CENTER FOR EDUCATION POLICY ANALYSIS at STANFORD UNIVERSITY

Using Student Test Scores to Measure Principal Performance

Susanna Loeb

Jason A. Grissom
Demetra Kalogrides

Research supported by Institute of Education Sciences Grant R305A100286
Instituto Nacional de Evaluación Educativa (INEE), March 21 2013

Motivation

- Substantial recent policy interest in using student test score data to evaluate the performance of school personnel
  - New York: Educational Law 3012-c (2010), 20-40%
  - Louisiana: House Bill 1033 (2010), 50%
  - Florida: Senate Bill 736 (2011), 50%
- A relatively large literature has focused on the issues surrounding the use of student growth models to measure teacher effects
- In contrast, very little research on using test scores to do the same for principals

Why Principals?

- Principals linked to teacher satisfaction and career choices (link)
- Principals central actors in most recent school reforms
  - Accountability
  - School-based budgeting
  - Mutual-consent agreements (teacher hiring)
  - Charter schools
- Some empirical evidence of effects on students
  - Students learn more in schools where
    - Principals spend more time on organizational management
    - Human resource practices are more conducive to hiring, retaining, assigning and developing good teachers.
- Increased policy attention on attracting and preparing effective school leaders

Aren't the principal issues like the teacher issues?

- In some ways, yes
  - Which test you use matters
  - Student sorting across schools can create bias if not well addressed
  - Test measurement error, sampling error, and other shocks introduces error in effect estimates

- But in some important ways, no
  - Principal effects dispersed over entire school, so the principal can affect a given student in more than one year
  - Indirect effects on students mediated by resources only partially under principal's control
  - For teachers, we use school fixed effects to combat sorting and control for school contextual factors—but only 1 principal per school

Goals of this study

1. Identify a range of possible value added-style models for capturing principal effects using student achievement data

2. Consider the conceptual issues associated with those models

3. Use longitudinal test score data to estimate the different models and examine their empirical properties

4. Compare the models' outcomes with non-test performance measures
   - Doesn't' show which is right but gives insights into validity of both test-score-based measures and other measures

How should we think about how principals affect school outcomes?

$$A_{ijs} = f(X_{ijs}, S(P_{js}, O_s, X_{ijs}))$$

- Student achievement is a function of his/her characteristics (*X*) and the effectiveness of the school for him/her (*S*), which is in turn a function of the performance of the school's principal (*P*), and other aspects of the school (*O*) that aren't under the principal's control
- How you measure principal effects depends on your beliefs about *S*
- Two main issues
  1. Timing: Do you expect that principal performance is reflected in student outcomes immediately (e.g., assigning teachers where they can be most effective, pushing everyone to work hard), or could good performance take time to show up (e.g., through teacher recruitment, changing the school culture)?
  2. School effects: How important is distinguishing *P* from *O*?

Approach 1: If principals' effects are primarily immediate and most relevant school factors are under their control

- Then principal effects are quite like teacher effects
- We would measure how much students learn while the principal is in the school, adjusting for student backgrounds
- Estimate a VA model with a principal-by-school effect

$$A_{ispt} = A_{is(t-1)}\beta_1 + X_{ispt}\beta_2 + S_{spt}\beta_3 + C_{spt}\beta_4 + \tau_y + \gamma_g + \delta_{sp} + \varepsilon_{ispt}$$

- Assumes that all differential growth in student learning from similar students in similar contexts is due to principal

Many alternative models with the same general idea

- Variation in how to separate principal/school effects from differences in student population
  - Could use additional prior scores
  - Could instrument prior scores, adjusting for measurement error
  - Could include student fixed effects
  - Could include non-linear school controls
  - Could include neighborhood characteristics
  - Could estimate in two stages
    - Big debate currently in US about teacher value-added.
    - In first stage control for school characteristics, save residuals in second stage estimate principal by school effects
- Also, could not control for immediate prior score (if early score is available)

Drawback of Approach 1

- Attributes the entire school to the principal
  - But this might not be fair—100% of school quality probably not attributable to current principal
    - A principal who steps into a new school inherits teachers hired by someone else (78% in our data)
    - Administrative team also (67%)

- Alternative: If principals' effects are primarily immediate but principals start with very different schools and these differences are not under their control

2. Relative School Effectiveness

- Potential solution: Compare the relative effectiveness of the school during the principal's tenure to the effectiveness of the same school at other times by adding a school fixed effect:

$$A_{ispt} = A_{is(t-1)}\beta_1 + X_{ispt}\beta_2 + S_{spt}\beta_3 + C_{spt}\beta_4 + \tau_y + \gamma_g + \phi_s + \delta_p + \varepsilon_{ispt}$$

  - Principal and School fixed effects
  - Identify school effects by differences in principals who move across schools
  - Identify principal effects by comparing principals serving in the same school

Similarly, many alternative specifications but the idea is the same

- Could make similar adjusts as in Approach 1
  - More prior test scores…
- Could run in two stages
  - First estimate school effects
  - Then estimate principals (or principal by school) effects from the residuals
  - Then principal effects would sum to one in a school
  - Allows for a principal by school effect
- Could control for prior school effectiveness instead of including a school fixed effect
  - Measurement bias but more variation (fewer small group comparisons).
  - Do this, but weak control means similar to first approach.

Drawbacks of Approach 2

- Need a long data stream so schools can have many principals
  - Not estimable when only 1 principal (in our data = 38% of schools) and no principal movement across schools
  - If only 1 additional principal, comparison is implicitly just to that person, which is difficult to justify (34%)
- Schools change over time, including changes in school composition and shocks to the culture.
  - Approach 2 may then still have omitted variables problems
  - Most used by researchers but, seems to be easily dismissed by practitioners
- Alternative to Approaches 1 and 2: If principals' effects accumulate over time

Approach 3: School Improvement

- Prior models do not explicitly incorporate long-run improvement
  - Building a productive work environment over time might be an important part of what a principal does
- Can incorporate improvement by estimating a principal-specific time trend in the school and taking its coefficient as improvement:

$$A_{ispt} = A_{is(t-1)}\beta_1 + X_{ispt}\beta_2 + S_{spt}\beta_3 + C_{spt}\beta_4 + \tau_y + \gamma_g + \delta_{sp} + \alpha_{sp}T_{spt} + \varepsilon_{ispt}$$

- Drawbacks
  - Can't estimate for principals who spend only one year in a school
  - Good face validity properties but substantial data requirements
  - Measurement error more of an issue—may be difficult to get reliable measures of improvement over time

The 3 Conceptual Approaches

- Approach 1: School Effectiveness during a principal's tenure
  - Assumes the effect is immediate and constant
  - Assumes principal controls all school effects
- Approach 2: Relative School Effectiveness to other principals serving in the same school
  - Assumes the effect is immediate and constant
  - Can only be estimated for some principals
  - Comparison to other principals may be idiosyncratic
  - Assumes no unobserved changes in schools that are outside the principal's control
- Approach 3: School Improvement
  - Measurement error
  - Need multiple years in a school

Prior Research Literature

- Studies have used student test score data to examine the impact of school leadership on schools
  - Often cross-section without proper controls
  - Some longitudinal studies – estimating effects of characteristics and processes
- Only 4 value-added studies
  - **Coelli & Green** (2012) only published paper: high school graduation and 12th grade final exam scores in British Columbia. No prior controls. Model like 2A, finds little effect. Model like 3A that allows growth: finds some effect of principals
  - **Branch, Hanushek, & Rivkin**, (2012): Model 1A (and a bit on 2A). Estimates the magnitude of the effect. About 0.05 standard deviations in math.
  - **Dhuey and Smith** (2012) elementary and middle schools in British Columbia. Model 2A. Standard deviation about 0.16 in math, 0.10 in reading
  - **Chiang, Lipscomb, & Gill** (2012). Alternative approach. Tries to separate the school effect from the principal effect using principal transitions.
- Small literature
  - 2A, model with principal and school fixed effects, most popular
  - Focus on separating principal from school effect
  - Little systematic discussion of the merits of different approaches
  - No comparison to other measures

Data for Value-Added Measures

- Longitudinal data on students and personnel in Miami-Dade County Public Schools (M-DCPS) for 2003-04 to 2010-11 (8 years)
  - 4th largest school district in U.S. (350,000 students)
  - 90% black or Hispanic, 60% FRL
  - Rich administrative files

- Test measure: Florida Comprehensive Assessment Test (FCAT) in math and reading, grades 3-10
  - Criterion-referenced, based on Sunshine State Standards
  - Standardized within grade and school year

Method

- Estimate models approximating each of the approaches
- Approach 1
  - **1A**: Basic Model with principal-by-school fixed effects
  - Many others as specification check
- Approach 2
  - **2A**: Basic model with school and principal fixed effects
  - **2B**: Basic model with controls for school effectiveness with other principals and principal fixed effects
  - Others as specification checks
- Approach 3
  - **3A**: Basic model with principal-by-school fixed effects and principal-by school time trends. Use principal-by-school time trend coefficients
    - only estimated if observe principals for 3 years in the same school
  - Others as specification checks

Basic Model

$$A_{ispt} = A_{is(t-1)}\beta_1 + X_{ispt}\beta_2 + S_{spt}\beta_3 + C_{spt}\beta_4 + \tau_y + \gamma_g + \delta_{sp} + \varepsilon_{ispt}$$

- X: Student Characteristics
  - whether the student qualifies for free or reduced priced lunch
  - whether currently classified as limited English proficient
  - whether repeating the grade in which they are currently enrolled
  - the number of days they missed school in a given year due to absence or suspension (lagged)
  - race
  - Gender
- C: Class characteristics
  - All of the above averages
  - Achievement
  - Standard deviation of achievement
- S: School characteristics
  - Like C but at the school level

Steps

1. Estimate principal effects
   - Create original
   - Shrink using Empirical Bayes approach (for some applications). Link
2. Coverage
3. Magnitude of principal effects
   - How important do principals look for student achievement using each method
4. Similarities among measures. Correlations
5. Consistency across subjects and schools. Correlations
6. Comparison to other measures
   - description to come

Results: Coverage

Table 2. Distribution of Principal Value-Added Estimates	
	Sample Size
Math	
1A (No Student FE)	781
2A (Principal & School FE)	484
2B (Principal FE, Control for Prior)	353
3A (Principal by School Time Trend)	263

- Approach 2 smaller because principal effects instead of principal by school effects
  - B smaller than A because have to adjust for prior effectiveness
- Approach 3 smaller because need 3 years of data

Network Sizes

- Need to compare principals to another group of principals and then set the groups equal to each other
  - For most models, choose to group elementary schools, middle schools, k-8 schools, high schools, and combination schools – 5 networks

Model 2A is setting a lot of small groups of principals equal to each other

	Math		Reading	
	# Networks	Av. Size	# Networks	Av. Size
1A	5	156	5	159
2A	145	3	144	3
2B	5	71	5	73
3A	5	53	5	45

Results: Magnitude

Table 2. Distribution of Principal Value-Added Estimates

	FE	EB	TRUE
Math			
1A (No Student FE)	0.109	0.095	0.105
2A (Principal & School FE)	0.064	0.058	0.059
2B (Principal FE, Control for Prior)	0.084	0.070	0.080
3A (Principal by School Time Trend)	0.058	0.032	0.050
Reading			
Model 1A (No Student FE)	0.083	0.069	0.077
Model 2A (Principal & School FE)	0.038	0.039	0.034
Model 2B (Principal FE, Control for Prior)	0.065	0.055	0.061
Model 3 (Principal by School Time Trend)	0.042	0.023	0.032

School Effectiveness almost 2X as big

Results: Similarities

Math, EB Estimates			
	1A	2A	2B
2A	0.45	1.00	
2B	0.58	0.60	1.00
3A	-0.05	0.16	0.09
Reading, EB Estimates			
	1A	2A	2B
2A	0.39	1.00	
2B	0.63	0.44	1.00
3A	-0.01	-0.04	0.14

Improvement not at all correlated with other measures

Controls for prior higher than effects (measurement error)

Results: Consistency

	Acoss subjects	Across Schools	
		Math	Reading
1A	**0.51**	**0.25**	**0.16**
2A	**0.61**	NA	NA
2B	**0.53**	NA	NA
3A	**0.44**	-0.39	-0.11

- Across subject
  - Subject to same shocks and omitted variables
- Only for 1A across schools
  - Could still be due to principals moving to similar schools

Comparing VA Estimates to Other Performance Measures

- Florida accountability grades (A-F)
- District's evaluation rating of principal (4-category from distinguished to below expectations)
- Student, staff, and parent climate survey scores ("Assign an overall grade to school…"). Administered by MDCPS.
  - 2004-2009 school year
  - 4 questions (first 3 Likert scale) collapsed to the school level
    - students are safe at this school
    - students are getting a good education at this school
    - the overall climate at this school is positive and helps students learn at this school.
    - Assign a letter grade (A–F) to their school that captures its overall performance.
  - Factor for each group

Comparing VA Estimates to Other Performance Measures

- Measures from our spring 2008 web-based surveys of principals and assistant principals
  - both asked about principal performance on a list of 42 areas of job tasks common to most principal positions. factored into 6 areas
  - principals' overall effectiveness, plus effectiveness in organizational management, given evidence of importance

- Measurement error in performance measures and no build in adjustments so, model with following regression.

$$O_{ps} = \beta_1 VA_{ps} + X_{ps}\beta_2 + S_{ps}\beta_3 + \varepsilon_{ispt}$$

  - Controls: principal race and gender, average school test scores (from first year first grade), percent white, percent black, percent suspended, and percent chronically absent

	Mean	Std Dev	N	
Average of Ratings Received From District (1-4)	3.54	0.51	659	Average While at School
Proportion of Years Received Highest Rating From District	0.59	0.41	659	Average While at School
School Accountability Grade, 0-4 Point Scale	2.79	1.19	755	Average While at School
School Climate Scale-Student Report	-0.16	0.99	775	Average While at School
School Climate Scale- Staff Report	-0.17	1.05	788	Average While at School
School Climate Scale- Parent Report	-0.21	1.05	781	Average While at School
AP Rating of Principal (Overall)	0.01	0.86	188	2008 Survey
AP Rating of Principal (Management)	0.02	0.91	236	2008 Survey
Principal Rating of Own Effectiveness (Overall)	-0.02	1.00	213	2008 Survey
Principal Rating of Own Effectiveness (Management)	-0.02	1.00	247	2008 Survey

Results: Comparison Summary

	District Eval	FL Acct'y Grade	Student Climate	Parent Climate	Staff Climate	AP Mang Rating	Principal Mang Self-Rating
1A	++	++	++	++	++	++	+
2A	+	++			+		
2B	++	++	++	++	++	+	
3A							

Summary

- Choice of model matters
- Need to be very thoughtful about how we model principal effects
  - School effect only: Attributes too much of the school effect to the current principal
  - Relative school effectiveness: May be biased by limited and unequal comparison groups. Assuming all groups are equal. Low coverage.
  - School improvement: Measurement error may produce very unreliable estimates. Low coverage.
- Simplest conceptual model has clearest relationships with non-test performance measures
  - it isn't clear whether that is because they are both biased

Implications

- Think carefully about the use of the measure
  - Is it for evaluation or research?
- Research
  - Cause is the most important factor
    - Worry about internal validity.
    - Not clear what the best approach is. Simulations might help
- Evaluation
  - Ehlert, Koedel, Parsons and Podgursky (2012) identify 3 key objectives for an evaluation measure:
    - elicit optimal effort from personnel
    - improve system-wide instruction by providing useful performance signals
    - avoid exacerbating pre-existing inequities in the labor markets between advantaged and disadvantaged schools
  - For this, basic school effectiveness measures (perhaps with more complete school characteristics adjustments) may be preferable if comparisons are made to similar schools.
- Just a start in understanding measures
  - Many possible measures and comparisons

CENTER FOR EDUCATION POLICY ANALYSIS at STANFORD UNIVERSITY

Using Student Test Scores to Measure Principal Performance

Susanna Loeb

Jason A. Grissom
Demetra Kalogrides

Research supported by Institute of Education Sciences Grant R305A100286
Instituto Nacional de Evaluación Educativa (INEE), March 21 2013

(perhaps and bit forced)

Transition

- My interest in principals, stemmed from their importance for teachers.
- In particular, relative strong evidence suggests that principals are key to teacher retention, and maybe to the differential retention of highly effective teachers
- But, surprisingly little literature on whether teacher retention matters for students…

CENTER FOR EDUCATION POLICY ANALYSIS at STANFORD UNIVERSITY

The Effect of Teacher Turnover On Student Achievement

Susanna Loeb

Matthew Ronfeldt and Jim Wyckoff

Background: Teacher Turnover

- Nationally in the US, about 30% of teachers leave the profession after 5 years
- A much higher proportion leave their school each year (approximately half a million teachers)
  - Only 16% of this teacher attrition at the school level can be attributed to retirement
  - The remaining 84% of the teacher turnover is due to teachers transferring between schools and teachers leaving the profession entirely
- About 20% in NYC
- Much higher in poorer schools

Effects of Turnover

- **Teacher turnover often assumed to harm student achievement**
  - Little empirical evidence for direct effect (Guin, 2004)
  - **Turnover rates are higher in lower-performing schools** (Guin, 2004; Hanushek et al. 1999) – Not Causal

- **Some turnover can be beneficial – new ideas, person-job match** (Organizational management lit, e.g. Abelson & Baysinger, 1984)

Possible Mechanisms

learning

$$L_{itcgsy} = f(\mu_{iy}, Q_{tcgsy})$$

the student

the quality of instruction

$$Q_{tcgsy} = g(C_{cgsy}, \tau_{ty}, \vartheta_{gsy})$$

classroom characteristics

who the teacher is

context/ supports

- **Context?**
  - e.g. time and effort to coordinate among teachers
  - e.g. time and effort needed to help other teachers
  - e.g. resources (other teachers' knowledge) available for teachers' own teaching

Possible Mechanisms

$$Q_{tcgsy} = g(C_{cgsy}, \tau_{ty}, \vartheta_{gsy})$$

TEACHER TURNOVER

- **Compositional** – turnover impact due to change in composition of teachers
- **Disruption / Innovation** – effects beyond changes in composition of teachers

"Compositional" Change

- **Effect depends on difference in effectiveness between "leavers" and replacements**
  - Beneficial when leavers less effective
- **More effective (e.g. higher VAM) teachers at least as or more likely to stay in teaching & in specific schools** (Boyd et al. 2007/2010; Goldhaber et al., 2007; Hanushek & Rivkin, 2010)
  - E.g. Miami estimates 19%, NYC link
  - Even in underserved schools
  - Calls into question turnover == harmful
- **Don't know whether teachers who filled vacancies were more effective than those they replaced, or whether worse teachers will continue to be poor teachers**
  - Net effect unknown

"Disruption" Change

- **Turnover can have a disruptive impact beyond changes in teacher composition**
- **Even when leavers == replacements, turnover may disrupt student learning**
  - Startup costs – recruiting, hiring, socializing new teachers
  - Lost relationships/collaborations (e.g. staff cohesion /community)
  - Lost institutional knowledge (e.g. instructional program coherence)
- **No research that we know of estimating this effect.**

Research Questions

- What is the average effect of teacher turnover on student achievement?
- Are the effects different for different kinds of schools? (e.g. low/high achieving)
- To what extent are the effects compositional or disruptive?

Data

- Data on NYC 4th and 5th grade students (NYCDOE; NY State Ed Dept)
  - Limit to these grades because have growth and teachers teach only one grade at a time
- Approximately 850,000 student observations over 8 years: 2001-2002, and 2005-2010
  - NYC 2003 data is low-quality
- Student test scores in math and ELA linked to student, class, school, and teacher characteristics

STUDENT CHARACTERISTICS	Mean
Proportion Female	0.51
Proportion Hispanic	0.37
Proportion Black	0.32
Proportion Asian	0.14
Proportion Free Lunch	0.64
Proportion Reduced Lunch	0.08
Proportion Home Language English	0.62
Proportion Suspended in Prior Year	0.01
Proportion Changing Schools from Prior Year	0.09
Average Number of Absences in Prior Year	10.66 (10.04)
Grade 4 Observations (Student-Year)	431,341
Grade 5 Observations (Student-Year)	432,765
Observations (Student-Year)	864,106

Methods

- **School-by-Grade-by-Year level turnover**
  - What happens to student achievement in year t as function of the percent of teachers in the students current grade that turned over between year t-1 and year t?
- **Two classes of fixed-effects regression models**
  - **Grade-by-School**: Leverage variation in turnover across years within the same grade level and school
  - **School-by-Year**: Capitalizes on turnover variation across grades within the same year and school

1st order potential biases

- Consider principal turnover or other school-level shock causing turnover:
  - Negative effect of shock would probably be in the prior year
  - Plus school-by-year effects should adjust for this (not school-by-grade)
- Consider teacher conflict in a given grade causing turnover
  - Negative effect of shock would probably be the prior year
  - School-by-grade effects should adjust for long-run within-grade conflict
- Consider teacher attrition because know next year will be bad.
  - If the whole school will be bad then school-by-year fixed effects should take care of this.
  - If just in their grade, problematic. (try an attempt to get around this but with too much noice)

Model

$$A_{itgsy} = \beta_0 + \beta_1 A_{itgs(y-1)}) + \beta_2 OtherA_{itgs(y-1)} + X_{itgsy}\beta_3 + C_{tgsy}\beta_4 + S_{sy}\beta_5 + \phi_y + \nu_{gs} + \beta_6 T_{gsy} + \varepsilon_{itgsy} \quad (1)$$

The test performance of individual, i, with teacher, t, in grade, g, in school, s, in time, y, is a function of:

- Prior scores in both math, $A_{itgs(y-1)}$, and ELA, $OtherA_{itgs(y-1)}$
- Student background characteristics, X_{itgsy}
- Time varying classroom, C_{tgsy}, and school characteristics, S_{sy}
- Grade-school-year turnover measure, $\boldsymbol{T_{gsy}}$
- Grade-by-school, ν_{gs}, and year fixed effects, ϕ_y
- Error term, ε_{itgsy}
- Clustered standard errors (level of turnover)

The alternative specification substitutes $\phi_y + \nu_{gs}$ for $\phi_g + \nu_{sy}$

Two Turnover Measures

- Want a measure at the grade-by-year level in order to use the fixed-effects approach
- Turnover not as straight forward to define as we initially thought

1. **Lagged Attrition**: Proportion of teachers in a given grade level in year t-1 who left the school by year t
2. **Proportion New to School-by-Grade**: Proportion of teachers in a given grade who were new to the school in year t
   - New to the school = first year or movers from other schools

- Note: not including grade switchers

Measuring Turnover When Teacher Population Changes

Hypothetical: Grade 4 in School A	Example	Turnover Rate Using Lagged Attrition (# who left in 04-05) /	Turnover Rate Using Proportion New (# new in 05-06) / 06)
Growth: Increase in Number of Teachers	200 200 (6 s		= 1/7 = 0.14
Decline: Decrease in Number of Teachers	200 200 (6 s		= 0/6 = 0
Constant: Number of Teachers is Constant	200 200 (5 stayers, 1 mover)		= 1/6 = 0.17

Also not the same when teachers change grade and they are replaced by a teacher who is new to the school:

2004-2005: 4 teachers in the grade
2005-2006: 2 teachers change grade and are replaced by new teachers

Lagged Attrition = 0

Proportion New = 1/2

TURNOVER DATA

TEACHER-YEAR CHARACTERISTICS	Mean
Experience	8.36 (7.16)
Proportion Stayers	0.86
Proportion Movers	0.04
Proportion First Years	0.09
Proportion Unknown Status	0.02
Observations (Teacher-Year)	42,170
GRADE-BY-YEAR-BY-SCHOOL CHARACTERISTICS	Mean
Teachers	4.80 (2.32)
Turnover Rate (Lagged Attrition)	0.11 (0.17)
Zero Lagged Attrition	0.58 (0.50)
Total Lagged Attrition	0.01 (0.08)
Turnover Rate (Proportion New to School)	0.13 (0.18)
Zero New to School	0.51 (0.50)
Total New to School	0.01 (0.09)
Observations (School-Grade-Year)	10,663

WHAT IS THE AVERAGE EFFECT OF TEACHER TURNOVER ON STUDENT ACHIEVEMENT?

Estimates of the Effects of Teacher Turnover on Student Achievement

Test	Turnover Measure	Model 1	Model 2
Math	Lagged Attrition	-0.074**	-0.074**
		(0.013)	(0.013)
	Proportion New To School	-0.096**	-0.093**
		(0.012)	(0.012)
ELA	Lagged Attrition	-0.060**	-0.064**
		(0.013)	(0.013)
	Proportion New To School	-0.083**	-0.082**
		(0.012)	(0.012)
	School-By-Year Fixed Effects	x	x
	Grade Indicators	x	x
	Student, Class, School Controls		x

Estimates of the Effects of Teacher Turnover on Student Achievement

Test	Turnover Measure	Model 1	Model 2
Math	Lagged Attrition	-0.086**	-0.082**
		(0.011)	(0.011)
	Proportion New To School	-0.102**	-0.096**
		(0.01)	(0.01)
ELA	Lagged Attrition	-0.049**	-0.049**
		(0.01)	(0.01)
	Proportion New To School	-0.060**	-0.051**
		(0.009)	(0.009)
	School-By-Grade Fixed Effects	x	x
	Year Indicators	x	x
	Student, Class, School Controls		x

Interpretation

- Student achievement is lower in years when turnover rates were higher
- Math scores are 7-9 percent of a standard deviation lower in years when there is 100 percent turnover (vs. no turnover)
  - ELA a bit smaller effect: 5-8 percent
  - Look by quartile of turnover to interpret more meaningfully …

Turnover Effect by Quartile - MATH

Lagged Attrition Q3	-0.009*	-0.008*
	(0.005)	(0.004)
Lagged Attrition Q4	-0.026**	-0.029**
	(0.005)	(0.004)
Proportion New Q3	-0.014**	-0.010*
	(0.005)	(0.004)
Proportion New Q4	-0.036**	-0.033**
	(0.005)	(0.004)
Student, Class, School Controls	x	x
School-by-Year Fixed Effects	x	
Grade Indicators	x	
School-by-Grade Fixed Effects		x
Year Indicators		x

Turnover Effect by Quartile - ELA

Lagged Attrition Q3	-0.012**	-0.012**
	(0.005)	(0.004)
Lagged Attrition Q4	-0.020**	-0.017**
	(0.005)	(0.004)
Proportion New Q3	-0.009*	-0.000
	(0.005)	(0.004)
Proportion New Q4	-0.029**	-0.016**
	(0.005)	(0.004)
Student, Class, School Controls	x	x
School-by-Year Fixed Effects	x	
Grade Indicators	x	
School-by-Grade Fixed Effects		x
Year Indicators		x

Interpretation: Magnitude

- Reducing attrition from 37% to none == 3 % of SD increase (math)
  - 6 teachers per grade level: From 2.2 leaving to 0

- Meaningful size? Roughly same magnitude of coefficient on free lunch eligibility

ARE THE EFFECTS DIFFERENT FOR DIFFERENT KINDS OF SCHOOLS?

Summary of Findings

- The negative effect of turnover on student achievement is larger in low-performing schools
  - though this varies across models
- Also larger in schools with more black students
- Not much difference between big and small schools
- Larger effects at old schools than new schools

WHAT EXPLAINS THE RELATIONSHIP BETWEEN TEACHER TURNOVER AND STUDENT ACHIEVEMENT?

Compositional (vs. Disruption) Explanations

- Changes in teacher experience?
- Being new to a school (mover)?
  - Being new difficult even for experienced teachers
- Teacher effectiveness (prior value-added)?
  - Prior value-added (average of prior estimates)
  - Reduced sample: 625,000 to 382,000 (math)
    - no first year teachers

Compositional (vs. Disruption) Explanations

- 2 approaches
  - Add controls for teacher characteristics to the model
    - Include experience of the current teacher
    - Include an indicator for whether the teacher is a mover
    - Include the prior value-added of the current teacher (much smaller sample – no first year teachers)
  - Run models separately just for stayers

Examining Whether Experience, Migration Explains Effects

Test	Turnover Measure	Model 1	Model 2	Model 3
Math	Lagged Attrition	-0.074**	-0.063**	-0.058**
		(0.013)	(0.013)	(0.013)
	Proportion New To School	-0.093**	-0.052**	-0.034**
		(0.012)	(0.012)	(0.013)
ELA	Lagged Attrition	-0.064**	-0.052**	-0.048**
		(0.013)	(0.013)	(0.013)
	Proportion New To School	-0.082**	-0.042**	-0.029*
		(0.012)	(0.012)	(0.012)
	School-By-Year Effects	x	x	x
	Grade Indicators	x	x	x
	Student, Class, School	x	x	x
	Experience Indicators		x	x
	Mover Indicator			x

Estimating Whether Prior Value-Added Explains Effects (reduced sample)

		All Schools		Low Achieving		High Achieving	
Test	Turnover Measure	Model 1	Model 2	Model 1	Model2	Model 1	Model 2
Math	Lagged Attrition	-0.055**	-0.034**	-0.063**	-0.049**	-0.045+	-0.016
		(0.015)	(0.013)	(0.019)	(0.017)	(0.025)	(0.021)
	Prop New To School	-0.045**	-0.033**	-0.101**	-0.075**	0.008	0.008
		(0.015)	(0.013)	(0.019)	(0.017)	(0.022)	(0.019)
ELA	Lagged Attrition	-0.036*	-0.034*	-0.042*	-0.041*	-0.031	-0.025
		(0.016)	(0.014)	(0.018)	(0.017)	(0.027)	(0.024)
	Prop New To School	-0.037*	-0.024+	-0.073**	-0.058**	-0.006	0.006
		(0.014)	(0.013)	(0.018)	(0.016)	(0.023)	(0.020)
	School-By-Year Effects	x	x	x	x	x	x
	Student, Class, School	x	x	x	x	x	x
	Average Prior VA		x		x		x

Compositional Effect?

- Some evidence for compositional mechanisms
  - Changes in experience, migration, and effectiveness explain much of the effect
  - Significant effect remains unexplained by compositional change

- To examine further – consider effect on students of "stayers"
  - In "compositional" mechanism stayers == bystanders; students unaffected
  - An effect is evidence for "disruption" mechanism

Just Stayers

	Math			ELA		
	All Schools	Low Perform	High Perform	All Schools	Low Perform	High Perform
School-by-Grade						
Lag Attr	-0.054**	-0.053**	-0.053**	-0.035**	-0.022	-0.059**
	(0.012)	(0.015)	(0.020)	(0.011)	(0.013)	(0.021)
Prop New	-0.030*	-0.058**	-0.000	0.008	-0.007	0.022
	(0.012)	(0.016)	(0.019)	(0.011)	(0.013)	(0.019)
School-by-Year						
Lag Att	-0.059**	-0.069**	-0.047+	-0.056**	-0.048**	-0.074**
	(0.014)	(0.017)	(0.024)	(0.014)	(0.017)	(0.025)
Prop New	-0.032*	-0.090**	0.026	-0.030*	-0.054**	-0.016
	(0.014)	(0.019)	(0.022)	(0.014)	(0.017)	(0.023)

Summary of Findings

- Teacher turnover, on average, has a negative effect on student achievement in ELA and math
- Effects strongest in schools with more low-performing and black students
- Teachers' experience, being new to a school, and effectiveness explain much (but not all) of the effect (compositional explanation)
- However, evidence for disruptive effect beyond changes in teacher composition
  - In low-performing schools, students of stayers do worse in years with more turnover

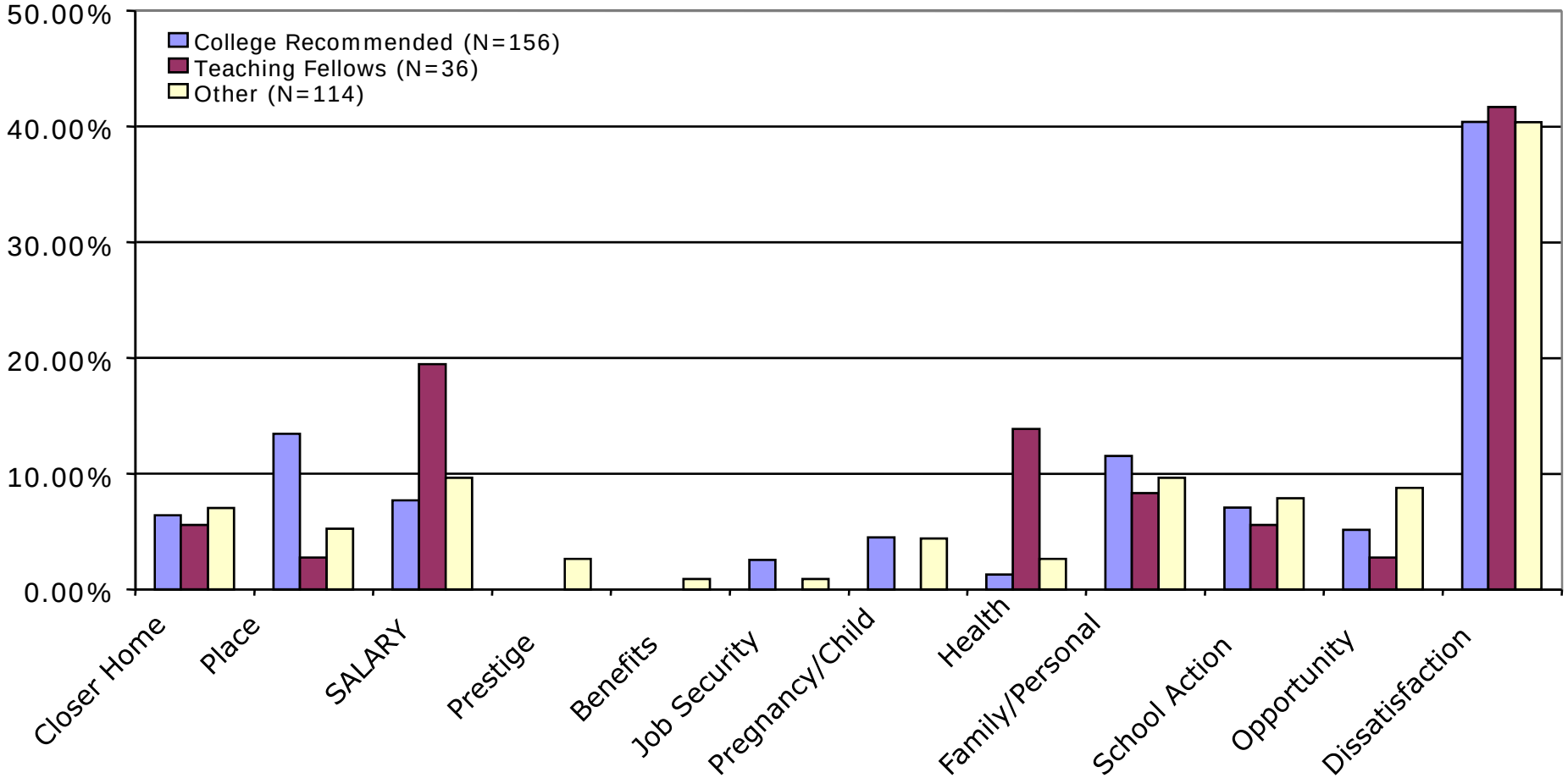
CENTER FOR EDUCATION POLICY ANALYSIS at STANFORD UNIVERSITY

The Effect of Teacher Turnover On Student Achievement

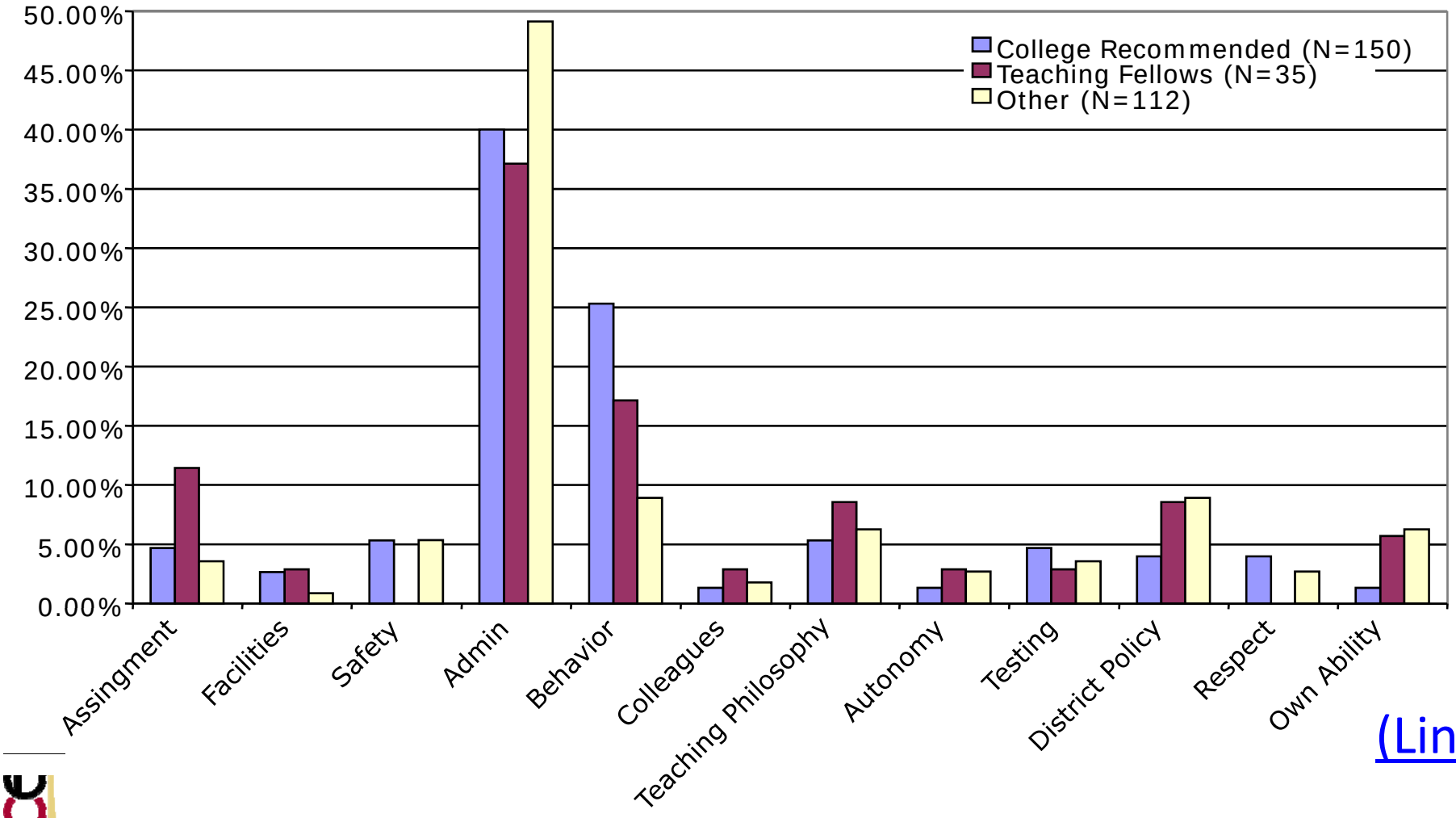
Susanna Loeb

Matthew Ronfeldt and Jim Wyckoff

Why Teachers Leave



When we look within school dissatisfaction…



(Link)

Empirical Bayes

- Our estimated principal effect is the sum of a "true" principal effect plus measurement error

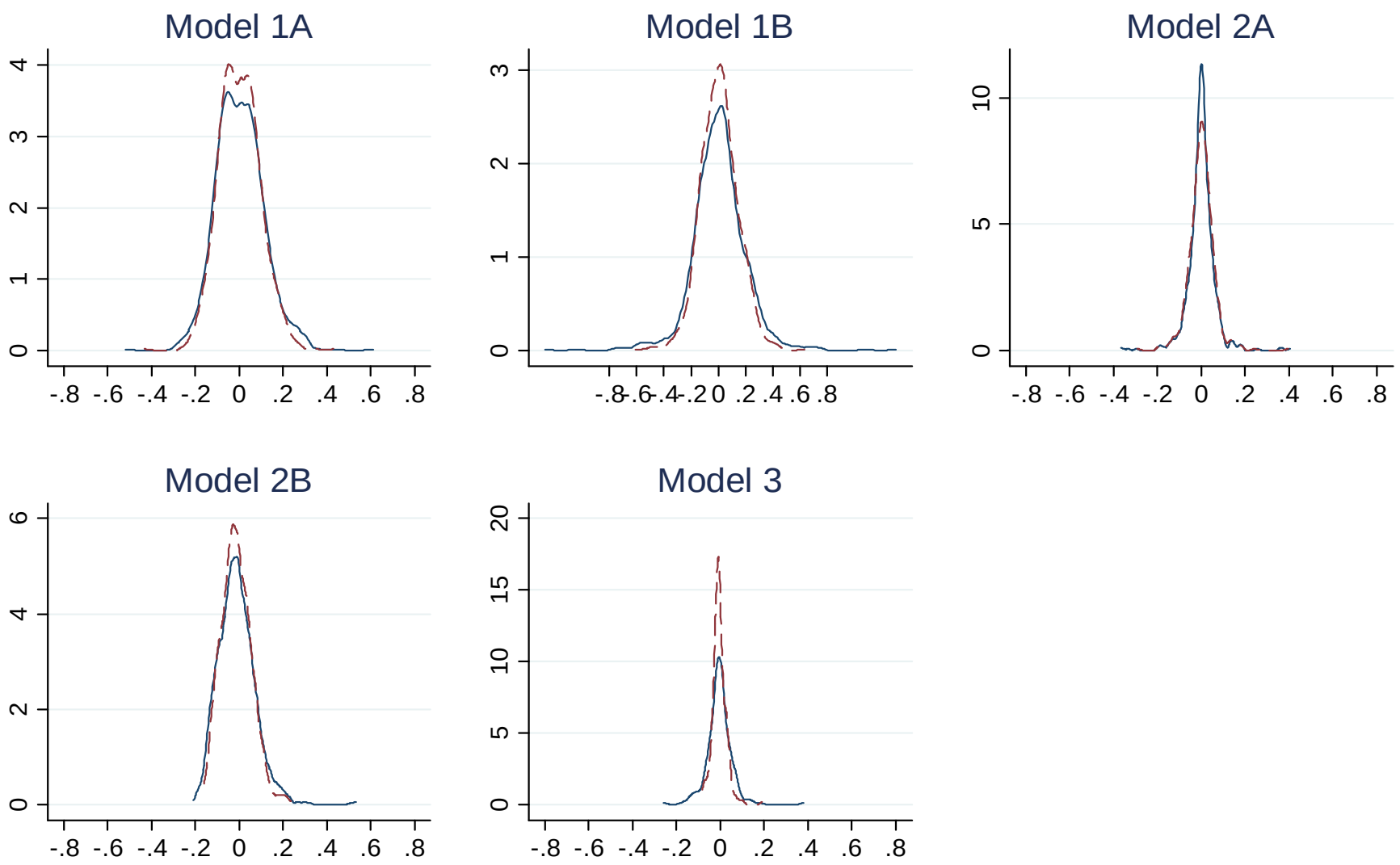
$$\hat{\delta}_{sp} = \delta_{sp} + \varepsilon_{sp}$$

$$E(\delta_{sp} \mid \hat{\delta}) = (1 - \lambda_{sp})(\bar{\delta}) + (\lambda_{sp}) * \hat{\delta}_{sp}$$

where $\lambda_{sp} = \dfrac{(\sigma_{\delta})^2}{(\sigma_{\varepsilon j})^2 + (\sigma_{\delta})^2}$ estimate with sample analogues

- the proportion of total variation in the teacher effects that is attributable to true differences between teachers

Link


Model 1A
Model 1B
Model 2A
Model 2B
Model 3
Estimated Effect
EB Effect

Link

Table 2: Average Within-School Differences in the Effectiveness of the Teachers Making Transitions Compared to Those Remaining in the Same School and Attrition Rates, First-Year Teachers in Grades 4-5

	NYC Transfer	NYS Transfer	Leave NYS
Relative Effectiveness in Math			
All Teachers	-0.026*	0.003	-0.041***
	(0.014)	(0.018)	(0.014)
Teachers in …			
Lower Scoring Schools	-0.004	0.010	-0.072***
	(0.021)	(0.039)	(0.023)
Middle group of schools	-0.025	0.005	-0.021
	(0.021)	(0.023)	(0.019)
Higher Scoring Schools	-0.072**	-0.008	-0.024
	(0.032)	(0.042)	(0.043)
Relative Effectiveness in English Language Arts (ELA)			
All Teachers	0.001	0.012	0.008
	(0.010)	(0.013)	(0.010)
Teachers in …			
Lower Scoring Schools	0.005	0.038	-0.012
	(0.016)	(0.029)	(0.017)
Middle group of schools	0.003	0.001	0.023
	(0.016)	(0.018)	(0.014)
Higher Scoring Schools	-0.007	0.019	0.013
	(0.024)	(0.031)	(0.032)

Link

	Principal Mobility				
	Same School	In District, New School	Out of District	Total Principals	First Year Principals
2003-2004	0.70	0.20	0.10	342	NA
2004-2005	0.66	0.25	0.10	346	55
2005-2006	0.67	0.23	0.10	356	75
2006-2007	0.64	0.28	0.08	365	69
2007-2008	0.63	0.28	0.08	379	76
2008-2009	0.60	0.29	0.11	398	68
2009-2010	0.67	0.27	0.06	413	70
2010-2011	NA	NA	NA	417	38

*NA indicates that we are unable to compute a given cell with the data that we have. The first year principal measure reflects the number of individuals who are principals in year t that were not principals in year t-1.

Number of Principals Per School in Our Data		
# of Principals	MDCPS	
	Freq.	Percent
1	166	37.64
2	131	29.71
3	92	20.86
4	38	8.62
5	12	2.72
6	2	0.45

*There is one observation per school and the figures represent how many different principals

Evaluations

	Average of Eval Ratings				% Yrs W/ Highest Rating			
	Math		Reading		Math		Reading	
1A	0.111	***	0.144	***	0.081	***	0.103	***
	(0.024)		(0.027)		(0.020)		(0.021)	
	637		641		637		641	
2A	0.067	*	0.027		0.040		0.012	
	(0.034)		(0.032)		(0.026)		(0.026)	
	390		377		390		377	
2B	0.132	***	0.128	***	0.085	**	0.081	**
	(0.038)		(0.038)		(0.030)		(0.029)	
	253		258		253		258	
3A	-0.009		-0.005		-0.017		-0.006	
	(0.032)		(0.029)		(0.029)		(0.026)	
	246		210		246		210	

	School Accnt Grade				Climate (Student Report)			
	Math		Reading		Math		Reading	
1A	0.325	***	0.300	***	0.153	***	0.082	**
	(0.026)		(0.027)		(0.030)		(0.029)	
2A	0.140	***	0.068	*	0.035		0.002	
	(0.033)		(0.033)		(0.034)		(0.033)	
2B	0.302	***	0.221	***	0.151	***	0.094	**
	(0.033)		(0.044)		(0.033)		(0.033)	
3A	-0.021		-0.008		-0.050		0.005	
	(0.032)		(0.032)		(0.031)		(0.031)	
	Climate (Staff Report)				**Climate (Parent Report)**			
1A	0.166	***	0.133	***	0.121	***	0.048	
	(0.034)		(0.035)		(0.033)		(0.036)	
2A	0.106	**	0.058		0.032		-0.023	
	(0.039)		(0.041)		(0.040)		(0.035)	
2B	0.160	***	0.120	*	0.145	***	0.107	**
	(0.047)		(0.051)		(0.042)		(0.038)	
3A	-0.032		-0.035		-0.051		0.013	
	(0.049)		(0.044)		(0.036)		(0.027)	

Prin and AP Evals

	AP Rating (Overall)				AP Rating (Mgmt)			
	Math		Reading		Math		Reading	
1A	0.189	+	0.073		0.160	+	0.161	+
	(0.100)		(0.097)		(0.084)		(0.093)	
2A	-0.009		-0.009		0.012		-0.014	
	(0.097)		(0.111)		(0.090)		(0.092)	
2B	0.183		0.180	+	0.178	+	0.114	
	(0.112)		(0.102)		(0.100)		(0.074)	
3A	-0.028		-0.001		-0.021		-0.003	
	(0.076)		(0.068)		(0.087)		(0.076)	
	Prin Rating (Overall)				**Prin Rating (Mgmt)**			
1A	0.173	+	0.093		0.163		0.240	**
	(0.094)		(0.100)		(0.099)		(0.087)	
2A	0.120		0.154		0.061		0.178	
	(0.113)		(0.123)		(0.102)		(0.117)	
2B	0.122		0.134		0.123		0.130	
	(0.132)		(0.116)		(0.133)		(0.119)	
3A	-0.143		-0.090		-0.121		-0.174	+
	(0.154)		(0.132)		(0.122)		(0.101)	

Estimates of Differential Effects of Turnover

Test	Turnover Measure	High Ach.	Low Ach.	Low Black	High Black
Math	Lagged Attrition	-0.060**	-0.085**	-0.047*	-0.095**
		(0.022)	(0.016)	(0.020)	(0.017)
	Proportion New	-0.055**	-0.130**	-0.048**	-0.129**
		(0.019)	(0.015)	(0.018)	(0.016)
ELA	Lagged Attrition	-0.058**	-0.072**	-0.037+	-0.087**
		(0.022)	(0.016)	(0.021)	(0.017)
	Proportion New	-0.058**	-0.112**	-0.043*	-0.114**
		(0.020)	(0.015)	(0.018)	(0.016)
	School-By-Year Effects	x	x	x	x
	Year Indicators	x	x	x	x
	Student, Class, School	x	x	x	x

Estimates of Differential Effects of Turnover

Test	Turnover Measure	High Ach.	Low Ach.	Low Black	High Black
Math	Lagged Attrition	-0.073**	-0.085**	-0.062**	-0.094**
		(0.018)	(0.014)	(0.018)	(0.014)
	Proportion New	-0.068**	-0.119**	-0.059**	-0.128**
		(0.016)	(0.013)	(0.016)	(0.014)
ELA	Lagged Attrition	-0.059**	-0.045**	-0.053**	-0.047**
		(0.019)	(0.012)	(0.018)	(0.013)
	Proportion New	-0.026	-0.073**	-0.018	-0.080**
		(0.016)	(0.011)	(0.015)	(0.012)
	School-By-Grade Effects	x	x	x	x
	Year Indicators	x	x	x	x
	Student, Class, School	x	x	x	x

Estimates of Differential Effects of Turnover

Test	Turnover Measure	New	Old	Small	Big
Math	Lagged Attrition	-0.101	-0.074**	-0.089**	-0.041+
		(0.067)	(0.013)	(0.016)	(0.022)
	Proportion New	-0.125*	-0.091**	-0.075**	-0.122**
		(0.063)	(0.013)	(0.014)	(0.022)
ELA	Lagged Attrition	-0.085	-0.063**	-0.063**	-0.062**
		(0.053)	(0.014)	(0.017)	(0.021)
	Proportion New	-0.054	-0.082**	-0.076**	-0.091**
		(0.058)	(0.012)	(0.015)	(0.021)
	School-By-Year Effects	x	x	x	x
	Year Indicators	x	x	x	x
	Student, Class, School	x	x	x	x

Estimates of Differential Effects of Turnover

Test	Turnover Measure	New	Old	Small	Big
Math	Lagged Attrition	-0.090*	-0.082**	-0.084**	-0.078**
		(0.045)	(0.012)	(0.014)	(0.019)
	Proportion New	-0.107*	-0.095**	-0.094**	-0.103**
		(0.043)	(0.011)	(0.013)	(0.018)
ELA	Lagged Attrition	-0.007	-0.051**	-0.052**	-0.046**
		(0.040)	(0.011)	(0.013)	(0.017)
	Proportion New	-0.033	-0.053**	-0.057**	-0.046**
		(0.036)	(0.010)	(0.012)	(0.016)
	School-By-Grade Effects	x	x	x	x
	Year Indicators	x	x	x	x
	Student, Class, School	x	x	x	x

Estimating Whether Prior Value-Added Explains Effects (much reduced sample)

- Measuring VA: Student Fixed-Effect Models

$$A_{itgsy} - A_{itgs(y-1)} = X_{itgsy}\beta_1 + C_{tgsy}\beta_2 + S_{sy}\beta_3 + \pi_g + \delta_y + \gamma_i + \tau_{jt} + \varepsilon_{itgsy}$$

Note: current work on VA that I'd be happy to talk about

- Similar with control model:

$$A_{itgsy} = \beta_4 A_{itgs(y-1)} + X_{itgsy}\beta_1 + C_{tgsy}\beta_2 + S_{sy}\beta_3 + \pi_g + \delta_y + \gamma_i + \tau_{jt} + \varepsilon_{itgsy}$$

cepa

Estimating Whether Prior Value-Added Explains Effects (reduced sample)

		All Schools		Low Achieving		High Achieving	
Test	Turnover Measure	Model 1	Model 2	Model 1	Model2	Model 1	Model 2
Math	Lagged Attrition	-0.059**	-0.040**	-0.067**	-0.049**	-0.044*	-0.024
		(0.013)	(0.012)	(0.016)	(0.015)	(0.021)	(0.020)
	Prop New To School	-0.048**	-0.034**	-0.082**	-0.065**	-0.015	-0.007
		(0.012)	(0.012)	(0.016)	(0.015)	(0.019)	(0.018)
ELA	Lagged Attrition	-0.033**	-0.027*	-0.035*	-0.028*	-0.031	-0.025
		(0.012)	(0.012)	(0.014)	(0.014)	(0.022)	(0.021)
	Prop New To School	-0.004	0.000	-0.027*	-0.023+	0.018	0.021
		(0.011)	(0.011)	(0.014)	(0.014)	(0.018)	(0.018)
	School-By-Grade	x	x	x	x	x	x
	Student, Class, School	x	x	x	x	x	x
	Average Prior VA		x		x		x

Examining Whether Experience, Migration Explains Effects

Test	Turnover Measure	Model 1	Model 2	Model 3
Math	Lagged Attrition	-0.082**	-0.065**	-0.059**
		(0.011)	(0.011)	(0.011)
	Proportion New To School	-0.096**	-0.055**	-0.036**
		(0.01)	(0.011)	(0.011)
ELA	Lagged Attrition	-0.049**	-0.035**	-0.031**
		(0.01)	(0.010)	(0.010)
	Proportion New To School	-0.051**	-0.012	0.000
		(0.009)	(0.010)	(0.010)
	School-By-Grade Effects	x	x	x
	Student, Class, School	x	x	x
	Experience Indicators		x	x
	Mover Indicator			x

Additional Checks – Not in Paper
1. Count grade leavers as leavers

	School-by-Grade FE		School-by-Year FE	
	Lag att	Prop new	Lag att	Prop new
	ELA			
Turnover	-0.041**	-0.037**	-0.093**	-0.089**
	(0.006)	(0.005)	(0.007)	(0.007)
N	1114855	1114855	1114855	1114855
adj. R-sq	0.573	0.573	0.576	0.576
	Math			
Turnover	-0.066**	-0.063**	-0.085**	-0.087**
	(0.006)	(0.006)	(0.007)	(0.007)
N	1114390	1114390	1114390	1114390
adj. R-sq	0.627	0.626	0.632	0.632

Additional Checks – Not in Paper
2. Instrument with Retirement

	Math		ELA	
	lagg att	prop new	lagg att	prop new
Percent 55	**0.029****	**0.022****	**0.029****	**0.022****
	(0.002)	**(0.002)**	**(0.002)**	**(0.002)**
Percent 62	**0.125****	**0.085****	**0.126****	**0.085****
	(0.004)	**(0.005)**	**(0.004)**	**(0.005)**
percent under 30	0.015**	-0.016**	0.015**	-0.016**
	(0.001)	(0.001)	(0.001)	(0.001)
percent over 50	0.024**	-0.013**	0.024**	-0.012**
	(0.001)	(0.002)	(0.001)	(0.002)
average age	0.000	0.001**	0.000	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)

Additional Checks – Not in Paper
2. Instrument with Retirement – nice idea

	Math		ELA	
	School, grade, and year FE		School, grade, and year FE	
	Lag att	Prop new	Lag att	Prop new
turnover	-0.099	-0.147	-0.181	-0.254
	(0.114)	(0.165)	(0.123)	(0.180)
N	1114390	1114390	1114855	1114855