Notes on Education Economics

From Dr. Chinhui Juhn's Lectures, Compiled by Brian Murphy*†

Contents

*These notes are from my time as a student in the University of Houston PhD Economics program.
†Typos may exist in these notes. If any are found, please contact me.

1 Introduction

These notes present key concepts from a Ph.D. course in the Economics of Education. Topics include, but are not limited to, human capital theory, estimating the returns to schooling, the GED, intergenerational mobility, early childhood investment, school spending, class size, peer effects, teachers, and other topics.

Note, some of the class days had students present their own research, or papers on the subject. These notes do not include those.

2 Human Capital Theory

2.1 Mincerian Earnings Equation

References: Willis' HB chapter, "Wage Determinants..."; Rosen, S., "Human Capital: Survey..."; Becker Woytinsky Lecture.

The Mincerian earnings function relates earnings to schooling and experience. In this formulation, s denotes years of schooling and x denotes years of labor market experience. The equation is expressed as:

$$\ln Y = \beta_0 + \beta_1 s + \beta_2 x + \beta_3 x^2 + \varepsilon \tag{1}$$

Empirical evidence consistently demonstrates a robust positive relationship between schooling and earnings across various countries, implying that $\beta_1 > 0$. Moreover, the positive coefficient on x and the negative coefficient on x^2 suggest a concave earnings profile over the life cycle. Human capital theory attributes these relationships to investment behavior in human capital.

There are several important distinctions between investment in human capital and investment in physical capital:

- Human capital is embodied in individuals; human beings cannot be bought or sold.
- The distinction between general and specific human capital is crucial.
- Borrowing constraints are significant because loans cannot be collateralized using human capital.

2.2 Simple Investment Problem

Consider an individual who chooses the optimal number of years of schooling, s, to maximize the present value (PV) of lifetime earnings, denoted by L. For simplicity, assume that:

1. All individuals are identical.
2. Individuals live indefinitely.
3. The only cost of schooling is foregone earnings.
4. The earnings function, $y = f(s)$, is taken as given.

The problem is formulated as:

$$\max L = \int_s^\infty f(s) e^{-rt}\, dt \tag{2}$$

Evaluating the integral:

$$\begin{aligned} L &= f(s)\left(\frac{-1}{r}\right) e^{-rt}\Big|_{t=s}^{t=\infty} \\ &= f(s)\left(\frac{-1}{r}\right)\left[0 - e^{-rs}\right] \\ &= \frac{f(s)}{r} e^{-rs} \end{aligned} \tag{3}$$

Differentiating L with respect to s and setting the derivative equal to zero yields the optimality condition. Applying the product rule,

$$\frac{d}{dx}[f(x) \cdot g(x)] = f(x)g'(x) + g(x)f'(x), \tag{4}$$

we have:

$$\frac{dL}{ds} = \frac{d}{ds}\left(\frac{f(s)}{r}e^{-rs}\right) = 0, \tag{5}$$

$$\frac{f(s)}{r}(-r)e^{-rs} + e^{-rs}\frac{f'(s)}{r} = 0, \tag{6}$$

$$\frac{1}{r}e^{-rs}\left[f'(s) - rf(s)\right] = 0, \tag{7}$$

$$f'(s) = rf(s), \tag{8}$$

$$\frac{f'(s)}{f(s)} = r. \tag{9}$$

This condition is analogous to the tree-cutting problem, which asks: *When should one harvest the tree?*

Insight #1: Investment continues until the internal rate of return equals the market interest rate.

2.3 Optimal Investment Over the Life Cycle

Suppose individuals allocate their time between working and investing in human capital. Let:

- $k(t)$ denote the fraction of time spent investing in human capital at time t,
- $1 - k(t)$ denote the fraction of time devoted to working,
- $H(0)$ denote the initial stock of human capital at birth,
- ρ denote the rate of return on each unit of time spent learning.

The instantaneous growth rate of human capital at time t is given by:

$$h(t) = \rho\, k(t).$$

Thus, the stock of human capital at time t is:

$$H(t) = H(0)e^{\int_0^t h(\tau)\, d\tau}. \tag{10}$$

Assuming the rental price of human capital is normalized to 1, earnings at time t are:

$$Y(t) = \left[1 - k(t)\right]H(t). \tag{11}$$

It is assumed that while in school, individuals invest full time in acquiring human capital (i.e., $k(t) = 1$ for $t < s$). Consequently, upon leaving school at time s, the stock of human capital is:

$$H(s) = H(0)e^{\rho s}. \tag{12}$$

If no further investment occurs after school, earnings are proportional to the human capital stock:

$$\ln Y = \ln H(0) + \rho s. \tag{13}$$

However, empirical evidence suggests that individuals continue to invest in human capital through on-the-job training, although the rate of investment declines over time as the returns diminish near the end of life. Suppose that during the working period, the fraction of time invested, $k(t)$, declines linearly from an initial value $k(0)$ to 0 at the terminal period T. Let $x = t - s$ denote years in the labor market (i.e., potential experience). Then,

$$k(x) = k(0) - \left(\frac{k(0)}{T-s}\right) x. \tag{14}$$

The human capital stock for an individual with x years of experience becomes:

$$\begin{aligned} H(x) &= H(s) e^{\rho \int_0^x \left(k(0) - \frac{k(0)}{T-s} t\right) dt} \\ &= H(s) e^{\rho \left[k(0)t - \frac{1}{2}\left(\frac{k(0)}{T-s}\right) t^2\right]_0^x} \\ &= H(s) e^{\rho k(0) x - \frac{\rho}{2}\left(\frac{k(0)}{T-s}\right) x^2}. \end{aligned} \tag{15}$$

Earnings, net of the cost of investment, are given by:

$$Y(x) = [1 - k(x)] H(x). \tag{16}$$

Taking logarithms yields:

$$\ln Y = \ln H(0) + \rho s + \rho k(0) x - \left(\frac{\rho k(0)}{2(T-s)}\right) x^2 + \ln(1 - k(x)). \tag{17}$$

The Mincerian earnings function can be interpreted as an approximation of the expression above.

Insight #2: Individuals tend to invest more heavily in human capital early in life because:

a) The stock of human capital is initially low, resulting in lower foregone earnings.

b) There are more periods over which the returns on early investments can be enjoyed. Consequently, older individuals typically invest less, and similarly, individuals with shorter working careers (e.g., women) tend to invest less in human capital (although this pattern may not hold for formal schooling).

Overall, the optimal investment model predicts a concave earnings profile. Early in life, earnings are low due to heavy investment and a small human capital stock; during mid-life, earnings peak as the human capital stock is high and the rate of investment declines; eventually, investment ceases and depreciation of the human capital stock may lead to a decline in earnings.

Discussion: In this model, wage growth with potential experience reflects the effect of ongoing investment in human capital. Alternative models that explain wage growth with potential experience may incorporate different mechanisms.

Some limitations of the model include:

a) The model focuses on the accumulation of skill quantity via $f(s)$, without fully exploring the determinants of the rental price of human capital, $R(t)$.

b) Under the assumption regarding $R(t)$, workers with s years of schooling are considered perfect substitutes for workers with $s+1$ years, implying that earnings differences arise solely from differences in the quantity of skills, not from differences in the prices of skills.

Importantly, when all individuals are identical, the model does not generate variation in schooling or income across people, which contradicts cross-sectional data that show heterogeneity in both schooling and earnings.

2.4 Schooling and the Distribution of Income

References: Willis; HB; Becker Woytinsky Lecture (1967); Rosen (1977).

The standard Mincer specification assumes that every individual experiences the same return to schooling. If all individuals are identical and face the same interest rate r, then they would all choose the same optimal schooling level, s^*, and thus the same lifetime earnings, L^* (present value of lifetime earnings). Consequently, this formulation does not generate a distribution of schooling, s, or lifetime income, L^*.

2.5 Becker Woytinsky Lecture

The distribution of schooling and earnings emerges when differences in returns are introduced, which may depend on individual ability as well as variations in borrowing costs (r).

Case i: Differences in Ability
Individuals differ in ability while facing the same interest rate r. For a given level of ability, it is assumed that the rate of return to schooling declines with additional schooling. Thus, the maximized present value of lifetime earnings, L^*, will differ across individuals such that:

$$L_3^* > L_2^* > L_1^*. \tag{18}$$

Case ii: Differences in Borrowing Costs
Even when individuals have the same ability, variations in borrowing costs can lead to unequal access to funds. Wealthier families typically have better access to capital, implying that individuals face different effective interest rates. Under the assumption of imperfect capital markets for human capital—where individuals cannot freely borrow against their future earnings—this variation in r generates differences in schooling decisions. In this context, the returns to schooling function, as formulated by Rosen, is traced out based on the relationship between earnings and schooling.

Rosen's formulation is particularly useful because earnings are observable. In this framework, earnings depend on both schooling and ability:

$$\ln Y = f(s, A).$$

2.6 Rosen's Formulation

Differences in Ability:
Consider a case where individual 2 possesses higher ability and, consequently, a larger return to schooling. As a result, individual 2 will acquire more schooling. In cross-sectional data, estimating the slope of the earnings function using observations from individuals with different abilities (e.g., points (a) and (b)) tends to overstate the returns for both individuals, a phenomenon known as *ability bias*. Notably, if ability only affects the intercept, it would be absorbed in β_0, but here ability also affects the slope, generating differences in s.

In cross-sectional earnings regressions of the form:

$$\ln Y = \beta_0 + \beta_1 s + \beta_2 x + \beta_3 x^2 + \varepsilon, \tag{19}$$

the estimator $\hat{\beta}_1$ is upward-biased because unobserved ability, present in ε, is correlated with s.

Differences in Borrowing Costs:
When individuals differ in borrowing costs, those facing the highest r will acquire the least amount of schooling. In this situation, the true schooling–earnings function, $f(s)$, can be identified.

3 Empirical Estimates of the Returns to Schooling

3.1 Becker's Woytinsky Lecture — Card (1999)

Background Reading: Card (HB 1999) and Angrist and Kruger (1991).

Suppose that the marginal returns and marginal costs have the following functional forms, as laid out in Becker (1967):

$$(0.1)\frac{Y'(s)}{Y(s)} = b_i - k_1 s, \tag{20}$$

$$h'(s) = r_i + k_2 s. \tag{21}$$

Here, b_i and r_i represent individual-specific ability and borrowing cost terms, respectively. Note that the marginal cost of funds is an increasing function of s. Additionally, the cost of schooling may be motivated more generally by considering factors such as a distaste for schooling, the discount rate, and borrowing costs.

The optimal schooling choice is defined by:

$$(0.1) \quad s_i^* = \frac{(b_i - r_i)}{k}, \quad \text{where } k_1 + k_2 = k. \tag{22}$$

From (1.1), the log earnings equation is given by:

$$\log Y_i = \alpha_i + bs_i - \frac{1}{2}k_1 s_i^2. \tag{23}$$

Rewriting this equation in deviations from the means yields:

$$\log Y_i = a_0 + \bar{b}s_i - \frac{1}{2}k_1 s_i^2 + a_i + (b_i - \bar{b})s_i, \tag{24}$$

where $a_i = \alpha_i - a_0$ and has mean zero.

Consider the linear projection of a_i and $(b_i - \bar{b})$ on schooling:

$$a_i = \lambda_0(s_i - \bar{s}) + u_i, \tag{25}$$

$$b_i - \bar{b} = \psi_0(s_i - \bar{s}) + v_i, \tag{26}$$

with $E[s_i u_i] = E[s_i v_i] = 0$.

The theoretical regression coefficients λ_0 and ψ_0 are defined as follows:

$$\lambda_0 = \frac{\text{cov}(a_i, s_i)}{\text{var}(s_i)} = k\,\frac{\sigma_{ba} - \sigma_{ra}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}}, \tag{27}$$

$$\psi_0 = \frac{\text{cov}(b_i - \bar{b}, s_i)}{\text{var}(s_i)} = \frac{\text{cov}\left(b_i - \bar{b},\ \frac{(b_i - r_i)}{k}\right)}{\text{var}(s_i)} = k\,\frac{\sigma_b^2 - \sigma_{br}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}}, \tag{28}$$

where σ_{br} denotes the covariance of marginal returns and marginal costs across individuals—that is, whether individuals with higher returns tend to face higher or lower costs.

An OLS regression of $\log Y_i$ on s_i results in:

$$\beta_{OLS} = \frac{\text{cov}(\log Y_i, s_i)}{\text{var}(s_i)} = \frac{\text{cov}(a_0 + \bar{b}s_i - \ldots,\ s_i)}{\text{var}(s_i)}, \tag{29}$$

(see Card (1999), Appendix A).

Summarizing, the OLS estimator can be written as:

$$\beta_{OLS} = \bar{b} + \lambda_0 + \psi_0\,\bar{s}, \tag{30}$$

$$\lambda_0 = \frac{\text{cov}(a_i, s_i)}{\text{var}(s_i)} = k\,\frac{\sigma_{ba} - \sigma_{ra}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}} \quad \text{(bias due to individual-specific intercepts } a_i\text{)}, \tag{31}$$

$$\psi_0\,\bar{s} = \frac{\text{cov}(b_i, s_i)}{\text{var}(s_i)} = k\,\frac{\sigma_b^2 - \sigma_{br}}{\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}}\,\bar{s} \quad \text{(bias due to individual-specific slopes } b_i\text{)}. \tag{32}$$

The term σ_{br} represents the underlying covariance of marginal returns and marginal costs across individuals. Typically, if this covariance is negative (i.e., if individuals with higher returns tend to have lower borrowing costs), then the bias is positive—implying that the endogeneity of schooling likely leads to an upward bias in the OLS estimates. More generally, the magnitude of the bias depends on both the covariance and the relative importance of the variance in schooling attributable to returns (b_i) versus the cost of funds (r_i). For simplicity, if we assume that $\sigma_{br} = 0$, then the bias increases with the variance in schooling due to differences in marginal returns across individuals.

3.2 Estimation Strategies Used to Estimate Returns to Schooling

Several strategies have been employed to estimate the returns to schooling:

1. Instrumental Variables (IV) using compulsory schooling law changes and quarter of birth (Angrist and Krueger, 1991).
2. IV using family background interacted with proximity to a nearby college.
3. OLS estimation controlling for family background.
4. Within-family comparisons, such as twin studies.

3.2.1 IV Strategies Using Compulsory Schooling Laws and Quarter of Birth

An IV strategy is one approach to address the endogeneity of schooling and the associated ability bias. The structural equation of interest is:

$$y_i = \alpha + \rho S_i + \eta_i, \tag{33}$$

where S_i (schooling) is endogenous because it may be correlated with the error term η_i. An instrument Z is introduced via the first-stage equation:

$$S_i = Z_i \, \delta + v_i. \tag{34}$$

A valid instrument must satisfy the following conditions:

a) Z is correlated with schooling.

b) $\text{cov}(Z, \eta) = 0$ (the exclusion restriction).

Angrist and Krueger (1991) Angrist and Krueger use quarter of birth dummies, interacted with cohort dummies, as instruments for education. The reasoning behind this instrument is:

- In most states, children enter school in the calendar year in which they turn six.
- Children born in December typically enter school at an age slightly below six, whereas those born in January enter school at about 6¾ years.
- For cohorts subject to binding compulsory schooling laws, individuals born earlier in the year (and therefore eligible to drop out at the legal minimum age of 16) tend to acquire less education than those born later in the year.

Empirical evidence reveals pronounced patterns in education and earnings by quarter of birth (see Figures I and II). The simplest IV estimator is the Wald estimator, defined as:

$$\rho_{IV} = \frac{E[y_i \mid Z_i = 1] - E[y_i \mid Z_i = 0]}{E[s_i \mid Z_i = 1] - E[s_i \mid Z_i = 0]}. \tag{35}$$

Reported estimates indicate that $\beta_{IV} = 0.102$ and $\beta_{OLS} = 0.071$. Notably, the IV estimate is larger than the OLS estimate, which is surprising given that the IV approach is intended to correct for the upward bias from unobserved ability (see Table III).

Several issues with the study include:

a) **Weak Instruments Problem:** Bound, Jaeger, and Baker (1995) show that when many weak instruments are used (e.g., fourth quarter of birth dummies crossed with 50 states and 9 birth cohorts), the IV estimates can be driven close to the OLS estimates (see Bound, Jaeger, and Baker, Table 3).

b) **Exclusion Restriction Concerns:** Buckles and Hungerman (2008) note that systematic patterns in outcomes by quarter of birth exist. For example, children born in winter are more likely to have mothers who are teenagers, unmarried, and high school dropouts, indicating worse family backgrounds.

This raises a central puzzle: Why do the IV estimates, which are intended to control for ability bias, yield returns to schooling that are as large as or larger than those estimated by OLS?

Local Average Treatment Effect (LATE) Interpretation An alternative explanation is provided by the Local Average Treatment Effect (LATE) interpretation. Consider that returns to schooling may be heterogeneous across individuals, where $\beta_i = \frac{\partial Y_i}{\partial S_i}$ and $E[\beta_i] = \bar{\beta}$. Following Angrist and Krueger (1991), using quarter of birth as an instrument leads to the following potential outcomes framework:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})S_i, \quad \text{with } (Y_{1i} - Y_{0i}) \text{ representing the causal effect of schooling,} \tag{36}$$

$$S_i = S_{0i} + (S_{1i} - S_{0i})Z_i, \quad \text{with } (S_{1i} - S_{0i}) \text{ representing the causal effect of the instrument on schooling.} \tag{37}$$

Here,

- S_{1i} denotes the level of schooling an individual would attain if $Z_i = 1$ (e.g., born in a later quarter),
- S_{0i} denotes the level of schooling if $Z_i = 0$ (e.g., born in the first quarter).

Two key assumptions are made:

Assumption 1: Independence The set of potential outcomes $\{Y_{1i}, Y_{0i}, S_{1i}, S_{0i}\}$ is independent of Z_i. Since quarter of birth is randomly assigned, the instrument is uncorrelated with unobserved determinants of both schooling and earnings.

Assumption 2: Monotonicity The effect of the instrument on schooling is nonnegative, meaning that $S_{1i} - S_{0i} \geq 0$ for all individuals; the instrument either has no effect or increases schooling.

Under these assumptions, the Wald IV estimator is equivalent to the ratio of the difference in mean earnings between individuals with $Z_i = 1$ and $Z_i = 0$ to the difference in their mean schooling. Formally, we have:

$$\begin{aligned} \rho_{IV} &= \frac{E[Y_i(S_{1i}) - Y_i(S_{0i})]}{E[S_{1i} - S_{0i}]} \\ &= E\Big[\omega_i \, Y_i'(\hat{S}_i)\Big], \end{aligned} \tag{38}$$

where ω_i is a weight proportional to the effect of the instrument on S_i, and $\hat{S}_i \in [S_{0i}, S_{1i}]$. In essence, ρ_{IV} is a local average derivative over the range of schooling variation induced by the instrument.

In practice, for most observations, compulsory schooling laws and quarter of birth have little impact on schooling decisions. Angrist and Krueger (1991) find that the largest schooling differences by quarter of birth occur among individuals with 8–12 years of education. Consequently, the IV estimator identifies the return to schooling predominantly for individuals at the lower end of the education distribution. This highlights a trade-off in IV estimation: while the IV approach may yield a clear causal interpretation, its generalizability is limited.

4 Estimating Returns to Schooling Using School Reforms

4.1 Meghir and Palme (2005)

Meghir and Palme (2005) and Duflo (2001) provide empirical analyses of the effects of school reforms on educational attainment and earnings.

Meghir and Palme (2005) examine the 1950 school reform in Sweden, which implemented the following changes:

1. Extension of compulsory schooling to nine years.
2. Abolition of grade-based tracking, allowing students to choose among three tracks: academic, general, and general with vocational training.
3. Standardization of curricula across all schools at the national level.

The reform was legislated in 1948 and subsequently rolled out gradually across municipalities.

Two surveys were conducted in 1961 and 1966, targeting sixth-grade students from the 1948 and 1953 birth cohorts. These surveys collected data on parental education, IQ scores, grades, and school type. The survey data were subsequently merged with administrative records on completed years of schooling and earnings.

Given the staggered roll-out of the reform, individuals within the same cohort were exposed to either the old or the new system depending on their municipality of residence. Additionally, within a given municipality, the older cohort followed the old system, whereas the younger cohort experienced the reformed system.

Meghir and Palme employ a difference-in-differences estimator:

$$Y_{idm} = b_0 + b_1 d_i + b_2 m_i + \alpha r_{idm} + \gamma x_{idm} + e_{idm} \tag{39}$$

where:

- Y_{idm} denotes the outcome of interest.
- $d_i = 1$ if individual i belongs to cohort d.
- $m_i = 1$ if individual i resides in municipality m.
- $r_{idm} = 1$ if individual i belongs to cohort d and lives in a municipality that adopted the reformed system.

The coefficient of interest is denoted by α. Meghir and Palme analyze the effects of the reform on educational attainment and earnings.

4.2 Results for Meghir and Palme (2005)

Table 1 presents the results on education:

- An increase in years of schooling, with a greater effect observed among women than men.
- A pronounced increase in schooling among individuals with less-educated fathers.
- Minimal differences in impact across ability levels.

Table 2 reports the results on earnings:

- No overall effect on earnings; however, significant heterogeneity is observed.
- A negative impact on earnings for individuals with highly educated fathers, a group for whom there was no observed effect on educational attainment.

Returns to education among individuals with less-educated fathers are estimated as follows:

- Across all ability levels: 3.36/0.405 = 8.3%.
- Among low-ability individuals: 2.62/0.468 = 5.6%.
- Among high-ability individuals: 4.53/0.355 = 12.8%.

The decline in earnings among individuals with highly educated fathers prompts further analysis through the lens of supply and demand for educated workers. The key questions pertain to shifts in equilibrium wages and labor market conditions.

Two primary mechanisms may underlie these effects:

1. Large-scale changes in the supply of educated workers may generate general equilibrium effects.
    - The elasticity of substitution between younger and older workers plays a crucial role.
    - The extent to which labor markets operate at a local (municipality-specific) level is also a factor.
2. Significant educational reforms may also influence the quality of schooling.
    - If teacher supply is inelastic, the influx of additional students may lead to a decline in educational quality.

4.3 Duflo (2001)

Duflo (2001) examines the 1973 school construction program in Indonesia. Between 1973-74 and 1978-79, over 61,000 primary schools were constructed, averaging two new schools per 1,000 children aged 5-14 in 1971. Enrollment rates increased from 69% in 1973 to 83% in 1978.

The study utilizes household survey data from the 1995 Census, covering men born between 1950 and 1972. Regional birth data on school construction between 1973-74 and 1978-79 were merged with this dataset.

Duflo (2001) estimates the following equation:

$$S_{ijk} = c_1 + \alpha_{1j} + \beta_{1k} + (P_j T_i)\gamma_1 + (C_j T_i)\delta_1 + \varepsilon_{ij} \tag{40}$$

where:

- i indexes individuals, j indexes regions, and k indexes birth years.
- $P_j T_i$ represents the treatment status, where T_i denotes individuals aged 2-6 in 1974, and P_j identifies regions experiencing substantial school construction.

The coefficient of interest is γ_1. This study provides valuable insights into the impact of large-scale educational investments on schooling attainment and labor market outcomes.

5 Education as Signaling

5.1 Spence (1973)

Background. The key observation in this model is that there is an information problem: employers cannot directly observe worker productivity.

5.1.1 Pooling Equilibrium

For simplicity, assume there are two levels of productivity, 1 and 2. If the population consists of two types in equal proportion (50/50) and no signal is provided, then the wage is set at 1.5. However, there exists an incentive for firms to separate the two types and for higher productivity individuals to distinguish themselves. In this context, a "signal" is any observable characteristic that can be used to identify high productivity versus low productivity individuals. Education may sometimes serve as such a signal.

5.1.2 Separating Equilibrium

For education to serve as an effective signal, the cost of acquiring education must be negatively related to productivity. In other words, the more productive types should be the ones to acquire education. Low productivity types will not find it worthwhile to acquire the signal, while high productivity types will. There is an optimal standard, denoted by s^*, which is set just high enough to separate the two groups. If the standard is set too low, then low productivity types will also acquire education; if it is set too high, even high productivity types might refrain from doing so. (Remember that in this model, schooling does not increase productivity.)

5.1.3 Social Return vs. Private Return to Education

In this framework, education yields a private return for those who pursue it. However, because education does not enhance workers' productivity overall, it represents a waste from society's perspective (unless one introduces the concept of social gains through sorting the right workers to the right jobs). Thus, the social return to education is less than the private return in the signaling model.

5.1.4 Discussion Question

The United States spends enormous amounts on education (education expenditures constitute a large percentage of GDP). Would it not be easier and cheaper for employers to administer an aptitude test?

6 Employer Learning

Suppose that education serves as a partial signal of worker ability. It is plausible to assume that employers gradually *learn* about the unobserved components of worker ability over time. Two key papers in this area are:

- Farber and Gibbons, "Learning and Wage Dynamics" (QJE 1995)
- Altonji and Pierret, "Employer Learning and Statistical Discrimination" (QJE 2001)

6.1 Farber and Gibbons (1995)

6.1.1 Model of Employer Learning and Wage Determination

Let η_i denote the innate ability of worker i, and let s_i denote the schooling of worker i. While η_i is unobserved by employers, s_i is observable.

In addition to s_i, consider three types of time-invariant worker characteristics:

(i) X_i: Characteristics observed by both employers and included in the data.

(ii) Z_i: Characteristics observed by employers but not recorded in the data (e.g., the quality of schooling, such as the specific institution attended).

(iii) B_i: Measures correlated with ability η_i that are not observed by the employer but are available in the data (e.g., aptitude tests such as the AFQT score).

There is a joint distribution, $F(\eta_i, s_i, X_i, Z_i, B_i)$, known by all employers. However, employers do not observe the particular type of worker who enters the labor market (analogous to the Spence setup).

6.1.2 Worker Output and Wage Determination

Let y_{it} denote the output of worker i in the t^{th} year after entering the labor market. Assume that the observed outputs $\{y_{it} : t = 1, \ldots, T\}$ are independent draws from the conditional distribution

$$G\left(y_{it} \mid \eta_i, s_i, X_i, Z_i\right).$$

Note that B_i is not included in this conditional distribution because it does not have a direct effect on output.

Farber and Gibbons (1995) simplify the model of output and wage determination by abstracting from other important channels of wage dynamics (e.g., why wages rise with experience). Note also that the variable t does not appear in the function $G(\cdot)$, implying that there is no worker productivity growth with experience, nor do returns to experience interact with s_i. As in the Spence signaling model, worker productivity is determined solely by time-invariant characteristics.

6.1.3 The Information Environment

The model assumes a "public information" environment in which all employers share the same information. Specifically, every employer knows the joint distribution $F(\eta_i, s_i, X_i, Z_i, B_i)$, the conditional distribution $G\left(y_{it} \mid \eta_i, s_i, X_i, Z_i\right)$, and also the individual values of s_i, X_i, Z_i, along with the stream of output $\{y_{i1}, \ldots, y_{it}\}$ for each worker.

In contrast, there are models with "private information" where only the employer of worker i observes the worker's output and infers the worker's innate ability, potentially leading to monopsony power over the worker.

The wage is determined as the expected productivity of the worker, given all available information at time t:

$$w_{it} = E\Big(y_{it} \mid s_i, X_i, Z_i, y_{i1}, \ldots, y_{it-1}\Big).$$

Labor markets are assumed to be competitive, and this is a spot market model in which workers are paid for their period-specific productivity (thereby ruling out incentive contracts).

6.1.4 Predictions from the Model

Based on these assumptions, Farber and Gibbons (1995) derive three predictions:

Prediction 1: Constant Effect of Schooling on Wages Consider the regression specification

$$w_{it} = \alpha_t + \beta_t s_i + X_i \gamma_t + \varepsilon_{it}.$$

Here, $(\hat{\alpha}_t, \hat{\beta}_t, \hat{\gamma}_t)$ denote the coefficients from the linear projection of w_{it} on s_i and X_i, such that

$$E^*(w_{it} \mid s_i, X_i) = \hat{\alpha}_t + \hat{\beta}_t s_i + X_i \hat{\gamma}_t.$$

Since the predicted wage equals the predicted output, i.e.,

$$E^*(w_{it} \mid s_i, X_i) = E^*(y_{it} \mid s_i, X_i),$$

and because $E^*(y_{it} \mid s_i, X_i)$ is independent of t, the effect of schooling on wages is invariant with respect to experience. In period 1, the forecast error is given by

$$\varepsilon_i = y_{i1} - E^*(y_{i1} \mid s_i, X_i),$$

which, by construction, is orthogonal to the variables used to predict output in period 1. Thus, future observations, on average, confirm the initial relationship between expected productivity and schooling.

Prediction 2: Increasing Effect of Unobserved Characteristics with Experience Since B_i (AFQT) is unobserved by the employer but present in the data, and may be correlated with observed variables (s_i, X_i, Z_i), define the residual

$$B_i^* = B_i - E^*(B_i \mid s_i, X_i, w_{i1}).$$

Regressing B_i on w_{i1} removes any correlation between B_i^* and Z_i (i.e., information already known by employers). Next, consider the wage regression:

$$w_{it} = \alpha_t + \beta_t s_i + X_i \gamma_t + B_i^* \pi_t + \varepsilon_{it}.$$

Because B_i^* is constructed to be orthogonal to the other regressors,

$$\hat{\pi}_t = \frac{\text{cov}(B_i^*, w_{it})}{\text{var}(B_i^*)}.$$

Given that

$$w_{it} = w_{it-1} + \xi_{it} = w_{i1} + \sum_{t=2}^{t} \xi_{it},$$

and since B_i^* is orthogonal to w_{i1} (implying $\hat{\pi}_1 = 0$), we have

$$\text{cov}(B_i^*, w_{it}) = \sum_{t=2}^{t} \text{cov}(B_i^*, \xi_{it}).$$

Because the wage innovations ξ_{it} are likely positively correlated with ability η_i, the estimated effect of B_i^* on wages should increase with experience as output signals gradually reveal worker ability.

Prediction 3: Wages are a Martingale The model predicts that wages follow a martingale process:

$$E(w_{it} \mid w_{it-1}) = w_{it-1}.$$

Although empirical evidence rejects this prediction, the primary focus remains on testing Predictions 1 and 2.

6.2 Altonji and Pierret (2001)

6.2.1 Combining Employer Learning and Statistical Discrimination

Altonji and Pierret extend the Farber and Gibbons framework by incorporating elements of statistical discrimination. They note that if B_i is used in the regression instead of B_i^*, then:

(1) The estimated effect of B_i on earnings increases with experience.

(2) The estimated effect of s_i on earnings decreases with experience.

This occurs when B_i and s_i are positively correlated.

The model can thus be reinterpreted as one of statistical discrimination. When workers enter the labor market, employers have access only to observable signals such as schooling. Lacking additional information, employers infer worker productivity from these signals. As more output signals (y_1, y_2, $\ldots$) become available with experience, the measured effect of B_i (which is correlated with unobserved ability) increases, while the signaling value of schooling diminishes.

Empirical Implications. Results are consistent with employers using statistical discrimination with respect to schooling. However, the evidence does not support a similar role for race. Instead, over time, race appears to behave like the AFQT variable—becoming correlated with unobserved ability and output signals—with the coefficient on race increasing over time and becoming more negative. This raises further questions regarding the underlying mechanisms.

7 Income Inequality vs. Instability/Mobility

Juhn, Murphy, and Pierce (1991) document rising inequality by percentile position. The analysis assumes that the individual at the 10th percentile in year t remains at the 10th percentile in year $t+1$. In this framework, while individuals maintain their relative positions, the absolute distances between them widen. Alternatively, if individuals switch positions, the interpretation is quite different.

With panel data, one can decompose the cross-sectional variance of earnings into permanent and transitory components via the equation:

$$y_{it} = p_t \alpha_i + \varepsilon_{it},$$

where:

- y_{it} is the log earnings of individual i in year t,
- α_i is an individual fixed effect (interpreted as "ability") with variance σ_α^2,
- p_t represents the price of skill/ability, which varies by year,
- ε_{it} denotes individual-level transitory shocks with variance σ_t^2.

Thus, the variance of log earnings is given by:

$$\text{Var}(y_{it}) = p_t^2 \, \sigma_\alpha^2 + \sigma_t^2.$$

The observed rise in cross-sectional income inequality could be attributed either to an increase in permanent differences across individuals or to an increase in the variance of the transitory component. The latter is referred to as income "instability," which is generally viewed as undesirable—even though instability implies mobility.

Consider the issue from an intergenerational perspective. Suppose we compare two economies, A and B:

- In both economies, cross-sectional income inequality within a given generation is the same.
- In economy A, sons maintain the same relative position as their fathers.
- In economy B, each generation resets, meaning that a son's relative position is independent of his father's.
- The higher mobility observed in economy B might be interpreted as indicative of a more equitable system.

8 Intergenerational Elasticity (IGE)

The most widely used measure of intergenerational mobility is the *intergenerational elasticity* (IGE). The standard specification is:

$$\log Y_1 = \beta \log(Y_0) + \varepsilon,$$

where $\log Y_1$ and $\log Y_0$ denote the log incomes of the child and parent, respectively (expressed as deviations from the population mean). The coefficient β represents the elasticity of the child's income with respect to the parent's income (i.e., the IGE).

An alternative formulation expresses β as:

$$\beta = \rho \frac{\sigma_1}{\sigma_0},$$

where ρ is the correlation between parental and child incomes, and σ_1 and σ_0 denote the standard deviations of child and parental incomes, respectively. This representation illustrates that the IGE is scaled by the relative variances of income in the two generations. If income inequality is larger in the child's generation than in the parent's generation, the IGE will be higher for a given correlation ρ.

8.1 Measurement Issues in Estimating the IGE

(1) **Permanent vs. Transitory Income:** Ideally, the correlation should reflect the permanent components of parental and child incomes. Averaging incomes over several years is recommended.

(2) **Age Bias:** The life-cycle earnings profile tends to stabilize around age 40. Using earnings at age 25 may lead to bias; thus, earnings at age 40 (if available) should be used.

(3) **Sample Selection:** Estimates may vary depending on the sample. For example, a homogeneous sample of white males may yield different results compared to those based on representative survey data.

8.2 Empirical Estimates of the IGE

(1) **Becker and Tomes (1986):** Early surveys indicate an IGE of approximately 0.2, though these studies often relied on single-year earnings data and selected samples.

(2) **Solon (1999):** Survey data discussed in the Handbook chapter suggest an IGE of approximately 0.4.

(3) **Mazumder (2005):** More recent estimates for the U.S. suggest an IGE in the range of 0.5 to 0.6.

It remains unclear whether the IGE has increased over time in the U.S. (implying lower mobility) or whether improved measurement techniques have led to higher estimates. In any case, intergenerational mobility in the U.S. appears to be lower than originally thought.

8.3 Chetty, Hendren, Kline, and Saez (2014)

This study utilizes tax records in which parents and children are matched via tax returns (when parents claim children as dependents). The advantages include a very large dataset that permits exploration of non-parametric relationships. However, there are several disadvantages:

- Data span the period 1996–2012.
- The focus is on a single cohort of children born between 1980 and 1982.
- The children are around 30 years old.

The relationship in levels is highly non-linear. The log-log specification of the IGE faces several issues:

- The IGE calculation drops zero values, which is significant for child income.
- Dropping zeros tends to overstate intergenerational mobility.
- The non-linearity results in a lower IGE when including outlier observations (approximately 0.344).
- Restricting the analysis to the 90th and 10th percentiles yields an IGE of approximately 0.452.

An alternative measure is the *rank-rank slope*, which involves regressing the child's rank in the income distribution on the parent's rank. This approach yields a linear relationship that is robust to outliers and facilitates cross-country comparisons.

Thus far, the focus has been on relative mobility measures, which are inherently zero-sum. In contrast, absolute mobility measures have also been examined. For example, CHKS study absolute mobility by analyzing the mean child rank for parents in the bottom 25th percentile. They further examine this measure at the commuting zone (CZ) level, finding considerable variation across CZs and regions (e.g., 46.2 in Salt Lake City versus 35.8 in Charlotte). Factors such as residential segregation, income inequality, and social capital also emerge as important correlates.

Unpacking the IGE and other intergenerational mobility measures is a challenging task. The subsequent discussion will introduce some basic theoretical frameworks underlying these measures.

9 Quantity-Quality Trade-off

9.1 Quantity-Quality Model of Parental Investment

We have discussed Becker and Tomes (1979) which suggests that children's outcomes depend not only on parental resources but also on parental choices.

Parents choose between their own consumption and investment in children.

Becker and Lewis (1973) suggested that parents choose not only child quality but also child quantity and invest accordingly.

Subsequent work has focused on identifying exogenous shocks to fertility based on the conjecture summarized in Becker (1981):

"If [prices and income] were held constant, an exogenous increase in quantity would raise the shadow price of quality, and thereby would reduce the demand for quality."

Studies use exogenous variation in family size using twin births and sibling sex composition as instruments.

Mixed evidence:

- Black, Devereux and Salvanes (2005)
- Angrist, Lavy, and Schlosser (2009)
- Rosenzweig and Wolpin (1980), Rosenzweig and Zhang (2009), Li, Zhang, and Zhu (2008)
- Conley and Glauber (2006), Caceres-Delpiano (2006)

This literature has almost exclusively focused attention on the first-born.

It has also ignored the *timing* of shocks to fertility.

9.2 Black, Devereux, and Salvanes (2005)

Black, Sandra, Paul J. Devereux and Kjell G. Salvanes. 'The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." The Quarterly Journal of Economics, Vol. 120, No. 2 (May, 2005), pp. 669-700.

Use data from Norway 1986–2000.

Can match parents, children, siblings.

Importantly they observe completed fertility for mothers.

Children's outcome measure is completed education as function of family size.

$$ED = \beta_0 + \beta_1 \text{ FAMSIZE} + X\beta_2 + \varepsilon \tag{1}$$

$$\text{FAMSIZE} = \alpha_0 + \alpha_1 \text{ TWIN} + X\alpha_2 + \nu \tag{2}$$

ED = education of child at nth birth-order

$TWIN$ = indicator that n+1 birth is twin birth

BDS (2005) find little effect of family size once they instrument for family size with twin birth and control for birth-order.

They postulate that what is important are birth-order effects and not overall effect of family size.

9.3 Quantity-Quality Model: Basic Setting

We build on Becker and Lewis (1973) where parents choose own consumption, the number of children, and level of human capital (quality) of each child.

Parents maximize the following log-linear utility function:

$$\ln U = (1-\alpha)\ln C + \alpha\Big(\ln N + \pi \ln \bar{H}\Big) \tag{1}$$

Subject to:

$$C + N(w\tau + I) = C + CLD = Y + w \tag{2}$$

where $w\tau$ is the foregone earnings ignoring child quality investments associated with raising a child.

Y stands for all other household income including partner's earnings.

$\bar{H}$ is the equality-adjusted measure of children's quality:

$$\bar{H} = \left(\sum_{i=1}^{N} \frac{H_i^{\beta}}{N}\right)^{\frac{1}{\beta}} \tag{3}$$

The parameter $\beta \in (-\infty, 1)$ represents parents' taste for equality between children. $\sigma_\beta = \left(\frac{1}{1-\beta}\right)$ is the elasticity of substitution across human capital of each child.

Technology of human capital production follows Heckman (2017):

$$H = A\Big[\gamma(I_1)^{\phi} + (1-\gamma)(I_2)^{\phi}\Big]^{\frac{1}{\phi}} \tag{4}$$

The optimal ratio of early to late investment is a function of γ and ϕ:

$$\frac{I_1}{I_2} = \left[\frac{\gamma}{1-\gamma}\right]^{\sigma_\phi} \tag{5}$$

The ratio of early to late investment increases with γ and more so when I is relatively easy to substitute between early and late investments ($\sigma_\phi > 1$).

The optimal mixture of child quantity and child quality is shown in the following equations:

$$N^* = \alpha \left(\frac{Y+w}{w}\right) \frac{(1-\pi)}{\tau} \tag{6}$$

$$I^* = \frac{w\pi\tau}{(1-\pi)} \tag{41}$$

$$N^* I^* = \alpha\pi(Y+w) \tag{42}$$

$$N^* w\tau = \alpha(1-\pi)(Y+w) \tag{43}$$

It is important to note that in the absence of shocks, quantity (N^*) and quality (H^*) are jointly determined.

The negative relationship between the optimal quantity and the optimal quality reflects the opposite influences π and $w\tau$ have on quantity vs. quality.

9.4 Quantity-Quality Model: Shock to Family Size

Following a shock to family size parents maximize the following log-linear utility function:

$$\ln U = (1-\alpha)\ln C + \alpha\pi \ln \bar{H} \tag{7}$$

Subject to:

$$C + \tilde{N} I = \left(Y + w - \tilde{N} w\tau\right) - \delta I^* \tag{8}$$

where $\tilde{N}$ represents the new family size and $\tilde{N} > N^*$.

δI^* represents investments in the first-born that had already taken place prior to the shock.

If $\delta I^* = 0$, parents split family resources between own consumption ($\tilde{C}$) and investments ($\tilde{N}\tilde{I}$) in the following way:

$$\tilde{N}\tilde{I} = \frac{\pi\alpha}{(1-\alpha+\alpha\pi)}\left(Y + w - \tilde{N} w\tau\right) \tag{9}$$

$$\tilde{C} = \frac{1-\alpha}{(1-\alpha+\alpha\pi)}\left(Y + w - \tilde{N} w\tau\right) \tag{44}$$

In this case parents invest equally less in each child:

$$\tilde{I} = \frac{\pi\alpha}{\tilde{N}(1-\alpha+\alpha\pi)}\left(Y + w - \tilde{N} w\tau\right) \tag{10}$$

Note in this case that higher $\tilde{N}$ reduces $\tilde{I}$ in the causal sense.

The change in investment per child is as follows:

$$\Delta\tilde{I} = -\frac{\pi w \tau}{(1+\alpha(\pi-1))(1-\pi)\left(\alpha\left(\frac{Y+w}{w}\right)\frac{(1-\pi)}{\tau} + 1\right)} < 0 \tag{45}$$

The drop $\Delta\tilde{I}$ is larger the higher the mother's wage (w) and smaller the higher other income (Y):

$$\frac{\partial \Delta\tilde{I}}{\partial w} < 0, \quad \frac{\partial \Delta\tilde{I}}{\partial Y} > 0 \tag{46}$$

If $\delta I^* > 0$, parents no longer invest equally in the first-born and newborn children.

The extent to which they deviate from equality depends on the parameters β and ϕ.

Ruling out extreme cases, $\beta = -\infty$ and $\phi = 1$, a shock to family size results in higher investment in the first-born.

Moreover, the drop in investment in the first-born is smaller (larger) as birth spacing increases.

9.5 Quantity-Quality Model: Takeaways

The adverse impact of a shock to family size on the first-born diminishes with birth spacing.

This is due to the mechanical effect of the production function, i.e., investments in early childhood are more productive.

Parents also amplify this effect by endogenously choosing to invest more in the first-born.

Firstborn children serve as credible proxies for all children only at short spacing intervals.

Most likely, focusing on the first-born *understates* the quantity-quality trade-off.

9.6 Quantity-Quality Model: Implications for Empirical Work

(1) Performance of firstborn children drop following a shock to family size.

(2) The drop in performance depends on birth spacing.

(3) The drop in performance is larger in more disadvantaged households (low Y).

9.7 Estimating Effect of Twin Births on Older Children by Birth Spacing

We estimate the following equation:

$$P_{it} = \beta_0 + \beta_T TWIN_i + \beta_{TS}(TWIN_i \cdot S_i) + \beta_s S_i + \beta_M M_i + \beta_x X_{it} + \varepsilon_{it} \tag{47}$$

P_{it} denotes child i's test score at time t.

$TWIN_i$ is an indicator if the next sibling is a twin.

S_i is an indicator if birth spacing is equal to or greater than 3 years.

We expect β_T to be negative and β_{TS} to be positive.

9.8 Quantity-Quality Recap

The original Becker and Lewis (1973) posited that parents make choices in terms of both child quantity and child quality, which are jointly determined.

While mother's wage and preference for quality have opposite influences, the model does not imply a causal relationship.

The empirical literature has focused on the causal relationship $\Rightarrow$ whether unanticipated shocks to family size (quantity) reduce investment in human capital of children.

The null result often found may be due to:

(1) The impact of shocks to fertility on education of first-borns may *understate* the negative impact on younger siblings and on the average.

(2) The impact may also depend on birth spacing (if early investments are what matter as Heckman (2017) postulates).

There is evidence that shocks matter in disadvantaged households.

10 Intergenerational Mobility – Simple Theoretical Model

10.1 Simple Theoretical Model

Based on Solon (1999), Becker and Tomes (1979).

Simplify by considering a family with 1 parent and 1 child.

$$y_{t-1} = C_{t-1} + I_{t-1} \tag{1}$$

y_{t-1} = parent's lifetime earnings
C_{t-1} = parent's consumption
I_{t-1} = investment in the human capital of the child

$$y_t = (1+r)I_{t-1} + E_t \tag{2}$$

y_t = child's lifetime earnings
r = return to human capital investment
E_t = child's "endowment" (could be genetic inheritance but also "culture", connections; could encompass both "nature" and "nurture")

Parents maximize the following Cobb-Douglas utility function:

$$U = (1-\alpha)\log C_{t-1} + \alpha \log y_t \tag{3}$$

α = altruism parameter

F.O.C. $\Rightarrow$

$$I^*_{t-1} = \alpha y_{t-1} - (1-\alpha)\frac{E_t}{(1+r)} \tag{4}$$

Substituting (4) for I^*_{t-1} into (2):

$$y_t = \beta y_{t-1} + \alpha E_t \tag{5}$$

where $\beta = \alpha(1+r)$

Although (5) looks like the original IGE regression, what is the problem?

Becker and Tomes (1979) go further to unpack E_t:

$$E_t = e_t + u_t \tag{6}$$

e_t = child's "endowment" which is inherited from parents (aside from parent's intentional investment via I_{t-1})
u_t = "market" luck

$$e_t = \lambda e_{t-1} + v_t \tag{7}$$

(7) is an AR(1) process with persistence $0 < \lambda < 1$.

v_t is the endowment shock that is orthogonal to parent's endowment.

Substituting (6) into (5):

$$y_t = \beta y_{t-1} + \alpha e_t + \alpha u_t \tag{8}$$

$$\sigma^2_e = \frac{\sigma^2_v}{1-\lambda^2} \tag{48}$$

E = e + u

e is the endowment and u is market luck.

$$e_t = \lambda e_{t-1} + v \tag{49}$$

λ is the persistence of endowment across parent and child.

How much variation in y is due to market luck and how much is due to endowment.

Extreme case: no variation in endowment shock, $\sigma_e^2 = \frac{\sigma_v^2}{1-\lambda^2}$ or $\lambda = 0$ (no inheritance of endowment) then back to β.

More generally, the correlation between y_t and y_{t-1} generated by the model in (8):

$$\text{Corr}(y_t, y_{t-1}) = \delta\beta + (1-\delta)\left[\frac{(\beta+\lambda)}{(1+\beta\lambda)}\right] \tag{9}$$

$$\delta = \frac{\alpha^2\sigma_u^2}{[(1-\beta^2)\sigma_y^2]} \tag{10}$$

The correlation is a weighted average where weights depend on relative variances σ_u^2 and σ_y^2, with $\frac{(\beta+\lambda)}{(1+\beta\lambda)} > \beta$ if $\lambda > 0$.

Unpacking the intergenerational correlation:

(1) In the simplest form, $\beta = \alpha(1+r)$
This term exists because parents invest, and care about children via α. Investment depends on parent's income, y_{t-1}.

(2) More complicated effects arise due to the transmission of "endowment" across generations.
What is "endowment"? Genes? Behavioral practices? Culture? Connections?
These forces increase the intergenerational correlation further.

10.2 Intergenerational Transmission: Role of Genetics

Houmark, Mikkel Aagaard, Victor Ronda, and Michael Rosholm. 2024. "The Nurture of Nature and the Nature of Nurture: How Genes and Investments Interact in the Formation of Skills." American Economic Review 114 (2): 385–425.

Genetic transmission:

Children inherit 2 sets of 23 chromosomes, one from each parent.

For most of DNA there is not variation across population but there are locations where there is variation, called SNPs (single nucleotide polymorphisms).

$g_{is} \in \{0, 1, 2\}$

- 0 = GG (all common allele G and no uncommon allele C)
- 1 = GC
- 2 = CC

Mother: 0 2 2
Father: 1 2 1
Child: 1 2 2

The child's genetic make-up is pre-determined, but conditional on both parents' genetic make-up, it is exogenous or random.

Houmark et al. (2024) use Avon Longitudinal Study of Parents and Children (ALSPAC) which is a British panel study of women recruited at pregnancy during 1991–1992.

Collected DNA information of mother, child, and some subset of fathers.

Questionnaire mailed to mothers on child development ("skills") and time spent in activities ("investment").

Estimate skill formation (production function):

$$\theta_{it+1} = f_t^{\theta}(\theta_{it}, I_{it}, G_i, G_i^M, G_i^F, X_{it}) \tag{50}$$

Estimate investment function:

$$I_{it} = f_t^{I}(\theta_{it}, G_i, G_i^M, G_i^F, X_{it}) \tag{51}$$

Child's genetic factor has direct effect on skill formation.

Child's genetic factor has indirect effect because parents respond to child's genetic factor.

But even more importantly, conditional on child's genetic factor, parents' genetic factors G_i^M, G_i^F have direct effect on children's skill formation and investment in children.

11 Returns to Parental Investment: Time and Money

11.1 Human Capital Production Function

So far we posited that parents make deliberate choices and invest in the human capital of their children. We considered the following type of human capital production function:

$$H = A\left[\gamma(I_1)^{\phi} + (1-\gamma)(I_2)^{\phi}\right]^{\frac{1}{\phi}} \tag{52}$$

where we contrasted the efficacy of early investment (I_1) vs later investment (I_2). Investment also may take different forms of family resources– money/income or time.

Question: What is the efficacy of parental income and time in the human capital development of children?

- Series of studies pioneered by Cunha and Heckman, Matt Wiswall
- Dahl, G. and L. Lochner (2012), "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit," *American Economic Review*, 102(5): 1927-56.
- Agostinelli, F. and G. Sorrenti (2022), "Money vs Time: Family Income, Maternal Labor Supply, and Child Development," HCEO Working Paper No. 2018-017, R&R at *Journal of Political Economy*

11.2 Dahl and Lochner (2012)

We might be interested in estimating the impact of family income on children's achievement such as the following:

$$y_{ia} = x_i'\alpha_a + w_{ia}'\beta + I_{ia}\delta_a + \cdots I_{ia-L}\delta_{a-L} + \mu_i + \varepsilon_{ia} \tag{53}$$

y_{ia} = measures of child achievement or skills
x_i' = permanent observable characteristics, i.e. gender, race
w_{ia}' = time varying characteristics
I_{ia} = family income and lagged incomes can also matter

In cross-section I_i is correlated with the error term. With FE, one can account for permanent unobserved characteristics:

$$\Delta y_{ia} = x_i'\alpha + \Delta w_{ia}'\beta + \Delta I_{ia}\delta_a + \cdots \Delta I_{ia-L}\delta_{a-L} + \Delta\varepsilon_{ia} \tag{54}$$

Still not out of the woods since unobserved shocks can generate income shocks and also affect children's achievement growth. Dahl and Lochner (2012) use EITC policy changes to instrument for change in family income. Earned Income Tax Credit (EITC) is a wage subsidy for working. EITC is a function of pre-tax income. Want to predict EITC using last year's pre-tax income and changes in the EITC schedule.

We might be interested in estimating the impact of family income on children's achievement such as the following:

$$\Delta\chi_{ia}^{IV}(P_{i,a-1}) = \chi_a^{s_{i,a-1}}(\hat{E}[P_{i,a} \mid P_{i,a-1}]) - \chi_a^{s_{i,a-1}}(P_{i,a-1}) \tag{55}$$

Back to the original estimating equation:

$$\Delta y_{ia} = x_i'\alpha + \Delta w_{ia}'\beta + \Delta I_{ia}\delta_a + \Delta\varepsilon_{ia} \tag{56}$$

δ_a may still be biased because $\Delta\chi_{ia}^{IV}(P_{i,a-1})$ is a function of $P_{i,a-1}$ which is likely to be correlated with the subsequent change in income due to measurement error, regression to the mean, and serially correlated income shocks.

Estimate the following:

$$\Delta y_{ia} = x_i'\alpha + \Delta w_{ia}'\beta + \Delta I_{ia}\delta_a + \Phi(P_{i,a-1}) + \eta_{ia} \tag{57}$$

where $\Phi(P_{i,a-1})$ is the "control function" and is in practice a dummy for positive lagged pre-tax income and a fifth-order polynomial of lagged pre-tax income.

Data: NLSY79 Mother-children data
Outcome measures are children's PIAT scores
Estimates suggest that a \$1000 increase in family income increases child's cognitive scores by ∼6% of std units. The effects appear to be larger for Blacks, for disadvantaged families, and for boys. The effects are larger than OLS, FE specifications:

- May be that the IV addresses measurement error
- EITC focuses on low-income populations where the effect may be larger
- The EITC changes may be reflecting more permanent income changes
- Check that it is deposited to family bank account

One issue that is ignored is that EITC schedule change may impact mother's labor supply and therefore also mother's time input. We need two instruments! One for income and one for mother's time.

11.3 Agostinelli and Sorrenti (2022)

Review of labor supply responses to EITC:

- In low wage regions ⇒ increase in incentives to work
- In mid to high wage regions ⇒ the income effect may dominate and may even reduce the incentive to work

The impact of EITC on children may have different effects depending on the labor incentives for mothers to increase labor supply.

Back to the human capital production function:

$$\theta_{t+1} = \left[\alpha_2\theta_t^{\phi_2} + (1-\alpha_2)X_t^{\phi_2}\right]^{\frac{1}{\phi_2}} \tag{58}$$

$$X_t = \left[\alpha_1(\tau - L_t)^{\phi_1} + (1-\alpha_1)g_t^{\phi_1}\right]^{\frac{1}{\phi_1}} \tag{59}$$

The extent to which money can compensate for reduction in mother's time will depend on the elasticity of substitution between the two inputs in the production function.

Estimate the following equation:

$$\Delta y_{it} = \alpha_0 + \alpha_1 \Delta I_{it} + \alpha_2 \Delta L_{it} + x_i' \beta_{1t} + x_{it}' \beta_2 + \Delta \varepsilon_{it} \tag{60}$$

The above has two endogenous variables ΔI_{it}, ΔL_{it}. Use Δ EITC schedule and local labor demand shock for female labor $LabDemShocks_{it}$ as the two instruments.

Children of low-wage, single mothers may experience large negative impact of EITC expansion through 2 channels:

(1) Mothers may have increased labor supply more in response.

(2) The elasticity of substitution between mother's time and income may be lower.

12 Education Production Function and Debate over School Resources

12.1 Debate over "Does Money Matter?"

- Proponents argue that school spending, per se, matters.
- The opposition emphasizes *incentives* – competition, school choice, and teacher incentives such as performance pay.

Hanushek, "The Failure of Input-Based Schooling Policies," *Economic Journal* 2002.
Time-series evidence

- Between 1890-1990, real expenditure per student grew 3 1/2 percent annually.
- Pupil-teacher ratio fell, fraction of teachers with MA increased, expenditures in 2000-2001 increased from \$2,235 in 1960 to \$7,500.
- However, there is not much improvement in test scores over time.
- Also, U.S. students perform in a mediocre fashion relative to other countries.

12.2 Cross-school evidence

Coleman Report (1966) — officially called the "Equality of Educational Opportunity," a government report prepared by sociologist James Coleman:

- Following the Civil Rights Act in 1964, the report was to investigate differences in schooling resources by race.
- Two surprising conclusions: (1) differences in schooling resources across race are not as large as we thought; (2) school resources had small and uncertain effects on student achievement once family background variables and peer variables were controlled for.

Cross-school comparisons — in U.S. parents choose schools and school districts — the variation in family background across schools and neighborhoods is highly correlated with variation in school resources.

12.3 Education Production Function

Education "Production" Function:

$$A_{it} = F(F_{it}, P_{it}, S_{it}) \tag{61}$$

A = student achievement
F = family inputs
P = peer inputs
S = school inputs

If schools are chosen by parents, it is hard to think of this as a causal relationship. If panel data are available, a value-added specification is often used:

$$A_{it-1} = \text{student achievement in the previous period}$$

This does not really help if parents are choosing schools. We will also investigate teachers where, within a school, random assignment of teachers to students is assumed. In a value-added specification, the hope is that A_{it-1} can account for the entire history of past inputs (i.e. a sufficient statistic).

Hanushek surveys cross-sectional studies (meta-analysis) and concludes there is not much evidence that increased spending increases student performance. His interpretation is NOT that school resources would never matter; rather, the interpretation is that more money in the current system will not produce results — the problem is the lack of incentives built into the current system.

12.4 Schooling Quality Does Matter

Card and Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment" (QJE 1992)

- So far, we have talked about studies that have largely relied on cross-sectional variation and looked at outcomes such as test scores.
- Another strand of the literature examines the effect of schooling quality on individual earnings (continual debate over test scores vs. earnings which we will revisit).

Card and Krueger conduct 2-step estimation.
First step:

$$y_{ijs} = \rho_s E_{ijs} + X_{ijs}\beta + \alpha_j + \mu_s + \varepsilon_{ijs} \tag{62}$$

E_{ijs} is years of education.
ρ_s is returns to education for individuals born in state s. Estimate the above by birth cohort and race for different Census years (1960, 1970, 1980), and end up with ρ^c_{str}.

They also focus on men born in the South and living in 9 Metro areas in the North.
Second step:

$$\rho^c_{str} = bQ^c_{sr} + \alpha^c_r + \mu_s + \upsilon_r + \eta_{str} \tag{63}$$

Q^c_{sr} is a measure of average school quality for a race-cohort-state group.

- They find large gaps in schooling across segregated white and black schools in the South.
- Between 1915-1964, there were dramatic improvements in schooling quality.
- Table VI shows that returns to schooling are much lower for blacks, but also the gap is larger for states where the schooling quality gap is large.

Estimates of 2nd step (Table VII):

- Pupil-teacher ratio reduces returns to schooling.
- Column (1) without pupil-teacher ratio: between 1910-1919 cohort and 1930-1939 cohort, the black-white gap in returns to schooling closed 40% from 3.31 to 2.02.
- Addition of pupil-teacher ratio, column (2): unexplained convergence goes from (1.29 = 3.31 - 2.02) in column (1) to (0.59 = 2.08 - 1.49) in column (2).
- Pupil-teacher ratio can account for 54 percent of the convergence.

What is the effect of school quality on earnings?
Card and Krueger fit reduced form equations of the following variety:

$$y^c_{btr} - y^c_{wtr} = \beta(Q^c_{br} - Q^c_{wr}) + \mu^c_r + \nu_t + \varepsilon^c_{str} \tag{64}$$

Reduced form specifications:

- The reduced form models exclude education.
- The coefficient, β, would reflect both the direct effect of schooling quality and the indirect effect through increases in years of schooling.

Table X:

- Compare cohort effects in column (1) and column (2) without and with pupil-teacher ratio.
- Cohort effects go up 0.066 for the 1930-1939 cohort but 0.038 with control (about 40%).

What Have We Learned?

- Which output measure we use matters — test scores vs. earnings.
- Inputs matter when there are big changes — evidence from segregated black schools (diminishing returns?).
- The effect of schooling inputs is swamped by the effect of other inputs, particularly if we use test scores as output measures and compare across schools.
- Schooling inputs may matter more for the poor and disadvantaged.

Jackson, Johnson, Persico (2016) — money *does* matter with a proper empirical strategy using court-mandated school finance reforms.

Background

The U.S. has a history of relying on local taxes to finance schools. This leads to a decentralized system and Tiebout sorting — localities decide how much to tax and spend on schools. In reality, this ties schooling choices with housing choices. One can interpret this as "willingness to pay" or "ability to pay" — low income students are segregated in school districts with low taxes and low spending on schools. This inequality was challenged in the courts.

Empirical Strategy

- Convert these state level changes to district level changes in per pupil spending.
- Individual level outcomes are measured by the Panel Study of Income Dynamics (PSID).
- Assume individuals are educated in their birth states.
- "Unexposed" cohorts are those that were 17 when the court-ordered SFR was passed.
- "Exposed" cohorts are those who were younger than 17 when the court-ordered SFR was passed.

- DiD is based on comparing exposed and unexposed cohorts in treated vs. untreated districts.
- There are dosage effects both for cohorts (some are treated for multiple years) and across districts within states (some districts have large positive change, some zero, some may have negative).

Main estimating equation

$$\ln\left(\overline{PPE_{5-17}}\right)_{idb} = \pi_1\,(EXP_{idb} \times Dosage_d) + \pi_2\,EXP_{idb} + \Pi\,C_{idb} + \rho_d + \rho_b + \xi_{idb} \tag{65}$$

$$Y_{idb} = \delta\,\ln\left(\overline{PPE_{5-17}}\right)_{idb} + \Phi\,C_{idb} + \phi_d + \phi_b + \varepsilon_{idb} \tag{66}$$

where $\left(\overline{PPE_{5-17}}\right)_{idb}$ is the average per pupil school spending during schooling age in the district of birth; EXP_{idb} is exposure measured as the number of school-age years occurring after the passage of a state court-ordered SFR; $Dosage_d$ is the district level measure of the amount of spending change caused by the court-ordered SFR.

Tables III, IV, V show the 2SLS estimates

- Table III — completed schooling: a 10% increase in district per pupil spending for each of the 12 years of exposure increases years of education by 0.315 (more for low income: 0.459).
- Table IV — ln(wage) at age 20-45: a 10% increase in district per pupil spending for each of the 12 years of exposure increases wage by 7.7% (more for low income: 9.6%).
- Table V — poverty status at age 20-45: a 10% increase in district per pupil spending for each of the 12 years of exposure reduces the annual incidence of adult poverty by 2.7 percentage points (more for low income: 6.1 percentage points).

13 Does School Spending Matter? New Literature

This section summarizes two strands of recent research examining the effects of school spending on educational and economic outcomes. The first part focuses on the paper by Jackson, Johnson, and Persico (2016), which exploits court-mandated school finance reforms in the United States. The second part reviews a meta-analysis by Jackson and Mackevicius (2024) that synthesizes findings from multiple studies.

13.1 Key Papers

- Jackson, Johnson, Persico (2016) "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms"
- Jackson and Mackevicius (2024) "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States"
  - A new meta-analysis

13.2 Jackson, Johnson, Persico (2016)

The central claim of Jackson, Johnson, and Persico (2016) is that money *does* matter when accompanied by a proper empirical strategy that exploits exogenous variation from court-mandated school finance reforms.

Background The U.S. historically relies on local taxes to finance schools. This reliance leads to a decentralized system and Tiebout sorting, whereby localities determine tax rates and school spending. In practice, schooling choices are closely linked with housing choices. This linkage can be interpreted as reflecting either "willingness to pay" or "ability to pay", resulting in low-income students often being concentrated in districts with lower taxes and lower school spending. These inequalities eventually prompted legal challenges.

Empirical Strategy State-level spending changes are translated into district-level changes in per pupil spending. Individual outcomes are measured using the Panel Study of Income Dynamics (PSID). It is assumed that individuals receive their education in their state of birth. Cohorts are classified as "unexposed" if they were 17 at the time of the court-ordered SFR passage, and as "exposed" if they were younger than 17. A Difference-in-Differences (DiD) strategy compares exposed and unexposed cohorts across treated and untreated districts. The analysis accounts for dosage effects both by cohort (with some cohorts being treated for multiple years) and across districts (as spending changes vary across districts).

13.2.1 Main Estimating Equations

$$\ln(\overline{PPE}_{5-17})_{idb} = \pi_1 \left(EXP_{idb} \times Dosage_d\right) + \pi_2 \left(EXP_{idb}\right) + \Pi C_{idb} + \rho_d + \rho_b + \xi_{idb}, \tag{67}$$

$$Y_{idb} = \delta \ln(\overline{PPE}_{5-17})_{idb} + \Phi\, C_{idb} + \phi_d + \phi_b + \varepsilon_{idb}. \tag{68}$$

Where:

- $(\overline{PPE}_{5-17})_{idb}$ denotes the average per pupil spending during school-age years in the district of birth.
- EXP_{idb} represents the exposure, measured as the number of school-age years occurring after the passage of the state court-ordered SFR.
- $Dosage_d$ is the district-level measure of the spending change induced by the court-ordered SFR.

13.2.2 Results

Table III – Completed Schooling A 10% increase in district per pupil spending over 12 years of exposure is associated with an increase of 0.315 years of education overall, and 0.459 years for low-income students.

Table IV – Log(Wage) at Age 20-45 A 10% increase in district per pupil spending over 12 years of exposure increases wages by 7.7%, with an even larger effect (9.6%) for low-income students.

Table V – Poverty Status at Age 20-45 A 10% increase in district per pupil spending over 12 years of exposure reduces the annual incidence of adult poverty by 2.7 percentage points overall, and by 6.1 percentage points for low-income individuals.

13.3 Jackson and Mackevicius (2024)

This study presents a meta-analysis of recent research (from 2009 to 2022) evaluating the impact of school spending on student test scores and educational attainment outcomes (e.g., high school dropout rates, high school graduation rates, and college enrollment).

Study Inclusion Criteria Studies must utilize plausible exogenous variation. For policy-induced variation, studies require a minimum acceptable first-stage effect on spending. Relevant sources of variation include school finance reforms, narrowly passed referendums, Title I spending adjustments based on low-income enrollment, and funding discontinuities using regression discontinuity designs. Effect estimates are standardized by computing the impact of a $1000 per pupil spending increase over 4 years. Outcome measures are reported in standardized units. The meta-analysis encompasses 32 studies, each with an estimated effect $\hat{\theta}_j$ and its corresponding standard error.

13.3.1 Meta-Analysis Method

Each study's estimated effect, $\hat{\theta}_j$, is modeled as:

$$\hat{\theta}_j \sim N(\theta_j, \sigma_j^2), \tag{69}$$

where θ_j is the true effect in study j. True heterogeneity across studies is captured by modeling:

$$\theta_j \sim N(\Theta, \tau^2), \tag{70}$$

with Θ representing the grand mean and τ^2 the variance of true effects. Combining these yields:

$$\hat{\theta}_j \sim N(\Theta, \sigma_j^2 + \tau^2). \tag{71}$$

If the marginal effects are independent of their precision, the grand mean estimate can be obtained as:

$$\hat{\Theta} = \frac{\sum_j \hat{\theta}_j \, \omega_j}{\sum_j \omega_j}, \tag{72}$$

with weights defined by:

$$\omega_j = \frac{1}{\sigma_j^2 + \tau^2}. \tag{73}$$

Additional Notes: The squared standard error se_j^2 is used as an estimate for σ_j^2. The between-study variance $\hat{\tau}^2$ is estimated; overlapping confidence intervals across studies suggest a small variance. The overall approach involves averaging the study-specific effects while down-weighting those with lower precision. Final estimates include $\hat{\tau}^2$, $\hat{\Theta}$, and the standard error of $\hat{\Theta}$.

13.3.2 Context and Comparison

A $1000 increase in per pupil spending over 4 years is estimated to raise test scores by 0.0316 standard deviations. For context, the Project STAR experiment, which reduced class size by approximately 7 students, increased test scores by 0.12 standard deviations.

Moreover, the same spending increase is estimated to boost high school graduation rates by approximately 2.05 percentage points (calculated as 0.357×0.0573) and college enrollment by 2.81 percentage points (calculated as 0.49×0.0573). In comparison, Project STAR reduced class size by around 7 students and increased college going rates by 2.7 percentage points by age 30.

Overall, the evidence suggests that school spending has larger effects on later life outcomes, possibly by influencing non-cognitive skills.

14 Class Size

When studying education production functions, we face significant challenges when using cross-sectional observational data. The standard education production function can be expressed as:

$$Y_{ij} = \alpha S_{ij} + \beta F_{ij} + \varepsilon_{ij} \tag{74}$$

where Y_{ij} represents the achievement of student i in school j, S_{ij} denotes school characteristics, and F_{ij} captures family background variables. An important consideration is that the entire history of school characteristics (S_{ij}) and family background (F_{ij}) can contribute to student achievement (Y_{ij}) in any given year. Randomization is crucial because it breaks this link between school and family factors.

14.1 Tennessee STAR Experiment (Krueger, 1999)

The Tennessee STAR (Student-Teacher Achievement Ratio) experiment represents an ideal randomized controlled trial in the education literature. This large-scale experiment cost approximately $12 million and was implemented for a kindergarten cohort in 1985-86. The study continued for four years and involved 11,600 children. To be eligible for participation, schools needed to be public and large enough to accommodate one class of each experimental type. The critical design feature was that randomization occurred within schools.

Students were randomly assigned to one of three class types: small classes (13-17 students), regular classes (22-25 students), or regular classes with a teacher aide. Teachers were also randomly assigned to different class sizes. Students who began in kindergarten remained in their assigned class size condition through third grade. Students who entered in subsequent grades (1st, 2nd, or 3rd) were also randomly assigned to treatment conditions upon entry.

While original "initial assignment" data is not available, researchers have access to a closely related variable—the class in which students were initially enrolled before the start of kindergarten—for a subset of participants. This variable was used as an instrumental variable in some specifications.

14.2 Methodological Challenges

The experiment faced several important challenges. First, re-randomization occurred for students not assigned to smaller classes, who were randomized to either have a teacher's aide or not. Meanwhile, students in small classes remained with the same classmates throughout the study. This presents an analytical problem if the constancy of classmates influences outcomes.

Second, attrition was substantial—approximately half of the students who began in kindergarten were missing from the sample in at least one of the subsequent four years. Children in small classes were 3-4 percentage points more likely to remain in the sample than those in regular classes. Without baseline test scores, researchers validated the randomization by checking whether other characteristics differed across treatment groups (shown in Tables I and II of the original paper).

14.3 Results of the STAR Experiment

14.3.1 Main Effects of Class Size

The experiment created a substantial contrast in class sizes—15 students in small classes versus 22 in regular classes on average. This reduction led to a 5-6 percentile point increase in test scores, representing approximately 20% of a standard deviation and 60-80% of the black-white achievement gap.

The statistical model included a class-specific component that could reflect either teacher effects or peer effects:

$$Y_{kcs} = \beta_0 + \beta_1 SMALL_{cs} + \beta_2 REG/A_{cs} + \beta_3 X_{kcs} + \alpha_s + \varepsilon_{kcs} \tag{75}$$

where $\varepsilon_{kcs} = \mu_{cs} + \varepsilon'_{kcs}$ and μ_{cs} represents a class-specific random component common to all students in the same class.

14.3.2 Addressing Attrition

Non-random attrition could potentially bias the estimates upward if, for example, stronger students assigned to regular classes were more likely to leave the study. To address this concern, researchers imputed percentile scores for students who left the study using their test scores from the latest available year. The estimated class size effect remained similar after this adjustment, suggesting that attrition did not substantially bias the results.

14.3.3 Initial and Cumulative Effects

Analyzing pooled data across grades revealed that initial year test scores increased by approximately 4 percentiles, with scores rising by about 1 percentile point per year thereafter. This pattern raises an interesting question: given that students remained in smaller classes through third grade, why was the largest effect observed in the first year? One explanation is a "one-time school socialization" effect that elevates the overall level of student achievement without changing the rate of learning. It is worth noting that a value-added specification would miss this important initial effect.

14.3.4 Heterogeneous Treatment Effects

The benefits of smaller classes appeared larger for certain subgroups, including boys, Black students, and students from lower-income backgrounds (those participating in free lunch programs). This suggests that class size reduction may be particularly effective for traditionally disadvantaged student populations.

14.3.5 Potential Hawthorne Effects

Researchers considered whether teachers in small classes might have responded differently because they knew they were part of an experiment, or whether the measured effects might be artificially small because teachers in regular classes exerted extra effort. This highlights the general principle that interventions can change incentives in unobserved ways. To investigate this possibility, Krueger examined the effect of class size variations within the regular class group (where random variation in class size existed) and found similar effect sizes, suggesting that Hawthorne effects were not driving the results.

14.4 Long-Term Effects: Project STAR 2.0 (Chetty et al., 2011)

Chetty et al. (2011) extended the original Project STAR analysis by examining long-term outcomes into adulthood. While previous research had established that small class sizes increased test scores by approximately 5 percentile points (0.2 standard deviations), these effects appeared to fade by eighth grade—similar to findings from Head Start research (Currie and Thomas, 1995).

The key question was whether early test score gains translated into improved adult outcomes. This study represented the first major attempt to link student outcomes in school directly to earnings in adulthood. The researchers utilized U.S. tax records to track outcomes for the STAR participants, most of whom were born in 1979-1980 and were approximately 27 years old in 2006.

14.4.1 Cross-Sectional Relationship Between Test Scores and Adult Outcomes

Analysis revealed that a 1 percentile point increase in test scores was associated with a $132 increase in earnings, although the R^2 was relatively low. For comparison, a 1 percentile point increase in parental income was associated with a $146 increase in earnings. Test scores were also correlated with college quality, marriage by age 27, and residence in zip codes with higher college graduation rates.

14.4.2 Effects of Class Size on Adult Outcomes

Small classes increased the probability of college attendance. However, conditional on attending college, the average quality of college attended actually decreased, likely because marginal students induced to attend college enrolled in less selective institutions. The researchers found no detectable effect on earnings at age 27, though this may have been too early in participants' career trajectories to observe earnings effects.

Teacher experience showed significant positive effects on both test scores and earnings, but only for students who entered the study in kindergarten. The researchers found limited evidence for observable peer effects.

14.5 Effects of Unobservable Classroom Characteristics

Chetty et al. investigated whether overall "class quality" influenced earnings and adult outcomes. Class quality could reflect peer effects, teacher effects, or common class-level factors such as classroom environment.

The conceptual model was specified as:

$$s_{icn} = d_n + z_{cn} + \alpha_{icn}$$

where s_{icn} represents the end-of-year test score of student i in classroom c in school n, d_n is a school fixed effect, z_{cn} is a classroom fixed effect, and α_{icn} captures student intrinsic ability.

For earnings, the model was:

$$y_{icn} = \delta_n + \beta z_{cn} + z^Y_{cn} + \rho\alpha_{icn} + \nu_{icn}$$

where z^Y_{cn} represents the direct effect of classroom quality on earnings, and β captures the effect of classroom quality that operates through test scores.

The researchers proxied z_{cn} using the end-of-year mean test scores of peers in the classroom (deviated from the school mean). With random assignment of students to classrooms, the class mean should equal the school mean in the absence of classroom effects.

Using a leave-one-out estimator:

$$\Delta s^{-i}_{cn} = s^{-i}_{cn} - s^{-i}_{n}$$

They estimated:

$$y_{icn} = \alpha_n + \beta_{LM}\Delta s^{-i}_{cn} + \varepsilon_{icn}$$

Results indicated that class quality significantly influenced earnings and other adult outcomes. Figure IV in the original paper illustrates these relationships graphically, showing strong positive correlations between classroom quality and both kindergarten test scores and adult earnings, despite a much weaker relationship with eighth-grade test scores.

14.6 Explaining Fade-Out and Long-Term Effects

An intriguing pattern emerged from the data: effects on test scores faded out by eighth grade, but effects on adult outcomes reappeared later in life. The researchers investigated whether non-cognitive skills might explain this pattern. Table IX shows that classroom quality in kindergarten affected both cognitive and non-cognitive skills (such as effort, initiative, engagement, and valuing school). While cognitive advantages faded, non-cognitive advantages persisted and may have contributed to the long-term effects on adult outcomes.

14.7 Conclusion and Policy Implications

The Tennessee STAR experiment and its long-term follow-up studies provide compelling evidence on the effects of class size and classroom quality on both academic achievement and adult outcomes. The fade-out of test score effects coupled with the reemergence of effects on adult outcomes suggests complex dynamics in how early educational interventions affect life trajectories—potentially operating through both cognitive and non-cognitive channels.

These findings have significant implications for education policy, suggesting that investments in early education quality (through smaller class sizes or other measures that enhance classroom environments) may yield substantial long-term benefits even when intermediate academic measures show diminishing returns. The research also highlights the importance of measuring both cognitive and non-cognitive outcomes when evaluating educational interventions.

15 Peer Effects

15.1 Introduction to Estimating Peer Effects

The study of peer effects in education represents a significant challenge in empirical economics. Several key papers have established methodological frameworks for addressing these challenges, including Hoxby (2000), Sacerdote (2001), and Carrell et al. (2013). This section explores the empirical strategies, methodological issues, and findings from these seminal works.

15.2 Challenges in Estimating Peer Effects

There are two fundamental problems when estimating peer effects. The first is the **selection problem**, which arises because parents and families of high-achieving children often choose environments with high-achieving peers, creating an endogeneity issue that confounds causal identification. The second problem concerns the **baseline model specification**. The standard approach assumes that peer effects operate linearly through the mean of peers' outcomes. This baseline model implies several distributional consequences: for instance, removing one good student from one classroom creates a negative effect that is exactly offset by a positive effect in another classroom; it rules out any efficiency gains or losses, thereby effectively excluding "one-bad apple" or "one shining light" effects; and it precludes the evaluation of the pros and cons of tracking students by ability.

15.3 Hoxby (2000): Peer Effects in the Classroom

15.3.1 Empirical Strategy

Hoxby's approach parallels the classroom size literature by exploiting adjacent cohort variation within a grade level within a school. This strategy requires extensive data, as demonstrated by the use of the entire population of public school students in Texas in grades 3, 4, 5, and 6. The empirical model is specified as:

$$A_{male,gjc} = \eta_{male,gjc} + \beta\, p_{female,jgc} + \varepsilon_{male,gjc}, \tag{76}$$

$$A_{female,gjc} = \eta_{female,gjc} + \beta\, p_{male,jgc} + \varepsilon_{female,gjc}. \tag{77}$$

The identification assumption is that the average achievement of males or females at grade level g in school j is stable, and that changes in adjacent cohorts' achievement should not be systematically related to idiosyncratic changes in the percentage of females or males. To implement this strategy, Hoxby estimates the model in first differences:

$$\Delta A_{male,gjc} = \beta\, \Delta p_{female,jgc} + \Delta\varepsilon_{male,gjc}, \tag{78}$$

$$\Delta A_{female,gjc} = \beta\, \Delta p_{male,jgc} + \Delta\varepsilon_{female,gjc}. \tag{79}$$

This approach works particularly well for gender, as there is no reason to suspect trends in the percentage of female students; however, for racial groups, additional controls for trends become necessary.

15.3.2 Results

The empirical findings, as illustrated in Table 4, demonstrate substantial cohort-to-cohort variation that supports the identification strategy. Further results presented in Table 5 indicate that a 10 percentage point increase in the female share leads to a 0.03 to 0.04 point increase in achievement scores. Moreover, being surrounded by peers who score 1 point higher on average is associated with a 0.3 to 0.5 point increase in one's own score. It is noteworthy that although female students score only slightly higher than males in math, if the effect operated solely through peer achievement, the estimated effect would be implausibly large. This observation suggests the presence of alternative channels for peer effects, such as reduced classroom disruption, teacher effects, or changes in academic standards.

15.4 Sacerdote (2001): Peer Effects with Random Assignment

15.4.1 Methodological Challenges

Sacerdote addresses three key challenges in estimating peer effects. First, the selection problem is mitigated through random assignment. Second, the reflection problem (as discussed in Manski, 1993) poses difficulties because peers may simultaneously affect one another, complicating the identification of peer effects without clear information about the reference groups. Third, there is the challenge of distinguishing between the effects of peer background variables (such as SAT scores or parental background) and peer behavior (such as GPA).

15.4.2 Empirical Design

Sacerdote exploits the random assignment of roommates at Dartmouth College while conditioning on housing preferences (including gender and responses to four specific questions). The validity of this randomization is verified in Table II, where a regression of student i's SAT score on roommate j's SAT score yields statistically insignificant coefficients.

15.4.3 Model Specification

Sacerdote estimates equations of the following form:

$$GPA_i = \delta + \alpha\, ACA_i + \beta\, ACA_j + \gamma\, GPA_j + \varepsilon_i, \tag{80}$$

$$GPA_j = \delta + \alpha\, ACA_j + \beta\, ACA_i + \gamma\, GPA_i + \varepsilon_j. \tag{81}$$

By substituting the second equation into the first, the model becomes:

$$GPA_i = \left(\frac{1}{1-\gamma^2}\right)\left[\delta(1+\gamma) + (\alpha+\gamma\beta)ACA_i + (\beta+\gamma\alpha)ACA_j + \gamma\varepsilon_j + \varepsilon_i\right]. \tag{82}$$

When estimating OLS regressions of GPA_i on ACA_i and ACA_j, the specification takes the form:

$$GPA_i = \pi_0 + \pi_1 ACA_i + \pi_2 ACA_j + \eta. \tag{83}$$

Here, the coefficients π_1 and π_2 represent reduced form relationships that combine both exogenous parameters (α, β) and endogenous effects (γ). Sacerdote also regresses student i's GPA directly on student j's GPA, though this approach is subject to the reflection problem and thus cannot be interpreted causally. Some specifications include dormitory fixed effects (denoted by θ_k) and introduce non-linearity by categorizing academic ability into the bottom 25%, middle 50%, and top 25% groups.

15.4.4 Results

The findings presented in Table III indicate significant peer effects on academic outcomes. The coefficient of 0.12 suggests that a one standard deviation (0.43) increase in a roommate's GPA is associated with a 0.05 increase in the student's own GPA. When dormitory fixed effects are included, the coefficient decreases slightly but remains statistically significant, indicating that the effect is not driven solely by common shocks. Interestingly, the freshman GPA peer effect is not observed by senior year, which suggests that these influences are transitory. Further, columns (4) through (6) demonstrate that having a roommate in the "top 25%" academic category significantly improves performance. An examination of the effect on the choice of major, as shown in Table IV, finds no significant impact since the distribution of roommate pairs with the same major is consistent with expectations under independence. Table V explores social outcomes and reveals that if a roommate joins a fraternity, the student is 8% more likely to join one as well; however, this effect disappears when dormitory fixed effects are included.

15.5 Carrell et al. (2009): Squadron-Level Peer Effects

15.5.1 Empirical Design

Carrell et al. study peer effects within "squadrons" at the Air Force Academy, where the freshman class is randomly assigned to 36 squadrons of 120 students each. This setting may provide a better measure of the relevant peer group since the freshman squadrons participate together in almost all activities. The empirical model is specified as:

$$G_{izt} = \phi_0 + \phi_1 X_{izt} + \phi_2 \frac{\sum_{i \neq l} X_{lzt}}{n_{zt} - 1} + \beta X_{izt} + \gamma_{ct} + \varepsilon_{izt}. \tag{84}$$

15.5.2 Results

Unlike the study by Sacerdote, Carrell et al. find that peers' verbal SAT scores exert a strong positive effect, whereas math SAT scores do not show significant effects. The magnitude of the effect is larger than in previous studies; for example, a one standard deviation increase in peer verbal SAT (approximately 11 points) is associated with a 0.05 increase in GPA (roughly 1/12 of a standard deviation), as indicated in Table 3. Moreover, the roommate effect is not statistically significant in this study, suggesting that the broader peer group may be more influential. When analyzed by course type, peer effects are most pronounced in math and science courses and are insignificant in foreign language and physical education, as reported in Table 4. Table 6 further suggests non-linearity in peer effects, with students in the bottom third of the ability distribution benefiting most from having peers with high verbal SAT scores.

15.6 Carrell et al. (2013): Optimal Policy Experiment

Building on their 2009 findings of non-linear peer effects, Carrell et al. (2013) conducted an experiment in which half of the incoming class served as controls, while the other half was randomized into squadrons designed to maximize peer effects for the lowest-third ability students. The optimal sorting resulted in two types of classes: one that was "bimodal," containing both high and low ability students, and another that was "homogeneous," comprising predominantly middle-ability students. Surprisingly, Table 6 shows that this "optimal" sorting had a negative impact on low-ability students while improving outcomes for middle-ability students. This unexpected result suggests that mere exposure to high-ability peers is different from endogenous friendship formation, thereby highlighting the complexity of peer group dynamics.

15.7 Further Perspectives on Peer Effects

The literature distinguishes between endogenous peer effects, which are influenced by peer behaviors, and exogenous or predetermined peer effects, which are influenced by peer characteristics. A typical model in this context is expressed as:

$$GPA_i = \pi_0 + \pi_1 ACA_i + \pi_2 ACA_j + \gamma\varepsilon_j + \varepsilon_i. \tag{85}$$

While Carrell et al. (2013) experimentally manipulated exogenous peer characteristics, experiments could also be designed to study endogenous effects. For example, researchers might experimentally shift peer behaviors (such as studying habits) and then observe the impact on the focal student's outcomes. Notable studies adopting this approach include those by Moffitt (2001) and Dahl et al. (2014).

16 Teacher Value-Added

16.1 Foundational Research

The literature on teacher value-added (VA) examines the impact of individual teachers on student achievement and long-term outcomes. Foundational studies in this area include Hanushek, Kain, O'Brien, and Rivkin (2005) with their work "The Market for Teacher Quality", Kane and Staiger (2008) in "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation", and the two seminal papers by Chetty et al. (2014) titled "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates" and "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood". These contributions establish the methodological and empirical framework for understanding how teachers influence student progress.

16.2 Value-Added Models

Value-added models are designed to isolate the contribution of teachers to student learning. A basic model as proposed by Hanushek et al. (2005) is represented by the equation

$$\Delta A_{isg} = A_{isg} - A_{isg-1} = f(X_{ig}, S_{ig}, \gamma_i, \varepsilon_{isg}), \tag{86}$$

where ΔA_{isg} denotes the gain in achievement for student i in school s in grade g. In this framework, X captures non-school factors such as family background, peers, and neighborhood influences, while S represents school and teacher factors. The term γ_i is an individual fixed effect, and ε_{isg} is the error term. A critical assumption underlying this model is that prior achievement A_{isg-1} is a sufficient statistic summarizing the entire history of past inputs.

An alternative specification employed by Hanushek et al. incorporates teacher effects more explicitly:

$$\Delta A_{isg} = f'(X_{ig}, \tilde{S}_{ig}) + \sum_j t_j T_{ijg} + (\gamma_i, \varepsilon_{sg}), \tag{87}$$

where T_{ijg} is an indicator variable denoting whether student i had teacher j in grade g, and $\tilde{S}_{ig}$ represents non-teacher school inputs. The parameter t_j reflects teacher fixed effects, allowing for an assessment of individual teacher contributions beyond other institutional factors.

16.3 Addressing Endogeneity in Teacher Selection

A major challenge in value-added models is the endogenous assignment of teachers to students. Endogeneity arises both between schools, where parents may actively seek out high-quality teachers, and within schools, as administrators make decisions regarding teacher-classroom assignments. While these selection effects can be partly controlled for by incorporating student fixed effects or accounting for initial achievement levels, complications emerge when placements depend on changes in student achievement—a phenomenon described by Rothstein (2007) as "dynamic selection". Hanushek et al. suggest that controlling for school fixed effects, by focusing on variations within individual schools, provides a potential "lower-bound" estimate of teacher quality.

16.4 Test Measurement Issues

Hanushek et al. (2005) identify several critical issues related to test measurement that may affect teacher value-added estimates. One prominent concern is standardization. For example, tests like TAAS encounter ceiling effects, whereby high-achieving students are unable to exhibit large gains, potentially biasing the estimation of teacher quality. To mitigate this, the authors propose the use of standardized gain scores defined as:

$$G_{isg} = [(A_{isg} - A_{isg-1}) - \mu_g^{cm}]/\sigma_g^{cm}. \tag{88}$$

This methodology involves partitioning the initial test score distribution into ten intervals and then computing the mean and standard deviation of gains for all district students beginning in each interval. The resulting gains are normalized to have a mean of zero and a standard deviation of one within each score interval for each year, thus allowing for evaluations of teacher performance that appropriately control for initial test scores.

Another important consideration is the adjustment for measurement error. Teacher effect estimates are subject to sampling error due to small class sizes and other factors. Hanushek et al. address this by modeling the observed teacher effect as

$$\hat{t}_j = t_j + v_j, \tag{89}$$

and using year-to-year correlations to extract the persistent component of teacher effects. Specifically, if r_{12} denotes the correlation of teacher effects across two years, then the true variance in teacher effects is given by

$$\text{Var}(t) = r_{12} \cdot \text{Var}(\hat{t}). \tag{90}$$

This approach effectively "shrinks" the estimates for teachers who experience significant class-level shocks, drawing them back toward the mean.

16.5 Key Findings from Hanushek et al. (2005)

Hanushek et al. (2005) report two critical findings in their analysis. First, there is substantial variation in teacher quality. The results, as illustrated in Table 1 of their study, indicate that even the most conservative estimate (0.047) implies that a one standard deviation increase in teacher quality corresponds to an increase in student test scores of 0.22 standard deviations. Second, observable teacher characteristics, including experience, certification exam scores, and educational background, appear to have minimal effects on teacher quality. The sole statistically significant finding regarding observable characteristics is that first-year teachers perform significantly worse compared to their more experienced counterparts.

16.6 Kane and Staiger (2008): Experimental Validation

Kane and Staiger (2008) tackle the essential question of how the average test scores of a class would differ if one teacher were replaced by another. Their study compares non-experimental value-added estimates with those derived from an experimental design based on random assignment. Conducted in LAUSD over two school years (2003–2004 and 2004–2005), their experiment involved teachers applying for National Board professional teaching certification. Each teacher was paired with a comparison teacher from the same school and grade level, with a minimum of three years of experience. Class rosters were created by principals for these teacher pairs, and teachers were randomly assigned to the rosters.

Their empirical model is formulated as:

$$A_{ijt} = X_{ijt}\beta + v_{ijt}, \quad \text{where} \quad v_{ijt} = \mu_j + \theta_{jt} + \varepsilon_{ijt}. \tag{91}$$

Here, A_{ijt} represents the test score level or gain for a student, while X_{ijt} includes student and classroom-level covariates. The term μ_j captures the teacher effect, θ_{jt} represents non-persistent classroom-by-year shocks, and ε_{ijt} is the student-by-year error component. Kane and Staiger employ empirical Bayes methods to generate estimates of teacher effects by estimating the relevant variance components and adjusting for reliability. Their analysis indicates that the most robust specifications for predicting experimental differences rely on student test score levels (as opposed to gains), with controls for prior achievement, student demographics, and peer characteristics. The study validates the use of value-added measures by demonstrating that non-experimental VA estimates closely predict experimental differences in teacher effectiveness, with coefficients near unity for the best models. Additionally, they observe that teacher effects tend to diminish by approximately 50% per year.

16.7 Chetty et al. (2014): Comprehensive Analysis

The Chetty et al. papers address two critical questions: (1) whether teacher VA measures are biased due to student sorting, and (2) whether teachers who raise test scores also improve students' long-term outcomes.

16.7.1 Data and Approach

Chetty et al. utilize extensive administrative data from grades 3–8 in a large urban school district spanning 1989–2009. This dataset, which includes teacher and class assignments for 2.5 million children, is merged with IRS tax returns from 1996 to 2010. The combined data enable the tracking of long-term outcomes such as earnings, college attendance, teenage birth rates, and neighborhood quality.

16.7.2 Testing for Bias in VA Estimates

To assess potential bias in VA estimates, Chetty et al. implement two tests. First, they investigate whether observable characteristics that are excluded from the VA model correlate with the VA estimates. Although they observe that students with higher-income parents tend to receive high-VA teachers, this does not bias the estimates because the models control for students' prior-year scores, 85% of the variation in teacher VA occurs within schools rather than across schools, and parents typically exert limited influence over teacher assignments within schools. Second, they employ a quasi-experimental approach based on teacher switching. Their event study analysis reveals that when a high-VA teacher enters or exits a school, the average test scores in the affected grade change in accordance with the predicted pattern, with coefficients very close to 1. This finding confirms that VA estimates are largely unbiased.

16.7.3 Impacts on Long-Term Outcomes

In the second part of their study, Chetty et al. (2014, Part II) examine whether improvements in test scores attributable to higher teacher VA translate into better long-term life outcomes. They adopt two methodological approaches. The first is a cross-sectional reduced form specified as:

$$Y_{it} = a + k_g m_{jt} + n_{it}, \tag{92}$$

where $m_{jt} = \mu_{jt}/\sigma_\mu$ represents normalized teacher VA. The second approach is a teacher switcher quasi-experimental design given by:

$$\Delta Y_{sgt} = a + k\Delta Q_{sgt} + \Delta n_{sgt}, \tag{93}$$

with ΔQ_{sgt} denoting the change in average teacher quality within a school-grade-year cell. Both approaches yield consistent results, demonstrating that a one standard deviation increase in teacher VA is associated with significant positive effects on several long-term outcomes. These include an increase in college attendance rates by 0.82 percentage points, higher enrollment in more selective institutions, an increase in earnings at age 28 of approximately $350, as well as reductions in teenage birth rates, improvements in neighborhood quality, and increases in retirement savings.

16.8 Policy Implications

Chetty et al. (2014) consider two potential policy interventions based on their findings. One proposal involves the deselection of low VA teachers. Their analysis indicates that replacing teachers in the bottom 5% of the VA distribution with those of median quality would yield substantial benefits, including a net present value gain of $407,000 for an average-sized class. The other policy measure focuses on the retention of high VA teachers. Retaining a teacher at the 95th percentile of VA for an additional year is estimated to generate present value earnings gains of $266,000. However, the authors caution that bonus payments aimed at increasing retention may not be cost-effective, as such payments would extend to many teachers who would have remained in their positions regardless.

16.9 Additional Policy Considerations

Rothstein (2015) underscores several additional factors that are critical for the design of teacher quality policies. First, if a policy involving the dismissal of low VA teachers were implemented, it is likely that teacher salaries would need to be raised to compensate for the increased job insecurity. Second, incentive pay could influence teacher effectiveness in several ways: it might motivate existing teachers to exert greater effort (acknowledging that VA is not a fixed characteristic), and it could affect the sorting of individuals into the teaching profession, as described by the Roy model. Third, performance-based, flexible pay is fundamentally different from policies that simply raise average salaries. Finally, the extent to which VA reflects teacher effort as opposed to innate ability has important implications; if VA primarily captures innate ability, then the role of sorting becomes even more significant in shaping policy design.

17 School Choice

17.1 Issues in School Choice

The idea of school choice dates back to Friedman, who proposed that introducing competition into the school system would improve productivity. The most "free market" approach would be to issue vouchers, allowing students to shop for schools that are the best match. As already discussed, there is substantial "choice" in the current public school system. This includes Tiebout choice, where local schools are financed by local school taxes, allowing parents to exercise choice through residential selection. However, state mandates and court cases that centralize spending may be undoing this channel.

Another form of choice is the introduction of "charter" schools, where states issue "charters" that allow parents, non-profits, and teachers to receive public funding to run schools. Other types of school choice within the traditional public school system include "magnet" school programs.

More on "charter schools": States, counties, and cities can issue a "charter" for which non-profits, parents, and teachers apply to open a publicly funded school. Funding typically comes from "sending" districts. These schools are managed by boards or privately and are not subject to the restrictions of public teachers' unions. Some well-known "brands" in the charter school sector include KIPP and YES.

There are several key issues regarding school choice:

1. What is the impact of school choice on those who exercise it? For instance, what is the impact of winning a lottery to attend an over-subscribed charter school?
2. What is the formula for success among those that are successful?
3. What is the overall impact of school choice, including on those who do not exercise choice?

17.2 Angrist et al. (2010)

Angrist et al. (2010), "Inputs and Impacts in Charter Schools: KIPP Lynn," addresses the issue that comparing outcomes for students in charter and traditional public schools leads to problems with selection on unobservables. The authors use admission lotteries at over-subscribed schools for identification. Note that the effect is still over the selected population who applied for the lottery.

The study focuses on KIPP Lynn, a middle school serving grades 5–8 in a low-income city north of Boston, where tuition is paid by sending districts. A comparison of 5th graders at KIPP Lynn shows they are no better than other public school students (Table 1).

The study uses the following model, where Y_{igt} represents test scores of student i in grade g in year t, S_{igt} is years of exposure in the KIPP school, and d_{ij} is a dummy for being in lottery cohort j:

$$Y_{igt} = \alpha_t + \beta_g + \sum_j \delta_j d_{ij} + \gamma' X_i + \rho S_{igt} + \varepsilon_{igt} \tag{94}$$

For the first stage:

$$S_{igt} = \theta_t + k_g + \sum_j \mu_j d_{ij} + \tau' X_i + \pi Z_i + n_{igt} \tag{95}$$

where Z_i is an indicator of whether the student received a lottery offer. Table 2 shows results, where the IV estimate is the ratio of the reduced-form to first-stage estimate.

The results indicate that lottery winners score about 0.4 standard deviations higher in math and 0.12–0.15 standard deviations higher in ELA (English Language Arts). An interaction with 4th grade scores shows that KIPP has a bigger effect on weaker students.

It is important to note that not all charter schools have a positive impact on test scores. KIPP is a network of schools operating as a "franchise" charter school. Their success formula includes: "no excuses" policies, long school days and years, selective (non-union) teacher hiring, strict behavior norms, and a strong work ethic.

17.3 Abdulkadiroglu et al. (2011)

Abdulkadiroğlu et al. (2011) compare charter schools and "pilot" schools. Charter schools are not subject to collective bargaining agreements with teachers' unions, tenure, or seniority provisions, and they follow a "no excuses" policy. Pilot schools, by contrast, are still part of the Boston Public School district.

Table 4 shows:

- Charter middle school reduced-form effects: 0.25σ (ELA) and 0.42σ (Math). Since the first stage is about 1, 2SLS estimates are similar.

- Pilot school elementary effects (reduced form): 0.21σ; first stage is around 3, so 2SLS estimate is 0.07σ.
- No effect on math in pilot elementary schools.
- Pilot middle schools have a negative effect on math (including baseline scores).

17.4 School Choice and Competition

Studies on school choice and competition include:

- Imberman (2011), "The Effect of Charter Schools on Achievement and Behavior of Public School Students" — instruments for charter location to examine spillover effects on nearby regular public schools and finds a negative effect.
- Hoxby (2000), "Does Competition Among Public Schools Benefit Students and Taxpayers?" — uses "rivers and streams" to instrument for the number of districts in a metro area as a proxy for competition, finding that productivity (test scores per pupil spending) is higher. This is a controversial study.

17.5 Abdulkadiroğlu et al. (2020)

Abdulkadiroğlu et al. (2020), "Do Parents Value School Effectiveness?," examines school choice in NYC public high schools, which has around 400 schools and a centralized assignment system using the Deferred Acceptance (DA) algorithm. Students rank schools; schools have priorities and offer seats based on these, rejecting others. The process repeats until all assignments are made. The data follow the 2003–2004 high school cohort.

Estimating Preferences. Students i are grouped into covariate cells $c(x_i)$ by borough, gender, race, subsidized lunch, census tract median income, and tercile of 8th grade scores. Let D_{ij} denote the distance between student i and school j. A logit model estimates utility:

$$U_{ij} = \delta_{c(x_i)j} - \tau_{c(x_i)} D_{ij} + \eta_{ij} \tag{96}$$

Here, $\delta_{c(x_i)j}$ is the mean utility from school j for group $c(x_i)$, and $\tau_{c(x_i)}$ is the marginal cost of distance.

Estimating Value-Added. The outcome for student i in school j is:

$$\begin{aligned} Y_{ij} &= \alpha_j + X_i'\beta_j + \varepsilon_{ij} && (97) \\ &= \alpha + X_i'\beta + \varepsilon_i + (\alpha_j - \alpha) + X_i'(\beta_j - \beta) + (\varepsilon_{ij} - \varepsilon_i) && (98) \end{aligned}$$

To correct for selection:

- Observables: sex, race, subsidized lunch status, log of median census tract income, 8th grade math and reading scores.
- Unobservables: control function approach using rank order. If students who rank school j highly do better than expected everywhere, this indicates positive selection to school j.

Control functions reveal that:

- More variation is attributed to peer quality.
- Less is attributed to value-added.
- Peer quality and value-added are highly correlated.
- Value-added is positively associated with higher returns for girls.

Linking Preferences to Peer Quality and Value-Added. Authors regress predicted utility on peer quality and value-added:

$$\delta_{cj} = \kappa_c + \rho_1 Q_j + \rho_2 ATE_j + \rho_3 M_{cj} + \varepsilon_{cj} \tag{99}$$

Each factor matters independently, but in a "horse race" regression, peer quality dominates.

Takeaways. School choice and competition were expected to improve school productivity. However, if parents do not choose based on value-added, schools have incentives to improve marketing and attract higher-performing students, rather than enhance teaching. Schools may be penalized for enrolling low-performing students, even though these students could benefit most from high value-added schools.