How Does Your Kindergarten Classroom Affect Your Earnings?
Evidence from Project STAR

Raj Chetty, Harvard
John N. Friedman, Harvard
Nathaniel Hilger, Harvard
Emmanuel Saez, UC Berkeley
Diane Schanzenbach, Northwestern
Danny Yagan, Harvard

February 2011
What are the long-term impacts of early childhood education?

Limited evidence because few datasets link information on childhood education with adult outcomes

We link data from the STAR experiment to U.S. tax records to analyze how class assignment in grades K-3 affects adult outcomes
Student/Teacher Achievement Ratio (STAR) experiment:

- Conducted from 1985 to 1989 in Tennessee
- One cohort of 11,571 children in grades K-3 at 79 schools
- Most children born in 1979-80 \( \rightarrow \) graduate high school in 1998

Students and teachers randomized into classrooms within schools

- Class size differs: small (15 students) or large (22 students)
- Classes also differ in teachers and peers
- Randomized into classroom upon entry into participating school and kept in small/large track from grades K-3

Only one cohort treated \( \rightarrow \) no repeat teacher observations
Project STAR Background

- Large literature on STAR shows that class size, teacher quality, and peer quality have causal impacts on scores
  - Students in small classes have 5 percentile point (0.2 sd) higher test scores in K-3 (Krueger 1999)
  - But test score gains fade out to 1-2 percentiles by grade 8
  - Similar fade out effects observed in other early childhood interventions (e.g. Currie and Thomas 1995, Deming 2009)

→ Do early test score gains translate into impacts on adult outcomes?
United States Tax Data

- Dataset covers full U.S. population from 1996-2008

- Approximately 90% of working age adults file tax returns

- Third-party reports yield data on many outcomes even for non-filers
  - Employer and wage earnings from W-2 forms
  - College attendance from 1098-T forms

- 95% of STAR records were linked to tax data
<table>
<thead>
<tr>
<th></th>
<th>STAR Sample (1)</th>
<th>U.S. 1979-80 cohort (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Wage Earnings (2005-07)</td>
<td>$15,912</td>
<td>$20,500</td>
</tr>
<tr>
<td>Zero Wage Earnings (2005-07)</td>
<td>13.9%</td>
<td>15.6%</td>
</tr>
<tr>
<td>Attended College in 2000 (age 20)</td>
<td>26.4%</td>
<td>34.7%</td>
</tr>
<tr>
<td>Attended College by age 27</td>
<td>45.5%</td>
<td>57.1%</td>
</tr>
<tr>
<td>Mean Parents’ Income (1996-98)</td>
<td>$48,010</td>
<td>$65,660</td>
</tr>
</tbody>
</table>
1. Test scores and adult outcomes in the cross-section

2. Impacts of observable classroom characteristics

3. Impacts of unobservable classroom characteristics

4. Fade-out, re-emergence, and non-cognitive skills

5. Cost-benefit analysis
Part 1: Cross-Sectional Correlations

- Begin by correlating KG test scores with adult outcomes
  - Useful to benchmark estimates from randomized interventions

- Estimate both raw correlations and OLS regressions with controls:
  - quartic in parental household income interacted with marital status
  - mother age at child’s birth
  - parent’s 401K contributions, home ownership
  - child’s gender, free lunch status, race, and age

- Test score: percentile score on Kindergarten Stanford Achievement Test (math + reading)
What is a kindergarten test?

- Instructions:

  - I’ll say a word to you. Listen for the *ending* sound.

  - You circle the picture that *starts* with the same sound.

“cup”
Figure 1a: Wage Earnings vs. KG Test Score

Mean Wage Earnings from Age 25-27 vs. KG Test Score Percentile

$10K
$15K
$20K
$25K

R² = 0.05
## Test Scores and Earnings in the Cross-Section

<table>
<thead>
<tr>
<th>Dependent Var.:</th>
<th>Wage Earnings</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KG Test Percentile</td>
<td>$132</td>
<td>$93.8</td>
<td>$90.0</td>
<td>$97.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($12.2)</td>
<td>($11.6)</td>
<td>($8.65)</td>
<td>($8.47)</td>
<td></td>
</tr>
<tr>
<td>Parental Income Percentile</td>
<td>$146</td>
<td>(8.15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry Grade</td>
<td>KG</td>
<td>KG</td>
<td>All</td>
<td>All</td>
<td></td>
</tr>
<tr>
<td>Class Fixed Effects</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Student Controls</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Parent Controls</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.05</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,621</td>
<td>5,621</td>
<td>9,939</td>
<td>9,939</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1b: College Attendance Rates vs. KG Test Score

Attended College before Age 27 vs. KG Test Score Percentile
We construct an index of college quality using tax data

Tuition paid to any higher ed. institution (Title IV) automatically generates a 1098-T form linking student and institution

Calculate average wage earnings (from W-2s) by college

For those who do not attend college, define college quality index as mean earnings for those not in college in 1999
### An Earnings-Based Index of College Quality

<table>
<thead>
<tr>
<th>US News Ranking</th>
<th>College</th>
<th>Mean Earnings at age 28</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Harvard</td>
<td>$79,643</td>
</tr>
<tr>
<td>2</td>
<td>Princeton</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Yale</td>
<td>$79,643</td>
</tr>
<tr>
<td>4</td>
<td>Cal Tech</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Stanford</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>U Penn</td>
<td>$75,570</td>
</tr>
<tr>
<td>8</td>
<td>Columbia</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>U Chicago</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Duke</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>Arizona St.</td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>Catholic U</td>
<td></td>
</tr>
<tr>
<td>123</td>
<td>MI Tech</td>
<td>$46,390</td>
</tr>
<tr>
<td>124</td>
<td>U Buffalo</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>U San Fran</td>
<td></td>
</tr>
</tbody>
</table>
Home Ownership vs. KG Test Score

Owned a Home by Age 27

KG Test Score Percentile

Owned a Home by Age 27

KG Test Score Percentile
Retirement Savings vs. KG Test Score

Made a 401(k) Contribution by Age 27

KG Test Score Percentile

- 20%
- 25%
- 30%
- 35%
- 40%
- 45%

0 20 40 60 80

KG Test Score Percentile

- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

Made a 401(k) Contribution by Age 27

- <20%
- 20-30%
- 30-40%
- 40-50%
- 50-60%
- 60-70%
- 70-80%
- 80-90%
- >90%

Graph shows a positive correlation between KG Test Score Percentile and the percentage of individuals who made a 401(k) contribution by age 27.
Cross-State Mobility vs. KG Test Score

Lived Outside TN before Age 27 vs. KG Test Score Percentile

KG Test Score Percentile

0 20 40 60 80

Lived Outside TN before Age 27

0 20 40 60 80 100
Percent College Graduates in ZIP code vs. KG Test Score

Percent College Graduates in 2008 ZIP:
- 14%
- 16%
- 18%
- 20%
- 22%

KG Test Score Percentile:
- 0
- 20
- 40
- 60
- 80
- 100
Part 2: Validity of the STAR Experimental Design

- Validity of experimental analysis rests on two assumptions:

- Assumption 1: *Randomization*
  
  - All pre-determined variables (e.g. parent characteristics) are balanced across classrooms

- Assumption 2: *No Differential Attrition*
  
  - 95% match rate $\rightarrow$ little attrition here
  
  - No evidence of differences in match rates across classrooms
  
  - No evidence of differences in death rates across classrooms
Part 2: Validity of the STAR Experiment Design

- Threat #1: *Failure of Randomization*

  - Prior studies had few baseline measures, limiting ability to evaluate randomization protocol (Schanzenbach 2006)

  - We test for balance across class types with an expanded set of parent/sibling characteristics in two ways:

  1. Do characteristics vary across small vs. large class types?

  2. Do characteristics vary across classrooms within schools?
<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Wage Earnings (%) (1)</th>
<th>Small Class (%) (2)</th>
<th>Teacher Exp. (%) (3)</th>
<th>Class Effects p-value (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent’s Income ($1,000s)</td>
<td>65.47 (-0.003)</td>
<td>-0.001</td>
<td>0.848</td>
<td></td>
</tr>
<tr>
<td>Mother’s Age at STAR Birth</td>
<td>53.96 (0.029)</td>
<td>0.022</td>
<td>0.654</td>
<td></td>
</tr>
<tr>
<td>Parents Have 401(k)</td>
<td>2273 (1.455)</td>
<td>0.111</td>
<td>0.501</td>
<td></td>
</tr>
<tr>
<td>Student Female</td>
<td>-2317 (-0.226)</td>
<td>0.236</td>
<td>0.502</td>
<td></td>
</tr>
<tr>
<td>Student Black</td>
<td>-620.8 (0.204)</td>
<td>0.432</td>
<td>0.995</td>
<td></td>
</tr>
<tr>
<td>p-Value of F Test</td>
<td>0.000</td>
<td>0.261</td>
<td>0.190</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>10,992</td>
<td>10,992</td>
<td>10,914</td>
<td></td>
</tr>
</tbody>
</table>

Note: Regressions include school-by-entry-grade fixed effects.
Validity of the STAR Experiment Design

- Threat #2: *Selective Attrition*

- Much less attrition than in prior studies of STAR because we follow 95% of the sample

- Test for selective attrition through two channels:
  1. Does match rate vary across treatment groups?
  2. Does death rate vary across treatment groups (Muennig et al. 2010)?
Table 3: Tests for Selective Attrition

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Matched</th>
<th></th>
<th></th>
<th></th>
<th>Deceased</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
<td></td>
<td>(%)</td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>Small Class</td>
<td>-0.019</td>
<td>0.079</td>
<td></td>
<td>-0.010</td>
<td>-0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.467)</td>
<td>(0.407)</td>
<td></td>
<td>(0.286)</td>
<td>(0.286)</td>
<td></td>
</tr>
<tr>
<td>p Value on F test on Class Effects</td>
<td>0.951</td>
<td>0.888</td>
<td></td>
<td>0.388</td>
<td>0.382</td>
<td></td>
</tr>
<tr>
<td>Demographic Controls</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Mean of dep. Var.</td>
<td>95.0</td>
<td>95.0</td>
<td></td>
<td>1.70</td>
<td>1.70</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>11,571</td>
<td>11,571</td>
<td></td>
<td>10,992</td>
<td>10,992</td>
<td></td>
</tr>
</tbody>
</table>
Part 3: Class Size Impacts

- Regress outcomes on dummy for small class assignment (intent to treat) with school fixed effects

- Analyze impacts on four outcomes:
  
  1. College attendance
  
  2. College quality index
  
  3. Mean earnings (ages 25-27)
  
  4. Standardized (SD = 1) summary index of other outcomes:

\[
\text{Index} = 401K + \text{Home Owner} + \text{Married} + \\
\text{Moved out of TN} + \text{Pct. College Grads. in Zip}
\]
Figure 2a: Effect of Class Size on College Attendance by Year

Percent Attending College

Year

10%

15%

20%

25%

30%

2000

2002

2004

2006

Large Class

Small Class
Figure 2b: College Earnings Quality by Class Size
Figure 2c: Effect of Class Size on Wage Earnings by Year

Wage Earnings

- $6K
- $8K
- $10K
- $12K
- $14K
- $16K
- $18K

Year

- 2000
- 2002
- 2004
- 2006

Large Class

Small Class
Table 5: Impacts of Class Size on Adult Outcomes

<table>
<thead>
<tr>
<th>Dependent Var.:</th>
<th>College In 2000 (1)</th>
<th>College Quality (2)</th>
<th>Wage Earnings (3)</th>
<th>Summary Index (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Class</td>
<td>2.02%</td>
<td>$119</td>
<td>$4</td>
<td>5.06%</td>
</tr>
<tr>
<td></td>
<td>(1.10%)</td>
<td>($97)</td>
<td>($327)</td>
<td>(2.16%)</td>
</tr>
<tr>
<td>Observations</td>
<td>10,992</td>
<td>10,992</td>
<td>10,992</td>
<td>10,992</td>
</tr>
<tr>
<td>Mean of Dep. Var.</td>
<td>26.4%</td>
<td>$27,115</td>
<td>$15,912</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: All specifications control for school-by-entry-grade effects.
Part 3: Teacher/Peer Effects

- Students randomly assigned to classes that differ in teacher and peer quality

  → Do teachers/peers affect adult outcomes?

- First test: does random assignment to a more experienced KG teacher improve adult outcomes?

  - Not necessarily causal effect of raising teacher experience *per se*

  - Experienced teachers may also differ on other dimensions such as dedication to teaching
Figure 3a: Effect of Teacher Experience on Test Scores
Figure 3b: Effect of Teacher Experience on Earnings

Mean Wage Earnings, 2005-2007

Kindergarten Teacher Experience (Years)

$16K

$17K

$18K

$19K

0

5

10

15

20
Figure 3c: Effect of Teacher Experience on Earnings by Year

- **Teacher Experience <=10 Years**
- **Teacher Experience > 10 Years**

Year

- 2000
- 2002
- 2004
- 2006

Wage Earnings

- $8K
- $10K
- $12K
- $14K
- $16K
- $18K
- $20K

- $1104
Table 6: Observable Teacher vs. Peer Effects

<table>
<thead>
<tr>
<th>Dependent Var.:</th>
<th>Test Score (1)</th>
<th>Wage Earnings (2)</th>
<th>Wage Earnings (3)</th>
<th>Wage Earnings (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher with &gt;10 Years Experience</td>
<td>3.18%</td>
<td>$1093</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher has post-BA deg.</td>
<td>-0.85%</td>
<td>-$261</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Black Classmates</td>
<td></td>
<td>-$1,757</td>
<td>($2,692)</td>
<td></td>
</tr>
<tr>
<td>% Female Classmates</td>
<td></td>
<td>-$67.5</td>
<td>($1,539)</td>
<td></td>
</tr>
<tr>
<td>% Free-Lunch Classmates</td>
<td></td>
<td>-$285</td>
<td>($1,731)</td>
<td></td>
</tr>
<tr>
<td>Classmates’ Mean Age</td>
<td></td>
<td>-$25.8</td>
<td>($1,359)</td>
<td></td>
</tr>
<tr>
<td>Classmates’ Mean Pred. Score</td>
<td></td>
<td></td>
<td>-$23.3</td>
<td>($93.7)</td>
</tr>
<tr>
<td>Entry Grade</td>
<td>KG</td>
<td>KG</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Observations</td>
<td>5,601</td>
<td>6,005</td>
<td>10,992</td>
<td>10,992</td>
</tr>
</tbody>
</table>

Note: All specifications control for school fixed effects and class size, as well as student and parent demographics.
Many elements of teacher and peer quality (e.g. clarity of instruction, enthusiasm) are not observable

- Well known problem in literature on teacher effects

Test for “class effects” on adult outcomes using analysis of variance

- Is there significant intra-class correlation in student’s outcomes?

- This class effect includes effect of teachers, peers, and any class-level shocks such as noise outside classroom

Formally, we are testing for clustering of outcomes by (randomly assigned) classroom
A Model of Class Effects

- Test scores and earnings for individual $i$ in class $c$:

$$s_{ic} = z_c + a_{ic}$$

$$y_{ic} = \beta z_c + \gamma z^Y_c + \rho a_{ic} + v_{ic}$$

- $z_c$ = class-level intervention (e.g. better teaching) that affects scores and earnings

- $z^Y_c$ = intervention that affects earnings but not scores
A Model of Class Effects

- Test scores and earnings for individual $i$ in class $c$:

  $$ s_{ic} = z_c + a_{ic} $$

  $$ y_{ic} = \beta z_c + \gamma z^Y_c + \rho a_{ic} + \nu_{ic} $$

- $z_c = \text{class-level intervention (e.g. better teaching) that affects scores and earnings}$

- $z^Y_c = \text{intervention that affects earnings but not scores}$

- $a_{ic} = \text{academic ability}$

- $\nu_{ic} = \text{earnings ability orthogonal to academic ability}$
A Model of Class Effects

- Test scores and earnings for individual $i$ in class $c$:

\[ s_{ic} = z_c + a_{ic} \]
\[ y_{ic} = \beta z_c + \gamma z_c^Y + \rho a_{ic} + v_{ic} \]

- $z_c =$ class-level intervention (e.g. better teaching) that affects scores and earnings

- $z_c^Y =$ intervention that affects earnings but not scores

- $a_{ic} =$ academic ability

- $v_{ic} =$ earnings ability orthogonal to academic ability

- $\beta + \gamma =$ impacts of interventions on earnings

- $\beta =$ covariance of class effects on scores and earnings
A Model of Class Effects

- Test scores and earnings for individual $i$ in class $c$:

  $$s_{ic} = z_c + a_{ic}$$
  $$y_{ic} = \beta z_c + \gamma y + \rho a_{ic} + v_{ic}$$

- Thus far, we have estimated $\beta$ directly by using observable $z$’s that affect test scores (e.g. teacher experience)

- How can we estimate $\beta$ and $\gamma$ when class-level interventions are unobserved?
Test for class effects on earnings \((\beta + \gamma > 0)\) using ANOVA

Do earnings vary across classes by more than what would be predicted by random variation in student abilities?

Two steps:

1. [Fixed effects] Test for significance of class fixed effects
2. [Random effects] Estimate class-level SD of outcomes assuming normally distributed class effects
<table>
<thead>
<tr>
<th>Dependent Var.: Grade K Scores</th>
<th>Grade 8 Scores</th>
<th>Wage Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>P-value of F-Test on KG</td>
<td>0.000</td>
<td>0.419</td>
</tr>
<tr>
<td>Class Fixed Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic Controls</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Large Classes Only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observable Class Chars.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,621</td>
<td>4,448</td>
</tr>
</tbody>
</table>

Note: All specifications control for school fixed effects and class size.
ANOVA does not tell us whether class effects on scores are correlated with class effects on earnings ($\beta > 0$)

- Do class-level interventions that raise test scores also improve adult outcomes?

→ Turn to a second strategy to measure covariance between class effects on scores and earnings ($\beta$)

- What is the correlation of class effects on scores and class effects on earnings?

- Derive estimator for $\beta$ and prove it is unbiased in paper; give a heuristic explanation here
Peer-Score Measure of Class Quality

- Average *end-of-year* test scores in class relative to school $s_c$ is a (noisy) measure of class effect on scores:

$$s_c = z_c + \frac{1}{I} \sum_{j=1}^{I} a_{jc}$$

- Motivates regression of the form:

$$y_{ic} = a + b^M s_c + \epsilon_{ic}$$

- Own-observation bias: with finite class size, $\mathbb{E} b^M > 0$ even if $b^M = 0$
  - Smart kid raises average class score and has high earnings
  - Analogous to bias in 2SLS estimate with weak instruments

$\Rightarrow$ Use jackknife (leave-out mean) to measure class effect on scores: $s^{\cdot i}_c$
Peer-Score Measure of Class Quality

- Regression specification:

\[ y_{ic} = a + b^{LM} s_{c-i} + \varepsilon_{ic} \]

- This regression does not estimate peer effects because we are using end-of-year test scores
  - Class quality \( s_{c-i} \) captures teacher quality + class-level shocks
  - Good teachers raise peers’ end of year scores

- Class quality \( s_{c-i} \) varies randomly within schools
  - Can test whether classes that generate test score gains also generate earnings gains
Peer-Score Measure of Class Quality

Regression specification:

\[ y_{ic} = a + b^{LM} s_{c}^{-i} + \varepsilon_{ic} \]

Three remaining sources of bias in \( b^{LM} \)

1. Mechanical: Peers below-avg. \( \rightarrow \) you are above avg. (Guryan, Kroft, Notowidigdo 2009). Solution: define intercept using leave-out mean

2. Attenuation: \( s_{c}^{-i} \) is a noisy measure of class quality

3. Reflection: with peer effects, smart kid raises peers’ scores and earns a lot, driving up \( b^{LM} \)

After presenting results, we bound reflection bias and show it is smaller than attenuation bias
Figure 4a: Effect of Early Childhood Class Quality on Own Score

Own Test Score Percentile vs. Class Quality (End-of-Year Peer Scores)
Figure 4c: Effect of Early Childhood Class Quality on Earnings

Mean Wage Earnings, 2005-2007

Class Quality (End-of-Year Peer Scores)
Figure 5a: Effect of Class Quality on Earnings by Year

Wage Earnings

Year

Below-Average Class Quality
Above-Average Class Quality
### Table 8a: Impacts of Class Quality on Earnings

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Wage Earnings ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Class Quality (peer scores)</td>
<td>50.61</td>
</tr>
<tr>
<td></td>
<td>(17.45)</td>
</tr>
<tr>
<td>Entry Grade</td>
<td>All</td>
</tr>
<tr>
<td>Observable class chars.</td>
<td>x</td>
</tr>
<tr>
<td>Observations</td>
<td>10,959</td>
</tr>
</tbody>
</table>

NOTE--All regressions control for student and parent demographics and school-by-entry-grade fixed effects.
### Table 8b: Impacts of Class Quality on Other Adult Outcomes

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>College in 2000 (%)</th>
<th>College by Age 27 (%)</th>
<th>College Quality ($)</th>
<th>Summary Index (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Quality</td>
<td>0.096</td>
<td>0.108</td>
<td>9.328</td>
<td>0.250</td>
</tr>
<tr>
<td>(peer scores)</td>
<td>(0.046)</td>
<td>(0.053)</td>
<td>(4.573)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>Observations</td>
<td>10,959</td>
<td>10,959</td>
<td>10,959</td>
<td>10,959</td>
</tr>
</tbody>
</table>

**NOTE**--All regressions control for student and parent demographics and school-by-entry-grade fixed effects.
Figure 6a: Fadeout of Class Effects
Effect of 1 SD of Class Quality on Test Scores by Grade
Figure 6b: Fadeout of Class Effects
Effect of 1 SD of Class Quality on Earnings

Wage Earnings

Grade

$0
$200
$400
$600
$800
$1000

95% CI

1 SD Class Quality Effect on Wage Earnings

95% CI
Bounding Reflection Bias

- Small impact of KG class quality on subsequent test scores places a tight upper bound on reflection bias.

- Smart kids score high on all tests (test scores highly autocorrelated).

- To have large reflection bias, smart kid must raise peer scores a lot.

  - Large correlation between peer scores and own score in later grades.

- We formalize this intuition in a linear-in-means model and derive a bound on the degree of reflection bias.

  - Observed correlation between KG peer scores and 8th grade score places an upper bound on reflection bias of 20%.

  - Variance in scores implies attenuation bias of 20% as well, implying preceding estimates are downward biased on net.
Fade-out and Re-emergence: The Role of Non-Cognitive Skills

- Why do effects of kindergarten class fade out and re-emerge?

- One explanation: non-cognitive skills (Heckman 2000)

- Data on non-cognitive measures (effort, initiative, disruption) collected for random subset of STAR students in 4th and 8th grade

Please consider the behavior of Jim Smith over the last 2-3 months. Circle the number that indicates how often the child exhibits the behavior.

<table>
<thead>
<tr>
<th>#1. Acts restless, is often unable to sit still</th>
<th>Never</th>
<th>Sometimes</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#2. Annoys or interferes with peers’ work</th>
<th>Never</th>
<th>Sometimes</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3 4 5</td>
</tr>
</tbody>
</table>

- Convert mean non-cog score to percentile scale as above
<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Wage Earnings</th>
<th>Grade 8 Scores</th>
<th>Grade 4 Scores</th>
<th>Grade 8 Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grade 8</td>
<td>Math+ Read</td>
<td>Math+ Read</td>
<td>Math+ Read</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Grade 4 Non-Cog. Score</td>
<td>$87.7</td>
<td>0.059</td>
<td>0.047</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>($20.4)</td>
<td>(0.017)</td>
<td>(0.035)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Grade 4 Math + Reading Score</td>
<td>$36.4</td>
<td>0.671</td>
<td>0.153</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>($24.7)</td>
<td>(0.023)</td>
<td>(0.065)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Class Quality (peer scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,360</td>
<td>1,254</td>
<td>4,023</td>
<td>4,448</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1,671</td>
<td>1,780</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Non-cognitive skills provide a simple explanation of our findings

High quality KG teachers raise KG test scores partly through good classroom management

- Good classroom management instills social skills
- Social skills not directly measured in standardized tests but have returns in the labor market

→ Rapid fadeout in math and reading tests after KG

→ But significant earnings gains from better KG class
Part 6: Cost-Benefit Analysis

- Assume: 3% real discount rate, constant percent income gains, income follows average US income profile, constant effects of class quality

1. One SD increase in KG class quality for a single year
   \[ \rightarrow \text{Total NPV earnings gain for class of 20 students of $782K} \]

2. 33% reduction in class size
   \[ \rightarrow \text{$4K$-$189K per class (very imprecisely estimated)} \]

3. One SD improvement in teacher quality
   \[ \rightarrow \text{$170$-$214K per class} \]

- Moving from below-avg (25\textsuperscript{th} pctile) to above-avg (75\textsuperscript{th} pctile) teacher generates NPV of $320K for a class of 20 students
Intergenerational income correlation of around 0.3 (Solon 1999)

How much of this can be explained by the fact that higher income families have access to better public schools?

In STAR data, each $10K of parents’ income increases class quality in each grade by 0.7% of a SD

Use our estimates of effect of class quality on child’s earnings and assume constant class-quality effects across grades

Roughly 1/3 of intergenerational income transmission runs through differences in school quality in K-12
Appendix Table 1: Correlation of Earnings Over the Life Cycle

Correlation between Earnings at Age and Age + 6
Table 6: Observable Teacher vs. Peer Effects

<table>
<thead>
<tr>
<th>Dependent Var.</th>
<th>Test Score</th>
<th>Wage Earnings</th>
<th>Test Score</th>
<th>Wage Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Teacher with &gt;10 Years Experience</td>
<td>3.18%</td>
<td>$1093</td>
<td>1.61%</td>
<td>-</td>
</tr>
<tr>
<td>Teacher has post-BA deg.</td>
<td>-0.85%</td>
<td>-$261</td>
<td>0.95%</td>
<td>-</td>
</tr>
<tr>
<td>% Black Classmates</td>
<td></td>
<td>-1,757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Female Classmates</td>
<td></td>
<td>-67.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Free-Lunch Classmates</td>
<td></td>
<td>-285</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates’ Mean Age</td>
<td></td>
<td>-25.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classmates’ Mean Pred. Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry Grade</td>
<td>KG</td>
<td>KG</td>
<td>Grade ≥1</td>
<td>Grade ≥1</td>
</tr>
<tr>
<td>Observations</td>
<td>5,601</td>
<td>6,005</td>
<td>4,270</td>
<td>4,909</td>
</tr>
</tbody>
</table>

Note: All specifications control for school fixed effects and class size, as well as student and parent demographics.