large literature studies the endogenous formation of (social) networks—e.g., Jackson and Wolinsky (*JET*, 1996), Bala and Goyal (*Econometrica,* 2000).
- Endogeneity of networks makes externalities and peer effects more interesting but also more complicated conceptually and more difficult to estimate.
- Let us return to Carrell, Sacerdote, and West (*Econometrica*, 2013).
- Recall that the peer effects they are estimating from the cadets within squadrons using random assignment from the U.S. Air Force Academy are non-linear.
    - Low (baseline) ability students appear to benefit significantly from being in the same squadron has high-ability students with limitednegative effect on high-ability students from such mixin.
- This suggests that optimally manipulating the composition of squadrons can lead to significant gains.

Endogenous Networks: Intervention Gone Wrong

- The authors convinced the U.S. Air Force Academy to allow such manipulation, and constructed "optimally designed" squadrons—in which the exposure of low-ability cadets to high-ability ones was maximized by creating *"bimodal"* squadrons.
- However, instead of the hypothesized gains, there were losses among low-ability cadets. Why?
- The authors hypothesize, and provides some evidence in favor of, the following story:
  - The real peer groups—the friendship networks—probably changed as a result of the intervention: low-ability and high-ability cadets may have stopped working and being friends together in the bimodal squadrons.
  - As a result, the peer effects from high-ability to low-ability cadets weakened or disappeared, leading to negative results.
- A cautionary tale on the endogeneity of social networks with respect to interventions.

Endogenous Networks: Bimodal Treatment



Endogenous Networks: Prediction Vs. Realization

TABLE IV

PREDICTED TREATMENT EFFECT[a]

Variables	(1) All Students	(2) Bottom $\widehat{GPA}$	(3) Middle $\widehat{GPA}$	(4) Top $\widehat{GPA}$
Student in Treatment Group	2.787	2.390	2.783	3.198
	(0.026)	(0.027)	(0.027)	(0.027)
Student in Control Group	2.772	2.336	2.767	3.195
	(0.026)	(0.027)	(0.027)	(0.026)
Predicted Treatment Effect	0.015	0.053[b]	0.016	0.003
	(0.037)	(0.037)	(0.037)	(0.037)
Observations	2653	881	884	888

TABLE VI

OBSERVED TREATMENT EFFECTS[a]

Variables	(1) All Students	(2) Low $\widehat{GPA}$	(3) Middle $\widehat{GPA}$	(4) High $\widehat{GPA}$
Student in Treatment Group	0.001	−0.061[c]	0.082[b]	−0.012
	(0.022)	(0.031)	(0.039)	(0.036)
Observations	4834	1571	1626	1637
R^2	0.357	0.136	0.067	0.151

Endogenous Networks: A Possible Explanation?

TABLE VII

LOW PREDICTED GPA STUDENTS: TREATMENT EFFECTS ON STUDY PARTNER AND FRIEND CHOICES[a]

	(1)	(2)	(3)	(4)	(5)	(6)
	Study Partners			Friends		
Treatment Effect On ...	Actual Peer Choices (sd)	If Peers Chosen Randomly (sd)	Actual Minus Random $P(A \prec R)$	Actual Peer Choices (sd)	If Peers Chosen Randomly (sd)	Actual Minus Random $P(A \prec R)$
Fraction Low $\widehat{GPA}$	0.171[b]	0.125[b]	0.046[c]	0.201[b]	0.135[b]	0.071[b]
	(0.061)	(0.020)	0.011	(0.060)	(0.019)	0.000
Fraction Middle $\widehat{GPA}$	−0.214[b]	−0.105[b]	−0.109[b]	−0.105[b]	−0.114[b]	0.008
	(0.046)	(0.015)	1.000	(0.038)	(0.014)	0.272
Fraction High $\widehat{GPA}$	0.042	−0.020	0.062[b]	−0.095[c]	−0.022	−0.074[b]
	(0.060)	(0.016)	0.000	(0.048)	(0.015)	1.000
Fraction High SAT-V	0.064	0.103[b]	−0.039[c]	0.004	0.112[b]	−0.107[b]
	(0.052)	(0.015)	0.984	(0.046)	(0.014)	1.000
Fraction Low $\widehat{GPA} > 0.50$	0.269[b]	0.170[b]	0.100[c]	0.236[c]	0.170[b]	0.066
	(0.097)	(0.042)	0.008	(0.092)	(0.041)	0.057
Observations	494	10,000	10,000	543	10,000	10,000

14.661 Labor Economics I

Fall 2017