Public Economics Lectures
Part 5: Education Policy

Raj Chetty

Stanford University

Fall 2016

Education Policy

- Education is one of the largest public goods provided by government
    - Approximately 5.5% of GDP or 1/6 of government expenditure
    - More than 90% at the state and local level
- Focus of an extensive body of research in the rapidly expanding field of economics of education
    - Excellent admin. data on inputs and outputs, sharp micro-level variation, and direct policy relevance

Education Policy

- Main Questions

1. Why is the government intervening in the market for education?
2. What is the optimal way to intervene in the market for education?
3. How can we improve the production quality of public education?

Two-Period Model of Human Capital

- Organize literature using a two-period model of human capital investment

$$u(c_1) + \delta \cdot u(c_2) \quad \text{where} \quad c_1 = Z - h \quad \text{and} \quad c_2 = w(h)$$

- Increasing and concave function $w(h)$ captures returns to human capital investment

- If individuals choose h to maximize utility, no reason for govt. intervention

    - Individuals will invest in education optimally

Motives for Government Intervention

- Five motives for government intervention

1. Fiscal externalities: wage income $w(h)$ subject to income tax τ
2. Externalities on other agents: social welfare is $W = \{u(c_1) + \delta u(c_2)\} + f(h)$
3. Divergence between parent and child prefs.: parents max $u(c_1) + \beta \cdot u(c_2)$ where $\beta < \delta$
4. Borrowing constraints: individuals cannot invest $h > Z$
5. Optimization failures: individuals misperceive $w'(h)$

Fiscal Externalities

- Government taxes wage earnings using a linear tax τ to meet a revenue requirement E
- Individual takes τ as given (large economy) and chooses h to maximize
$$u(Z-h)+\delta u((1-\tau)w(h))$$
- Individual sets h such that
$$(1-\tau)\delta u'(c_2)w'(h)=u'(c_1)$$

Fiscal Externalities and Optimal Subsidy

- In first best, social planner would set h and τ to maximize:

$$W = \{u(Z-h) + \delta u((1-\tau)w(h))\} + \lambda[\tau w(h) - E]$$

- First order conditions for h and τ imply that planner would set h s.t.

$$\delta u'(c_2)w'(h) = u'(c_1)$$

- This is not incentive-compatible when agents face a tax τ
- If govt. introduces subsidy s for education, then individuals maximize

$$u(Z-(1-s)h) + \delta u((1-\tau)w(h))$$

- Individual now sets h such that

$$(1-\tau)\delta u'(c_2)w'(h) = (1-s)u'(c_1)$$

- Optimal subsidy $s = \tau$ replicates first-best

Fiscal Externalities and Optimal Subsidy

- Optimum requires *full deductibility* of education expenses from income tax bill
  - Equivalent to Ramsey model with taxes on all goods, so no distortions
- Intuition: marginal dollar of investment in education has a positive fiscal externality on government's budget
  - Therefore optimal for the government to subsidize education and correct this externality

Fiscal Externalities and Optimal Subsidy

- More general setting: individuals endogenously choose hours of work in period 2 and leisure is untaxed
  - Here first best can no longer be achieved
- But optimal policy is still full deductibility of education expenses even though taxes distort labor supply (Bovenberg and Jacobs 2005)
  - Example of "production efficiency" (Diamond and Mirrlees 1971)

Fiscal Externalities and Optimal Subsidy

- Stantcheva (2014) generalizes this to a dynamic model with non-linear taxes
  - Optimal education subsidy depends upon various factors including wage elasticities and insurance motives
  - Simulations suggest that full deductibility still close to optimal
- Note that all these results assume that individuals fully perceive net-of-tax returns to education over their lifetimes
  - But empirical evidence discussed below suggests that even *pre-tax* returns $w(h)$ may not be correctly perceived

Externalities on Other Agents

- Several studies have quantified non-private returns to education
  - Ex: crime, voting behavior, others' wage rates
- Typical research design: changes in state compulsory schooling laws that affect cohorts differentially
- Focus here on Lochner and Moretti (2003), who study effects of schooling on imprisonment using 1960-80 Census data

Effect of Years of Schooling on Imprisonment

	IV Estimates			Control Function
	(1)	(2)	(3)	(4)
WHITES				
Second-Stage				
Years of Schooling	-0.11	-0.09	-0.14	-0.09
	(0.02)	(0.05)	(0.06)	(0.05)
First Stage				
Compulsory Attendance = 9	0.278	0.222	0.202	
	(0.026)	(0.024)	(0.024)	
Compulsory Attendance = 10	0.213	0.199	0.176	
	(0.035)	(0.034)	(0.033)	
Compulsory Attendance $\geq$ 11	0.422	0.340	0.329	
	(0.037)	(0.033)	(0.033)	
First Stage F-test (d.o.f. = 3)	52.5	38.6	36.2	

Source: Lochner and Moretti 2003

Externalities: Lochner and Moretti (2003)

- Lochner and Moretti show an extra year of schooling reduces incarceration rates significantly
  - 0.1 pct decline for white males relative to a mean of 1%
  - 0.3 pct decline for black males relative to mean of 3%
- Gap in schooling between whites and blacks accounts for more than one-fourth of difference in crime rates
- Externality from crime reduction is about 20% of private return

Divergence in Preferences

- Hard to quantify child's preferences directly
- Duflo (2000) studies divergence in preferences between adults for investment in children
- Studies impacts of pension reform in 1992 in South Africa that gave black families larger pensions
  - Women over age 60 and men over age 65 became eligible for significant pension benefits
  - Pension benefits were twice median per capita income and 25% of children live with a pension recipient
- Studies effect of pension benefit receipt on child's weight using survey data on 3,500 households in 1993

Effect of Cash Grants to Grandparents on Children

	Dependent variable: Weight for Height Z-score						
	OLS						2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Girls							
Eligible household	0.14	0.35*	0.34*				
	(0.12)	(0.17)	(0.17)				
Woman eligible				0.24*	0.61*	0.61*	1.19*
(in col. 7: woman receives pension)				(0.12)	(0.19)	(0.19)	(0.41)
Man eligible				-0.011	0.11	0.056	-0.097
(in col. 7: man receives pension)				(0.22)	(0.28)	(0.19)	(0.74)

Source: Duflo 2000

Borrowing Constraints

- U.S. govt. disbursed $47 billion in grant aid and loans in 2000
  - Does this have a significant causal effect on college attendance?
- Dynarski (2003) studies elimination of SSA program to provide aid to students with deceased or disabled SSA beneficiaries in 1982
  - Average annual payment to children attending college with deceased parent pre-1982 was $6,700
- DD estimates of impacts on college attendance using NLSY data
  - Treatment group: children with deceased father
- Estimates imply that $1000 of grant aid increases probability of attending college by 3.6pp relative to mean of 50 pct.

Effect of SSA college aid on probability of attending college

TABLE 2—OLS, EFFECT OF ELIGIBILITY FOR STUDENT BENEFITS ON PROBABILITY OF ATTENDING COLLEGE BY AGE 23

	(1) Difference-in-differences	(2) Add covariates
Deceased father × before	0.182	0.219
	(0.096)	(0.102)
Deceased father	−0.123	Y
	(0.083)	
Before	0.026	Y
	(0.021)	

Source: Dynarski 2003

Behavioral Motives

- Standard model assumes that individuals are fully aware of rate of return to education
- Jensen (2010): randomized experiment in the Dominican Republic providing information about rates of return to completing secondary school to 2000 boys in 8th grade
- Measures impacts of presenting information about mean incomes by education level
- Tests for changes perceptions 6 months later
- Tests for changes in attendance and school completion 4 years later
    - 60% attend next year and 30% complete school in control group

Effects of Information on Perceived Returns

	Panel A. Perceived returns to school				
	Round 1		Round 2		
	Control	Treatment	Control	Treatment	Difference-in-difference
Expected earnings (self):					
Primary (only)	3,548	3,484	3,583	3,230	−284***
	(116)	(124)	(118)	(92)	(43)
Secondary (only)	3,884	3,806	4,001	3,995	82*
	(132)	(145)	(132)	(114)	(44)
Implied perceived returns	336	322	418	765	366***
	(25)	(27)	(24)	(34)	(29)
Expected earnings (others):					
Primary (only)	3,509	3,447	3,546	3,204	−274***
	(112)	(120)	(113)	(92)	(41)
Secondary (only)	3,802	3,728	3,892	3,916	102**
	(126)	(143)	(120)	(111)	(45)
Implied perceived returns	293	281	346	712	377***
	(23)	(29)	(22)	(31)	(26)
Number of observations	1,003	1,022	922	977	1,859

Source: Jensen 2010

Effects of Information on Schooling

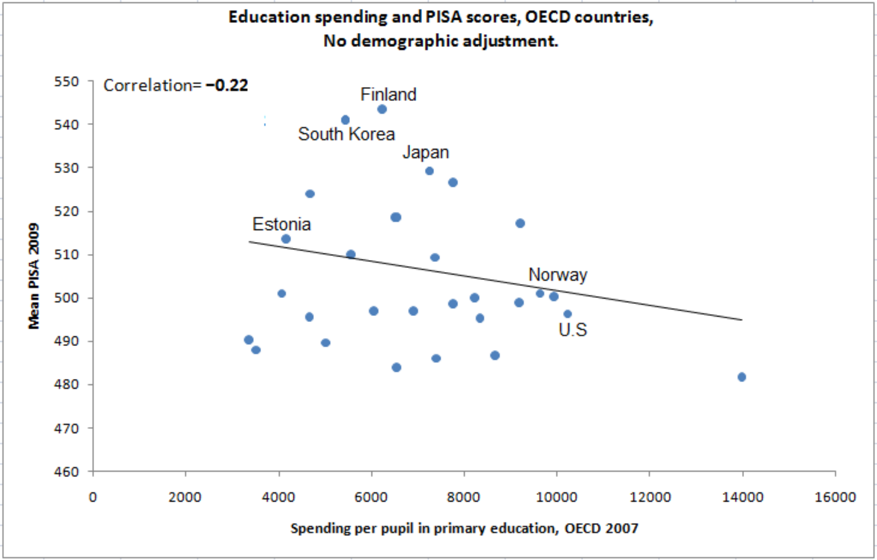
	Full sample				Poor households				Least poor households			
	(1) Returned next year	(2) Finished school	(3) Years of schooling	(4) Perceived returns	(5) Returned next year	(6) Finished school	(7) Years of schooling	(8) Perceived returns	(9) Returned next year	(10) Finished school	(11) Years of schooling	(12) Perceived returns
Treatment	0.041*	0.023	0.20**	367***	0.006	−0.01	0.037	344***	0.072*	0.054*	0.33***	386***
	(0.023)	(0.020)	(0.082)	(28)	(0.034)	(0.026)	(0.11)	(41)	(0.038)	(0.031)	(0.12)	(41)
Log	0.095**	0.23***	0.79***	29.0	0.054	0.26***	0.69***	188**	0.047	0.10	0.51	23
(inc. per capita)	(0.040)	(0.044)	(0.16)	(47)	(0.068)	(0.062)	(0.23)	(87)	(0.12)	(0.13)	(0.45)	(133)
School	0.011	0.019**	0.086**	0.74	0.001	0.015	0.064	−9.5	0.025*	0.024*	0.10**	8.2
performance	(0.010)	(0.009)	(0.034)	(14)	(0.014)	(0.012)	(0.048)	(13.5)	(0.013)	(0.012)	(0.048)	(22)
Father	0.074**	0.050*	0.26**	−24	0.056	0.019	0.16	−29.1	0.096**	0.096**	0.36**	−3.8
finished sec.	(0.030)	(0.030)	(0.12)	(32)	(0.045)	(0.043)	(0.18)	(62)	(0.038)	(0.038)	(0.14)	(40)
Age	−0.010	0.004	−0.006	−42*	−0.042	0.002	−0.071	−46	0.005	0.005	0.025	−35
	(0.016)	(0.015)	(0.059)	(21)	(0.030)	(0.019)	(0.088)	(32)	(0.025)	(0.035)	(0.087)	(29)
R^2	.016	.040	.049	.090	.007	.019	.014	.094	.020	.020	.029	.090
Observations	2,241	2,205	2,074	1,859	1,055	1,055	1,007	920	1,056	1,056	1,002	939

Source: Jensen 2010

Optimal Level of Investment in Education

- Preceding results suggests that private investment in education is likely to be below social optimum
- Optimal to implement Pigouvian corrective taxes and possibly regulations to increase level of education
  - Motivates policies such as compulsory school laws and financial aid for college
  - Potentially motivates public education system
- But reason for public *provision* less clear
  - Moreover, simply increasing spending h is not necessarily the solution

Test Scores vs. Spending Across Countries



Education Production Function

- Human capital h often modeled as a one-dimensional choice
- Education production function literature: model h as a function of inputs such as class size, teacher quality, computers, length of school day,...

$$h(z_1, ..., z_N)$$

- From this perspective, return to "education" is ill-defined
  - Important to measure and improve production efficiency
- One leading approach: rely on private market competitive forces
  - Government does not try to figure out how to best produce energy-efficient cars; just sets regulations
  - Why not do the same with schools?

Tiebout Competition

- Classic paper on competition and provision of public goods: Tiebout (1956)
- Considers a model with many small communities, each of which choose public goods provision (e.g. school quality)
- Individuals are free to choose where to live frictionlessly and with no transporation costs
- House prices are set to equate supply and demand for each community in equilibrium
- Main result: with a large number of communities, public good provision will be efficient
  - Intuition: individuals "vote with their feet" by leaving inefficient areas

Tiebout Competition

- Large subsequent literature shows that efficiency is only obtained under special conditions
  - But point that option to move across communities creates competitive pressure to improve schools is quite general
- Key practical constraint: moving is costly and it is difficult to unbundle school choice from other factors such as transport costs
- School vouchers and school choice (ability to take tax payment to any school) can be viewed as efforts to improve Tiebout competition
- Does this approach work in practice?
  - Modern literature on school choice focuses on impacts of new charter schools that compete with public schools

Effects of Charter Schools

- Several studies estimate effects of charter schools using lottery designs
  - Charter schools are often oversubscribed and have lotteries for admission
  - Compare outcomes of winners vs. losers to identify causal effects
- Abdulkadiroglu et al. (2011): compare effects of charters and pilot schools in Boston
  - Charter schools are exempt from all public school regulations
  - Pilot schools are like charters but covered by BPS union regulations
  - Financed by payments from students' home district (tax payments transferred to charter school)

Effects of Boston Charter Schools on Test Scores

			Charter Schools				
			Basic controls			2SLS w/Additional controls	
Level	Subject		First Stage (1)	Reduced Form (2)	2SLS (3)	Demographics (4)	Demographics + Baseline (5)
Elementary School	ELA		—	—	—	—	—
		N					
	Math		—	—	—	—	—
		N					
Middle School	ELA		1.000***	0.253***	0.253***	0.203***	0.198***
			(0.099)	(0.066)	(0.067)	(0.056)	(0.047)
		N		3157		3157	3101
	Math		0.967***	0.401***	0.415***	0.376***	0.359***
			(0.094)	(0.065)	(0.067)	(0.059)	(0.048)
		N		3317		3317	3258

Source: Abdulkadiroglu at al 2011

Effects of Boston Pilot Schools

Pilot Schools				
Basic controls			2SLS w/Additional controls	
First Stage (6)	Reduced Form (7)	2SLS (8)	Demographics (9)	Demographics + Baseline (10)
2.945***	0.209**	0.071**	0.062**	—
(0.189)	(0.084)	(0.028)	(0.026)	
	1141		1141	
2.950***	0.110	0.037	0.033	—
(0.190)	(0.085)	(0.029)	(0.028)	
	1139		1139	
1.526***	0.022	0.014	0.010	−0.041
(0.172)	(0.065)	(0.042)	(0.040)	(0.103)
	4314		4314	3024
1.450***	−0.065	−0.045	−0.041	−0.223**
(0.167)	(0.064)	(0.044)	(0.041)	(0.090)
	4777		4777	3348

Source: Abdulkadiroglu at al 2011

Effects of Charter Schools

- Subsequent study by Angrist et al. (2013) show that Boston charters have significant effects on college attendance rates
- General finding in literature: small positive mean effects on test scores on average, but substantial heterogeneity across schools
  - Ex: Hoxby and Murarka (2009) estimates for 42 schools in NYC
  - Dobbie and Fryer (2009): very large effects for Harlem Children Zone
- In general, "no excuses" schools (extra hours, discipline, academic focus) tend to have positive impacts

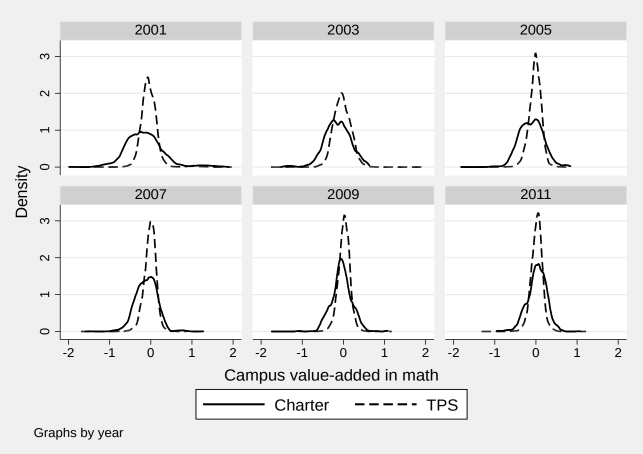
Market Competition and Charter Schools

- Does market discipline lead to growth of better schools and improvement in performance over time in equilibrium?
- Baude, Casey, Hanushek, and Rivkin (2014) study evolution of quality of charter schools in Texas
- Difficult to estimate causal effect of 500 schools using lottery-based methods
- Instead use observational approach: calculate "value-added" of each school from a regression of the form

$$A_{it} = a + bA_{i,t-1} + \gamma X_i + \varepsilon_{it}$$

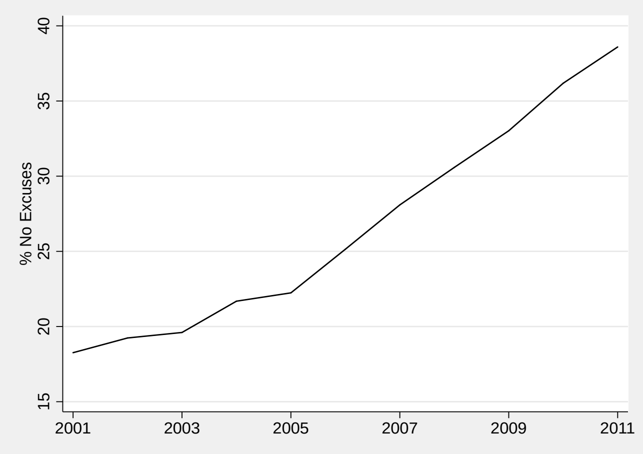
- Define "value-added" of school s in year t as mean value of ε_{it} for students enrolled in school s

Distribution of School Math VA by Year



Source: Baude et. al. 2014

Market Share of “No Excuses” Charter Schools



Source: Baude et. al. 2014

Limitations of Market Competition

- Three limitations of relying purely on private market competition

1. Markets may function poorly when quality is not well observed
    - Difficult to gauge quality when outcomes (e.g. college, earnings) are realized 10+ years after treatment
    - Concern about excessive focus on short-term test scores
2. Cream skimming of students and teachers
    - Private schools have an incentive to reject less qualified applicants
    - Can exacerbate inequality by leaving less qualified students behind in schools with fewer resources and weaker peers
3. Optimization failures
    - Parents may not make the best choices for kids

Information and School Choice

- Hastings, Kane, and Staiger (2007) study introduction of school choice in Charlotte, NC in 2002
  - Parents allowed to submit three choices for schools
  - Guaranteed admission to local "home" school; lottery for other schools
  - Find that low income parents are much less likely to choose schools with high test scores than high income parents

Information and School Choice

- Hastings and Weinstein (2008) test whether providing information to low-income households about test-score differences improves choices
- Ex ante: information on schools provided in a 100 page book and complex websites
- Natural experiment: district forced to send students at 16 NCLB-failing schools simplified info. about alternatives in mid 2004
- Randomized experiment: researchers sent 1 page brochures with test score info. of nearby schools
- Similar results from both interventions; focus on the first here

Effect of NCLB Information on Active School Choice

Variable	Spring 2004 choice round[a] (1)	July 2004 NCLB choice round[a] (2)	Difference: July–spring[b] (3)
All parents of NCLB students[c]			
Fraction choosing school and program other than	0.112	0.163	0.051***
NCLB school and program first	(0.315)	(0.369)	(0.006)
Test score of first-choice school and program	0.053	0.100	0.047***
minus test score of NCLB school and program[d]	(0.207)	(0.267)	(0.004)
Number of students	6,695	6,695	6,695
Parents who chose NCLB school and program first in spring 2004 choice round[e]			
Fraction choosing school and program other than	0	0.145	0.145***
NCLB school and program first	—	(0.353)	(0.005)
Test score of first-choice school and program	0	0.088	0.088***
minus test score of NCLB school and program	—	(0.251)	(0.003)
Number of students	5,946	5,946	5,946

Source: Hasting and Weinstein 2008

Effect of NCLB Information on Type of School Chosen

Variable	All students[a] (1)	African American[b] (2)	Not African American (3)
Test score at first-choice school and program[c]			
Spring 2004 choice round	−0.502	−0.513	−0.421
July 2004 NCLB choice round	−0.017	−0.034	0.108
Average test score of schools and programs within five miles[d]			
Spring 2004 choice round	−0.322	−0.328	−0.277
July 2004 NCLB choice round	−0.247	−0.253	−0.206
Number of students	1,092	963	129

Source: Hasting and Weinstein 2008

Estimating the Education Production Function

- Alternative approach: government directly provides public education
- Here, estimating education production function $h(z_1, ..., z_N)$ becomes very important
  - Should we try to hire better teachers or reduce class size?
  - At what ages should we invest the most?
  - What is the optimal length of school days and years?
- Literature most developed on class size and teachers

Class Size

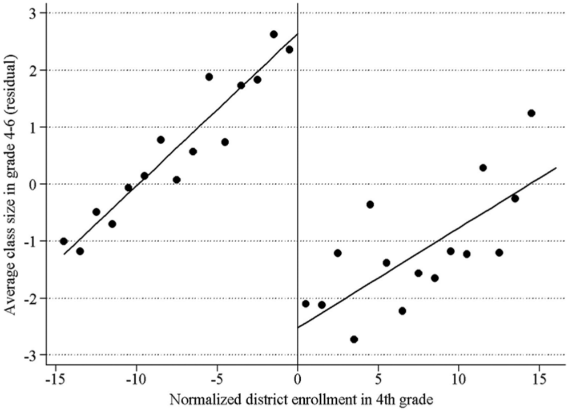
- Robust evidence that smaller class sizes improve outcomes
- Quasi-experimental evidence: RD estimates using maximum class size limits (Angrist and Lavy 1997, Fredriksson et al. 2013)
  - Angrist and Lavy: test score impacts in Israel
  - Fredriksson et al.: long-term impacts in Sweden
- Experimental evidence: Project STAR (Krueger 1999, Chetty et al. 2011)
  - Random assignment of 12,000 kids in Tennessee to classrooms in grades K-3 in mid 1980's
  - Small classes: 15 students, large classes: 23 students

STAR Experiment: Impacts of Class Size

Dep Var:	Test Score	College in 2000	College Quality	Wage Earnings	Summary Index
	(1)	(2)	(3)	(4)	(5)
Small Class	4.81	2.02%	$119	-$4	5.06%
	(1.05)	(1.10%)	($97)	($327)	(2.16%)
Observations	9,939	10,992	10,992	10,992	10,992
Mean of Dep. Var.	48.67	26.4%	$27,115	$15,912	0.00

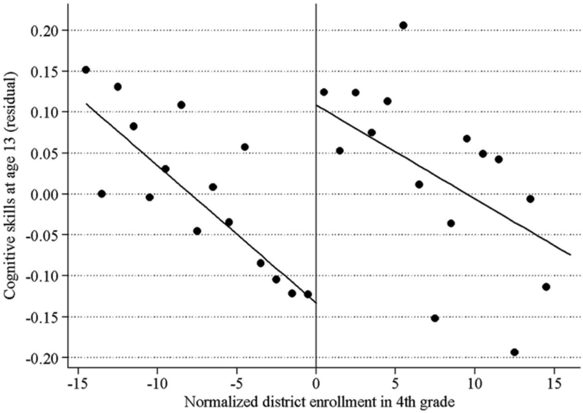
Source: Chetty et al. (QJE 2011)

RD Evidence: Class Size vs. Enrollment in Grade 4



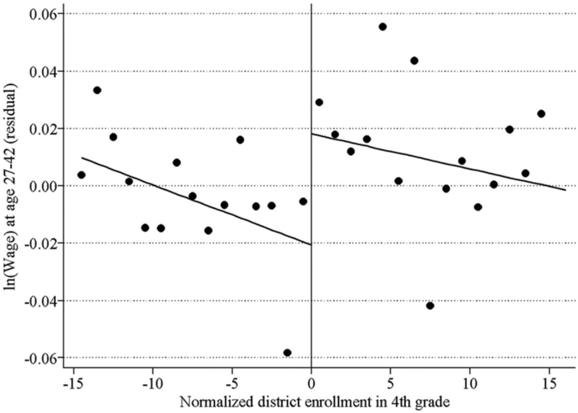
Source: Fredriksson et al. (QJE 2013)

Test Scores at Age 13 vs. Enrollment in Grade 4



Source: Fredriksson et al. (QJE 2013)

Earnings vs. Enrollment in Grade 4



Source: Fredriksson et al. (QJE 2013)

Teachers

- Large literature has studied impacts of observable teacher characteristics
  - Experience, certification, credentials, grades
- Basic design: do students' test gains vary depending upon observable chars. of teacher?
  - General conclusion: little evidence that observables matter (e.g., Kane, Rockoff, and Staiger 2008)
  - Except experience in the first 2-3 years
- Important caveat: this is based on variation within the sample of individuals *currently* applying for teaching
  - If only weak students from top colleges apply for teaching, then it will appear that college is not predictive

Input vs. Output-Based Policies

- Hanushek (2003): "Failure of Input-Based Schooling Policies"
  - Characterizing the production function based on inputs z has not been very successful
- Alternative: output-based approach
  - Identify teacher quality from ex-post *outcomes* such as student test scores
  - Performance-based measures

Value-Added Metrics

- Value-added (VA) measures rate teachers based on impacts on students' test scores
    - Long academic history: originally proposed by Hanushek (1971) and Murnane (1975)
- School districts have recently started to use VA to evaluate teachers
    - Ex: Washington DC put 50% weight on VA measures in making teacher layoff and bonus decisions under Michelle Rhee
    - *Vergara v. California* case on teacher tenure focused on VA measures

Debate About Teacher Value-Added

- Debate about value-added stems primarily from three issues:

  1. Potential for bias in VA estimates [Kane and Staiger 2008, Rothstein 2010]
     - Do differences in test-score gains across teachers capture causal impacts of teachers or are they driven by student sorting?
  2. Lack of evidence on teachers's long-term impacts
     - Do teachers who raise test scores improve students' long-term outcomes or are they simply better at teaching to the test?
  3. Measurement error in VA estimates
     - Are estimates of teacher quality based on a few years of data too noisy to be useful for policy?

Chetty, Friedman, Rockoff 2014

- Chetty, Friedman, Rockoff (2014a,b) answer these questions by tracking 2.5 million children from childhood to early adulthood

    1. Develop new quasi-experimental tests for bias in VA estimates
    2. Test if children who get high VA teachers have better outcomes in adulthood
    3. Measure monetary gains from selecting teachers with higher estimated VA given observed measurement error

Constructing Value-Added Estimates

- Model the estimation of VA as a forecasting problem
- Simplest case: teachers teach one class per year with N students
- All teachers have test score data available for t previous years
- Objective: predict test scores for students taught by teacher j in year $t+1$ using test score data from previous t years

Constructing Value-Added Estimates

- Three steps to estimate VA ($\hat{\mu}_{j,t+1}$) for teacher j in year $t+1$

1. Form residual test scores A_{is}, controlling for observables X_{is}
    - Regress raw test scores A^*_{is} on observable student characteristics X_{is}, including prior test scores $A^*_{i,s-1}$

2. Regress mean class-level test score residuals in year t on class-level test score residuals in years 0 to $t-1$:

$$\bar{A}_{jt} = a + \psi_{t-1}\bar{A}_{j,t-1} + ... + \psi_0\bar{A}_{j0} + \varepsilon_{jt}$$

3. Use estimated coefficients $\psi_1, \ldots, \psi_t$ to predict VA in year $t+1$ based on mean test score residuals in years 1 to t for each teacher j:

$$\hat{\mu}_{j,t+1} = \sum_{s=1}^{t} \psi_s \bar{A}_{js}$$

Constructing Value-Added Estimates

- Two special cases:

  1. Forecast VA in year t using data from only year $t-s$:

     $$\hat{\mu}_{jt} = r_s \bar{A}_{j,t-s}$$

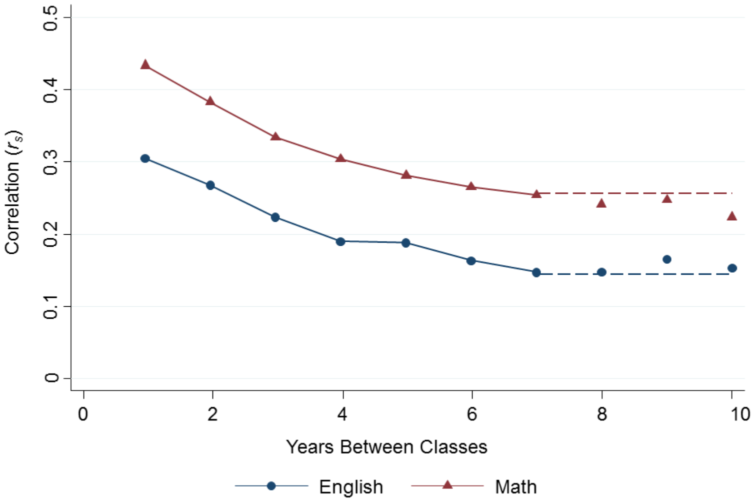
     where $r_s = Corr(\bar{A}_t, \bar{A}_{t-s})$ is autocorrelation at lag s

  2. Without drift, put equal weight on all prior scores:

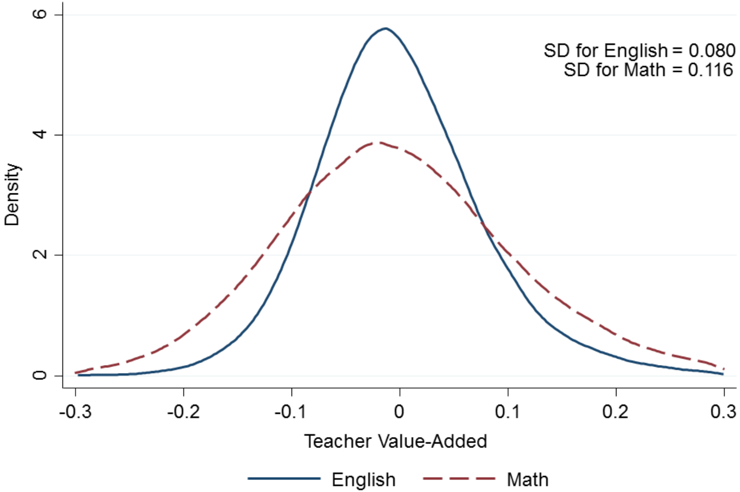
     $$\hat{\mu}_{jt} = \bar{A}_j^{-t} \frac{\sigma_\mu^2}{\sigma_\mu^2 + (\sigma_\theta^2 + \sigma_\epsilon^2/n)/T}$$

     - Bayesian interpretation: shrinkage based on signal-noise ratio (Kane and Staiger 2008)

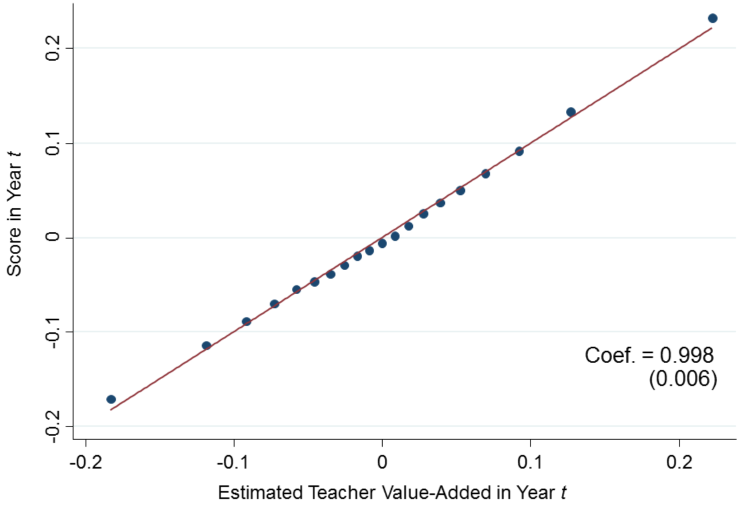
Autocorrelation Vector in Elementary School



Distribution of VA Estimates



Test Score Residuals vs. VA in Cross-Section



Econometric Challenges in Value-Added Models

- Typically identify causal effects of VA by estimating models such as

$$y_i = a + b\hat{\mu}_{jt} + \varepsilon_i$$

- Key econometric complication: $\hat{\mu}_{jt}$ is itself *estimated*
  - Differs from class size, which is measured without error
- Examples of problems that arise:
  1. Estimating long-term impacts using the same data used to estimate VA leads to upward bias in b
     - Having a very smart set of students will lead to high $\hat{\mu}_{jt}$ and those students will have high earnings
  2. Standard placebo tests can fail
     - Ex: current teacher VA can be correlated with past scores if VA is estimated using data on prior scores, as in Rothstein (2010, 2014)

Are VA Estimates Biased?

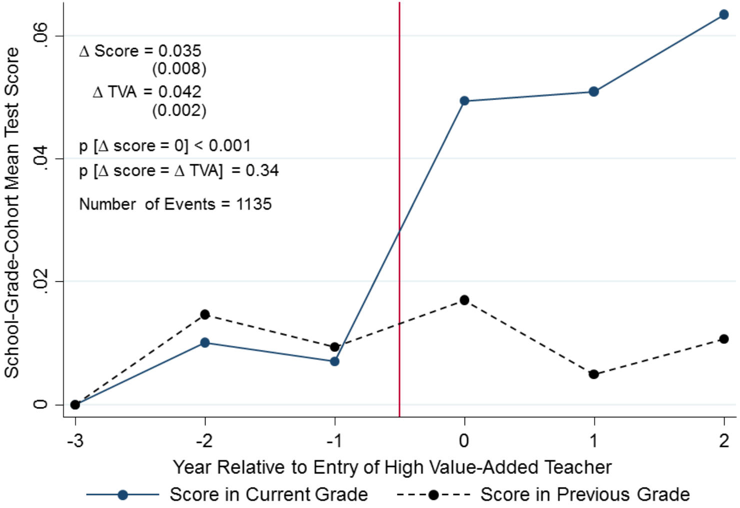
- Let γ denote causal impact of 1 unit increase in teacher's estimated VA on student's test score
  - Define forecast bias as $B = 1 - \gamma$
- Ideal experiment to estimate forecast bias (Kane and Staiger 2008): randomly assign students to teachers with different VA estimates
  - Does a student who is randomly assigned to a teacher previously estimated to be high VA have higher test score gains?
- We use teacher switching as a quasi-experimental analog

Teacher Switchers in School-Grade-Subject-Year Data

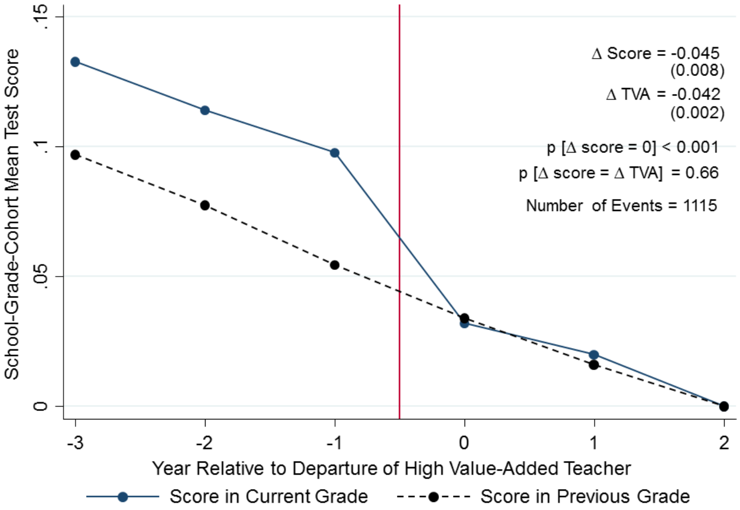
School	Grade	Subject	Year	Teachers	Mean Score	Mean Age 28 Earnings
1	5	math	1992	Jones, Heckman, ...	-.09	$15K
1	5	math	1993	Jones, Heckman, ...	-.04	$17K
1	5	math	1994	Jones, Heckman, ...	-.05	$16K
1	5	math	1995	Katz, Heckman, ...	0.01	$18K
1	5	math	1996	Katz, Heckman, ...	0.04	$17K
1	5	math	1997	Katz, Heckman, ...	0.02	$18K

- Jones switches to a different school in 1995; Katz replaces him

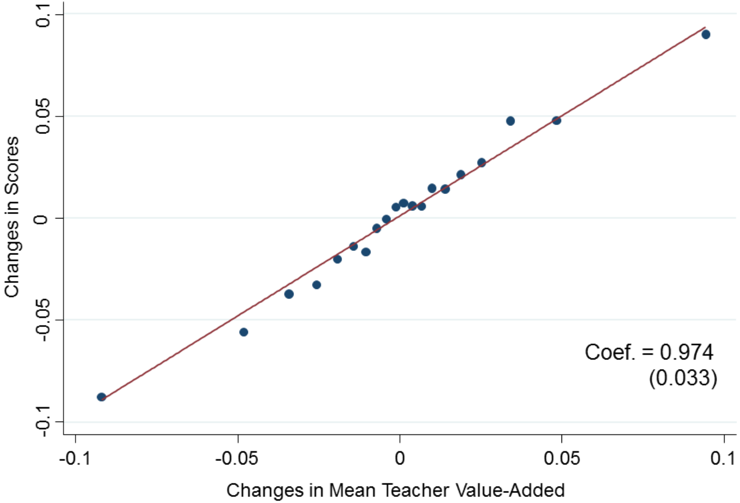
Impact of High VA Teacher Entry on Cohort Test Scores



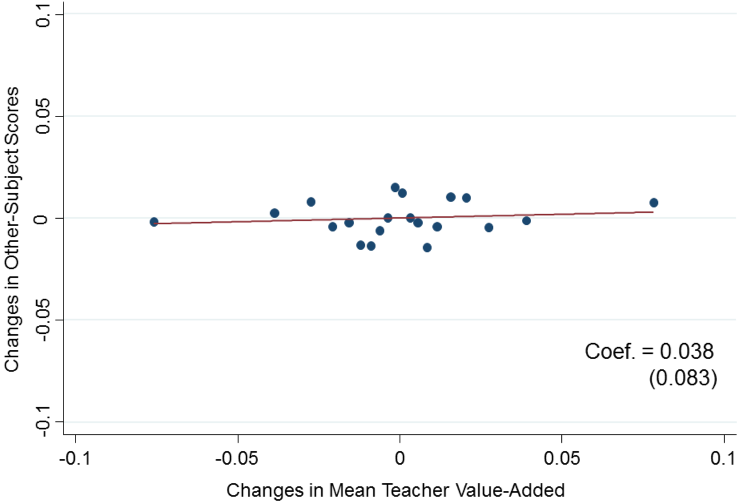
Impact of High VA Teacher Exit on Cohort Test Scores



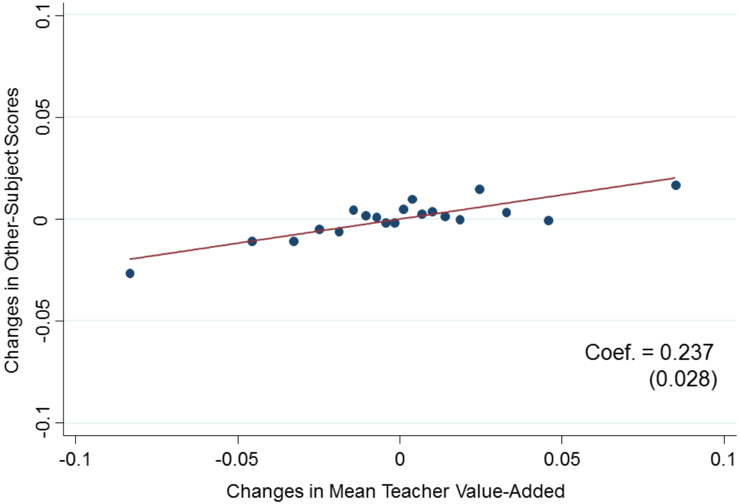
Changes in Mean Scores vs. Changes in Mean Teacher VA



Changes in Other-Subject Scores vs. VA: Middle Schools



Changes in Other-Subject Scores vs. VA: Elem. Schools



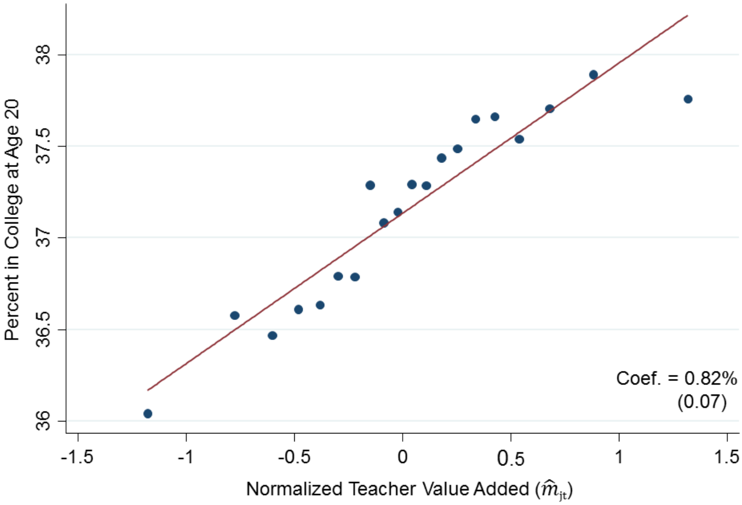
Estimates of Forecast Bias with Alternative Control Vectors

Control Vector	Quasi-Experimental Estimate of Bias (%)
Baseline	2.58 (3.34)
Student-level lagged scores	4.83 (3.29)
Non-score controls only	45.39 (2.26)
No controls	65.58 (3.73)

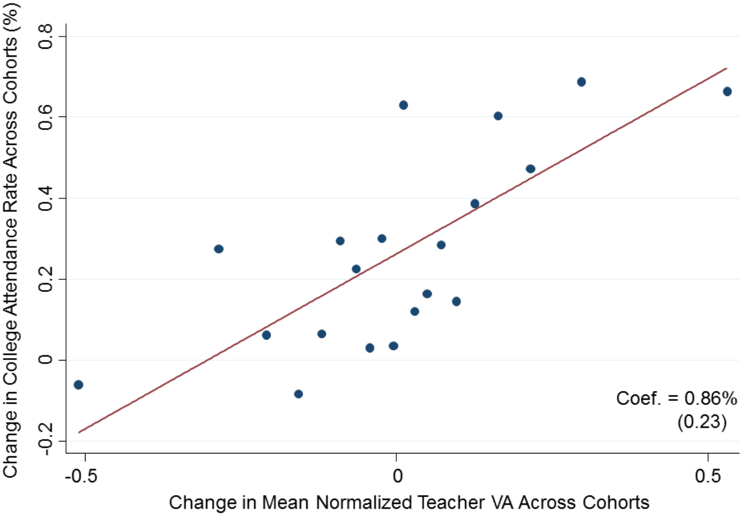
Impacts on Outcomes in Adulthood

- Do teachers who raise test scores also improve students' long-run outcomes?
- Regress long-term outcomes on teacher-level VA estimates
  - Then validate using cross-cohort switchers design
- Interpretation of these reduced-form coefficients (Todd and Wolpin 2003):
  - Impact of having better teacher, as measured by VA, for single year during grades 4-8 on earnings
  - Includes benefit of better teachers, peers, etc. in later grades via tracking

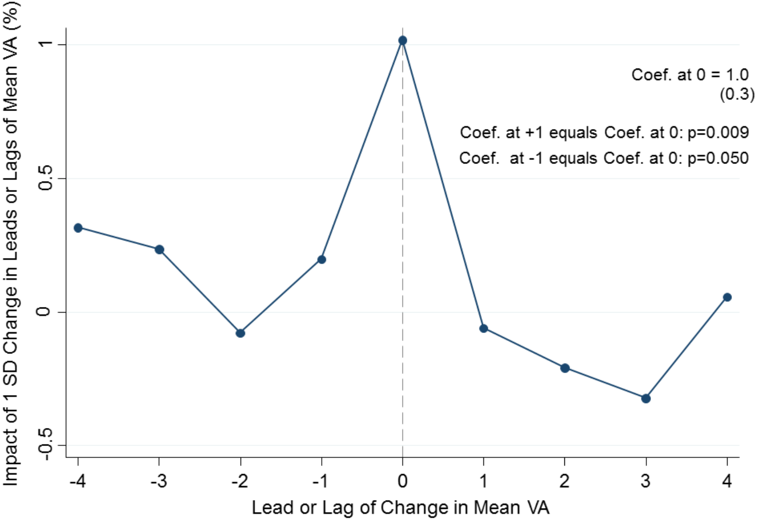
College Attendance at Age 20 vs. Teacher Value-added



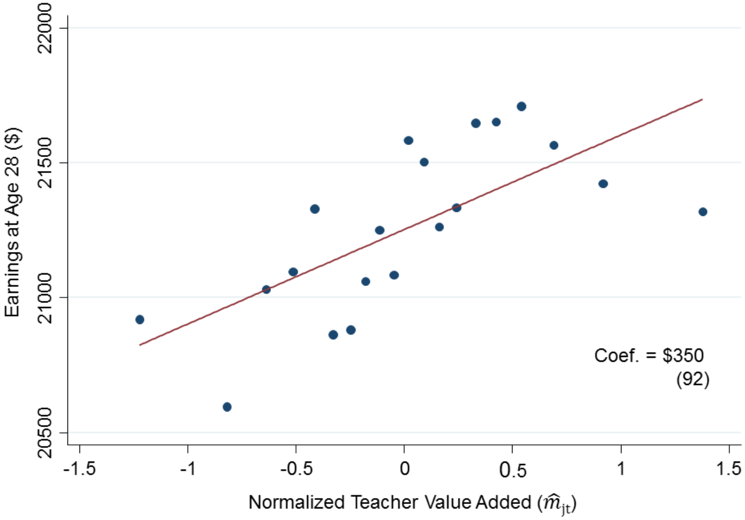
College Attendance: Cross-Cohort Design



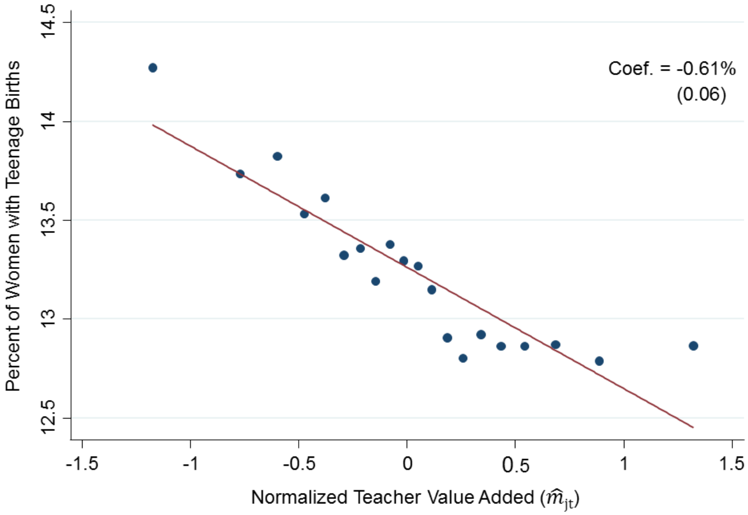
Event Study of Coefficients on College Attendance



Earnings at Age 28 vs. Teacher Value-Added



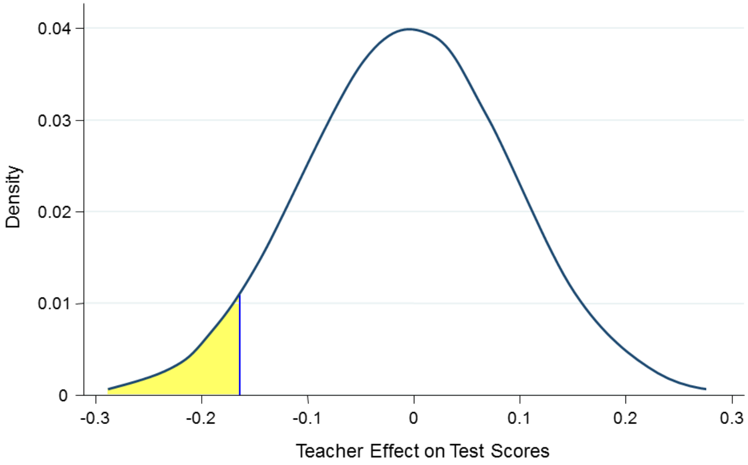
Women with Teenage Births vs. Teacher Value-Added



Policy Analysis and Noise in VA Estimates

- What is the impact of selecting teachers based on their VA on students' lifetime earnings?
- Any evaluation of teachers based on VA must rely on only a few years of classroom data
  - This generates noise in VA estimates, potentially reducing its utility for performance evaluation
- Simulate gains from selection policies accounting for noise

Deselecting Teachers on the Basis of Value-Added



Assumptions

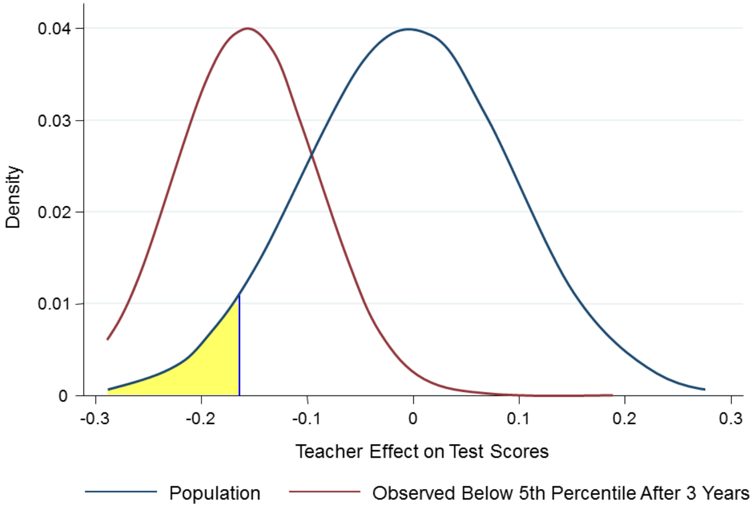
1. Ignore general equilibrium effects and non-monetary gains
2. Constant percentage impact on earnings over life
3. 2% wage growth with 5% discount rate back to age 12
    - Undiscounted lifetime earnings gains are roughly 5 times larger

Selection on True Value-Added

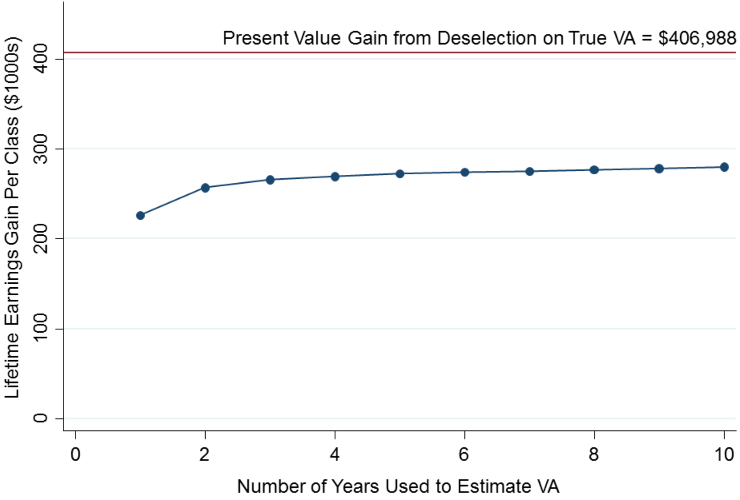
- Deselecting bottom 5% of teachers on true VA yields NPV earnings gain for a classroom of average size = \$407,000

- In practice, gains are reduced by estimation error in VA

Deselecting Teachers on VA Estimates



Earnings Impacts Accounting for Noise



Policy Implications

- Optimal teacher evaluation policy can be modelled as a stopping rule problem (Rockoff and Staiger 2010)
- Rapidly diminishing gains to information suggest that it is best to make decisions quickly
- Optimal policy weighs gains from teacher selection against cost imposed by higher risk
  - Rothstein (2014) studies a structural model of the teacher labor market
  - Estimates that deselecting bottom 5% of teachers based on VA would require a salary increase of $700 for all teachers
- Avg. gain from deselection policy is $184,000 x 5% = $9,250
- Gain 10 times as large as cost $\Rightarrow$ selection on VA could potentially be a useful policy tool

Open Questions

- Two key issues remain to be resolved before one can determine optimal way to use VA for policy

1. Gains may be eroded when VA is actually used (Lucas critique, Campbell's "Law")
    - Using VA in high-stakes evaluation could lead to teaching to the test or cheating
2. Need to compare VA to other metrics
    - Classroom observation, principal/peer evaluation, non-cognitive assessments
    - What is the optimal weight on each measure to predict earnings impacts?

- Teacher switching methodology can be used for both purposes

Directions for Future Research

- Many important questions remain
  - Are some teachers better at teaching some types of students?
  - Could resort teachers and students instead of hiring new teachers
- Complementarities in production across teachers
  - Are effects of having a good teacher in grade g and grade $g+1$ super-additive?
  - Which grades are most important?
- Bringing in theory to structure policy analysis
  - E.g. sufficient statistic approach to optimal policy design

Directions for Future Research

- More broadly, substantial scope to implement similar methods in improving other sectors
  - Tax preparer effects, manager effects, doctor effects,...
  - Developing methods to measure and improve efficiency of govt. agencies
- Public perception of government inefficiency focuses on production inefficiency rather than just deadweight loss
  - Using tools of public economics to improve production has great potential returns