American Economic Journal: Applied Economics 2009, 1:4, 85–108
http://www.aeaweb.org/articles.php?doi=10.1257/app.1.4.85

Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers[†]

By C. Kirabo Jackson and Elias Bruegmann*

Using longitudinal elementary school teacher and student data, we document that students have larger test score gains when their teachers experience improvements in the observable characteristics of their colleagues. Using within-school and within-teacher variation, we show that a teacher's students have larger achievement gains in math and reading when she has more effective colleagues (*based on estimated value-added from an out-of-sample pre-period*)*. Spillovers are strongest for less experienced teachers and persist over time, and historical peer quality explains away about 20 percent of the own-teacher effect, results that suggest peer learning.* (*JEL* I21, J24, J45)

Economists have long been concerned with human capital spillovers, given that these have strong implications for the optimal distribution of workers both within and across firms. When workers and their colleagues are complementary inputs in production, improvements in coworker quality may increase a worker's own productivity. There is evidence of such spillovers. Workers' wages are higher in firms with more educated coworkers (Harminder Battu, Clive R. Belfield, and Peter J. Sloane 2003), and wages for educated workers are higher in cities where the share of educated workers is higher (Enrico Moretti 2004b). Using direct measures of productivity, Pierre Azoulay, Joshua S. Graff Zivin, and Jialan Wang (2008) find that scientists have fewer grants and publications after a high-profile scientist leaves their institution. Peer quality may affect worker productivity, even if worker output is independent, by changing the social context. It has been documented that supermarket checkout workers work faster while in the line of sight of a high-productivity worker (Alexandre Mas and Moretti 2009), the productivity of berry pickers converges to the productivity of their close friends when those friends are present (Oriana Bandiera, Iwan Barankay, and Imran Rasul forthcoming), and the shirking of workers who move branches is positively correlated with the average shirking of their coworkers (Andrea Ichino and Giovanni Maggi 2000). However, Jonathan E. Guryan, Kory Kroft, and Matthew J. Notowidigdo (2009) find no evidence of peer

* Jackson: Department of Labor Economics, ILR School, Cornell University, 345 Ives Hall East, Ithaca, NY 14853-3901, (e-mail: ckj5@cornell.edu); Bruegmann: Cornerstone Research, 1000 El Camino Real, Suite 250, Menlo Park, CA 94025-4327 (e-mail: ebruegmann@cornerstone.com). We thank David Cutler, Li Han, Lawrence Katz, Andrew Oswald, Gauri Kartini Shastry, Kate Emans Sims, Daniel Tortorice, and participants at the Harvard Labor and Organizational Economics Lunches. Elias gratefully acknowledges financial support from the Graduate Society Dissertation Completion Fellowship of Harvard University. The order of the authors' names was determined by a coin flip.

† To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/app.1.4.85.

effects between randomly assigned golf partners in professional tournaments, suggesting the importance of context.

Although much research on empirical peer effects has focused on motivation and shirking, peer learning is an important mechanism. According to modern macroeconomic growth models (Robert E. Lucas, Jr. 1988; Paul M. Romer 1990), knowledgeable and skilled individuals increase the skill and knowledge of those with whom they interact, generating more ideas and faster macroeconomic growth. Despite the importance of peer knowledge spillovers for the personnel practices of firms and for the economy as a whole, there is little documented evidence of their existence.[1] Documenting peer learning is difficult because output may be produced jointly, there may be self-selection such that observed peer ability may be endogenous to unobserved ability, peer knowledge is difficult to observe, and unobserved factors could affect both output and peer quality.

We fill this gap in the literature, providing evidence of peer learning among teachers, using a unique longitudinal dataset of student test scores linked to teacher characteristics in North Carolina. Specifically, we test whether changes in a teacher's peers affect the test score growth of her own students, and we investigate possible mechanisms. Our empirical strategy is to estimate a student achievement value-added model with the inclusion of teacher peer attributes as covariates. To avoid the reflection problem (Charles F. Manski 1993), we use two measures of peer quality that are not determined by contact with peers: observable peer characteristics that change exogenously, such as experience and certification test scores; and unobservable peer quality based on teacher-specific, time-invariant, value-added estimates from pre-sample data. We ensure that spillovers are not driven by students having direct contact with their teacher's colleagues by focusing on elementary school students who only have one teacher for the entire year.[2] To ensure that we do not use changes in peer quality due to teacher self-selection, we identify the *changes* in the performance of a teacher's students that are correlated with *changes* in the composition of her peers *within the same school* by including teacher-school fixed effects. Lastly, we define a teacher's peers to be all other teachers at the same school with students in the same grade. This allows us to deal with the possibility that changes in the attributes of a teacher's peers could be correlated with changes in school attributes or school policies (i.e., a school decides to de-emphasize math and gets rid of its best math teacher) by including year fixed effects for each school. Because teachers may be affected by teachers in other grades, our narrow definition of peers will provide a *lower* bound on the estimate of the importance of peers. In our preferred specification, we identify the effect of peers by comparing the changes in the test scores of a teacher's students over time as her peers (and therefore the characteristics of her peers) change within the same school, while controlling for school-specific time shocks.

[1] There is evidence of learning-by-doing spillovers across firms in the same industry (Martin B. Zimmerman 1982; Douglas A. Irwin and Peter J. Klenow 1994; Rebecca Achee Thornton and Peter Thompson 2001).

[2] We also remove all classrooms with teacher aides or team teachers to further eliminate the possibility of direct contact between students and their own teacher's colleagues. Students in an alternative education program may be exposed to guidance counselors and special educators. This is not a problem, however, because none of these other teachers are used in our data to form the peer group.

We find that students perform better when their teachers' peers have better observable characteristics. In models that use teacher value-added (based on historical student achievement gains) as a measure of teacher quality, we find that students experience greater test score gains when their teachers' peers have higher mean estimated value-added in both math and reading. These effects are robust across a variety of specifications and to our two distinct measures of teacher peer quality. Despite the predictive power of a teacher's peers, a failure to account for *contemporaneous* peer quality has a negligible effect on the own-teacher effect. Although we are careful to control for a variety of possible confounding influences, we do not have random assignment of students to teachers or of teachers to peers. Because the possibility of spurious correlation remains, we present several specification and falsification tests. These indicate that our results are not driven by endogenous peer quality changes across grades within schools, or the nonrandom dynamic sorting of students into classrooms.

To help disentangle peer learning from other forms of spillovers, we test for empirical predictions that are most consistent with peer learning. We find that less experienced teachers who are still acquiring "on-the-job" skills are most sensitive to changes in peer quality, teachers with greater labor-market attachment are more sensitive to peer quality, both current and historical peer quality changes affect current student achievement, and historical peer quality explains away between 18 and 25 percent of the own-teacher effect. These findings are consistent with either direct learning from peers or what we refer to as peer-induced learning (learning induced by one's peers influencing one's decision to acquire work-related skills). This paper provides some of the first credible empirical evidence of learning associated with one's peers in the workplace.

This paper contributes to the nascent literature questioning the validity of standard value-added models by evaluating the assumption of no spillovers across teachers—a key identification assumption in teacher value-added models.[3] Also, the findings here should give pause to advocates of *individual-level* merit-based pay that relies on comparing teachers within schools because such pay schemes could reduce teachers' incentives to help their colleagues and could undermine peer learning.

The remainder of the paper is organized as follows. Section I presents the theoretical framework. Section II presents the identification strategy. Section III presents the data. Section IV presents our different measures of peer quality. Section V presents the results. Section VI presents specification and falsification tests. Section VII presents evidence supporting the learning hypothesis, and Section VIII concludes.

[3] Jesse Rothstein (2007) finds that value-added models may perform poorly in the presence of student tracking, such that future teachers have as much predictive power as current teachers in many standard value-added models. In contrast, Thomas J. Kane and Douglas O. Staiger (2008) use data from a random-assignment experiment and find that several nonexperimental specification estimates of teacher effectiveness have strong predictive power in an experimental setting where students are randomly assigned to teachers. They find that patterns of fade-out over time are very similar across experimental and nonexperimental settings. Cory Koedel (2008) tests for joint production among secondary school teachers but finds no evidence of cross-subject spillovers among high school teachers.

I. Theoretical Framework

We aim to observe how, and try to explain why, the performance of an individual teacher's students is affected by arguably exogenous changes in the quality of that teacher's peers.[4] In this section, we outline three potentially important sources of spillovers between teachers and outline a framework for thinking about learning between teachers.

A. *Joint Production and Shared Resources*

Even when teachers have direct contact only with their own students, they may affect the time and other resources available to their peers' students. Teachers may share duties outside the classroom that require time and effort, so better peers may reduce the burden of these shared tasks. Similarly, the resources that teachers get from the school may be affected by the activities of their colleagues. The direction of this effect is ambiguous because more effective teachers may be better at lobbying for shared resources, increasing the amount available for each teacher, or may take a greater share of the resources available to the grade. A joint production explanation should yield a very simple prediction that a teacher may be positively or negatively affected by the quality of her contemporaneous peers. Under such an explanation, there may be substantial response heterogeneity, reflecting the fact that particular types of teachers are likely to be given certain types of tasks. Another prediction is that all peer effects should be contemporaneous, such that they do not persist over time.

B. *Motivation and Effort*

A teacher's peers can also affect her classroom performance by changing her own teaching effort. The presence of good teachers may motivate colleagues through contagious enthusiasm or through embarrassment over the unfavorable direct performance comparison. Because overall school or grade performance may be used to evaluate schools, the introduction of a better teacher to the grade could make free-riding more attractive. However, Eugene Kandel and Edward P. Lazear (1992) suggest that peer pressure may force teachers to internalize their spillovers. If peer pressure is sufficiently strong, it could push teachers with better peers toward higher performance. A motivation or effort explanation will have ambiguous empirical predictions; however, the empirical work on such mechanisms in the workplace suggests that teachers are likely to perform better if they have better peers. A simple motivation or effort explanation implies that all peer effects should be contemporaneous.

[4] There is a large literature on peer effects for students (this includes Caroline Hoxby 2000; Bruce Sacerdote 2001; Joshua D. Angrist and Kevin Lang 2004; Hoxby and Gretchen Weingarth 2005; Victor Lavy and Analia Schlosser 2007). There is also a literature documenting the importance of social networks (this includes Esther Duflo and Emmanuel Saez 2003; Ron Laschever 2005; Alan T. Sorensen 2006; Dora L. Costa and Matthew E. Kahn 2007).

C. *Peer Learning*

Improvement in teacher effectiveness over time, particularly in the first few years of teaching, is a consistent finding in the literature. This finding suggests that on-the-job learning is very important for teachers. Therefore, we are interested in whether learning is a major avenue for the transmission of peer effects. We believe that learning has several important features that help distinguish it from other peer-effects explanations, and we examine these empirically:

- On average, one can learn more from better peers, so we should observe positive correlation between peer quality and own-student performance.
- Learning requires investment, so teachers with greater labor-force attachment and less experience (who have more years of teaching remaining and therefore have more years in which to benefit from investing in their teaching skills and learning from their peers) should be more likely to invest in learning and be more sensitive to peer quality.
- Learning is cumulative, so students should be affected by the composition of their teacher's past peers.
- Because teaching ability is a combination of innate ability and learned skills, historical peer quality should explain some of the own-teacher effect.

One can easily distinguish a simple motivation story or a simple shared-task story from a learning explanation by testing the empirical features listed above. However, although these patterns imply a learning explanation, they do not necessarily imply learning directly from one's peers. It is possible that having better peers allows teachers to spend less time on other shared tasks and more time learning how to be a better teacher. Also, it is possible that when teachers have good peers, they are motivated to be better teachers and therefore invest in learning how to be a better teacher. Both of these explanations are learning stories, but they entail peer-induced learning rather than direct learning from peers. Because understanding how teachers acquire human capital is important and relatively understudied, being able to distinguish any kind of learning from other explanations is useful. Because all learning explanations could yield the same empirical predictions, we are unable to distinguish a peer-learning story from the peer-induced learning story. However, we are able to test for peer-related learning (either through peers inducing a teacher to learn or through peers teaching their peers).

II. Identification Strategy

In our analysis, a teacher's peers are defined as those teachers in the same school who teach students in the same grade in the same year. As discussed later, excluding peers from other grades is crucial to our identification strategy because that allows us to control for school-specific time shocks that could affect student outcomes and teacher peer quality. Using variation in the quality of *all* of a teacher's potential peers (teachers in all grades in the school) could lead one to confound school shocks with changes in peer quality. This is clearly undesirable. Teachers are more likely to be affected by their peers in the same grade, but because teachers in one grade may be affected by teachers in other grades, our estimates, using own-grade teachers, will provide a *lower* bound of the full effect of peers. Because establishing that peer effects exist is of first-order importance, and quantifying the full effect is secondary, we focus only on that variation that is credibly exogenous to other changes (that is, variation in own-grade peer quality conditional on school-specific shocks).

To infer the effect of a teacher's peers on student test scores, we begin with our baseline specification, a value-added model augmented to include measures of teacher peer quality.

$$A_{it} = \delta A_{it-1} + \varphi \mathbf{X}_{it} + \eta \mathbf{W}_{jt} + \mu \bar{P}_{j't} + \xi_{gt} + \varepsilon_{ijgst}. \tag{1}$$

In (1), we simplify the notation so that A_{it} represents A_{ijgst}, which is the achievement score of student i with teacher j in grade g of school s in year t. Similarly, A_{it-1} represents $A_{ij_{t-1}g_{t-1}s_{t-1}t-1}$, which is the score of student i with teacher j_{t-1} in grade g_{t-1} of school s_{t-1} in the previous year. $\mathbf{X}_{it}$ is a vector of student characteristics such as ethnicity, gender, and parental education level; $\mathbf{W}_{jt}$ is a vector of characteristics of teacher j in year t; ξ_{gt} is a grade-by-year fixed effect; and ε_{ijgst} is the idiosyncratic error term. $\bar{P}_{j't}$ is a measure of teacher peer quality. We discuss our measures of peer quality in Section IV.

One of the major problems in identifying credible peer-effect estimates is the fact that individuals often self-select to their peers. To avoid bias due to self-selection to peers, we remove all potentially problematic variation in peer characteristics that occurs as a result of the teacher's own movement by adding a teacher-school fixed effect to (1). As such, we identify our effects based on *changes* in the characteristics of a teacher's peers and *changes* in the performance of her students, when the teacher has remained at the same school over time. By relying only on variation within the scores of students of a given teacher within a given school, all variation in peer quality comes from either a teacher being reassigned to another grade within the same school, or the movement of peers into or out of her school and grade.

Another major difficulty in identifying peer effects, particularly where individuals are not randomly assigned to peers, is that changes in peer quality may be correlated with omitted factors that also affect own outcomes. For example, a disruptive event, such as a hurricane, could cause good teachers to leave the school at the same time that students perform poorly. Any school-specific shock that has a deleterious effect on both peer quality and student achievement would produce results that look like positive peer effects. To address this concern, we make comparisons only within groups of teachers at the same school at the same time (i.e., teachers who are subject to the same school-level shocks but teach in different grades and therefore have different peers). We do this by also adding a school-by-year fixed effect to (1). The school-by-year effect removes those confounding factors that affect all grades in the school that could also have an effect on the peer quality of teachers. Because peer quality for each teacher in a particular school is identified at the school-grade-year level, we cannot include school-grade-year effects. With the inclusion of school-year, grade-year, and school-grade fixed effects, our estimates will be biased only in the unlikely event that higher quality teachers are added to grades within schools at the same time as other improvements are made that are particular to that grade within the school. We present evidence that this was not the case in Section VI.

Our preferred model is, therefore, an augmented version of the student achievement model in equation (1) that includes teacher peer quality as an input, while also controlling for teacher-school fixed effects and school-by-year fixed effects. Specifically, we estimate (2) below.

$$A_{it} = \delta A_{it-1} + \varphi \mathbf{X}_{it} + \xi_{js} + \mu \overline{P}_{j't} + \xi_{gt} + \xi_{st} + \varepsilon_{ijgst}. \tag{2}$$

All variables are as before, ξ_{js} is a teacher-school fixed effect, and ξ_{st} is a year fixed effect for each school. Although it is tempting to include as many fixed effects as possible to remove confounding factors, such an approach often leads to weak identification, undermining the overall objective of identifying the parameter of interest (William Anderson and Martin T. Wells 2008). Although our preferred specification includes teacher-by-school fixed effects and school-by-year fixed effects, to show that our results are robust across a variety of empirical specifications, we report results from a series of regressions with the same basic specification described in (1) but with different sets of fixed effects.

III. Data

We use data on all third-grade through fifth-grade students in North Carolina between 1995 and 2006 from the North Carolina Education Research Data Center.[5] Our student data include demographic characteristics, standardized test scores in math and reading, and codes allowing us to link the data to information about the schools the students attend and the teachers who administered their tests. We use *changes* in student test scores as the dependent variable, so our regression analysis is based on the outcomes of fourth and fifth graders. According to state regulation, the tests must be administered by a teacher, principal, or guidance counselor. Discussions with education officials in North Carolina indicate that tests are always administered by the students' own teachers when these teachers are present. Also, all students in the same grade take the exam at the same time. Thus, any teacher teaching a given subject in a given grade will almost certainly be administering the exam only to her own students. This precludes our misspecifying a teacher as her own peer. We take several steps to limit our sample to teachers who we are confident are the students' actual teachers. We include only students who are being administered the exam by a teacher who teaches math and reading to students in that grade, and we remove teachers who are co-teaching or have a teaching aide. This process

[5] These data have been used by other researchers in different contexts to look at the effect of teachers on student outcomes (Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor 2005, 2007; Clotfelter et al. 2007; Rothstein 2007), the effect of schools on student achievement (Justine S. Hastings, Richard Van Weelden, and Jeffrey Weinstein 2007; Hastings and Weinstein 2007), the effect of student demographics on teacher quality (Jackson 2009), and the effect of schools on housing prices (Kane, Stephanie K. Riegg, and Staiger 2006).

Table 1—Summary Statistics

Variable	Observations	Mean	SD
Unit of observation: student-year			
Math scores	1,361,473	0.033	0.984
Reading scores	1,355,313	0.022	0.984
Change in math score	1,258,483	0.006	0.583
Change in reading score	1,250,179	0.001	0.613
Black	1,372,098	0.295	0.456
White	1,372,098	0.621	0.485
Female	1,372,098	0.493	0.500
Parent education: no high school degree	1,372,098	0.107	0.309
Parent education: high school degree	1,372,098	0.428	0.495
Parent education: some college	1,372,098	0.315	0.464
Parent education: college degree	1,372,098	0.143	0.350
Same race	1,372,098	0.649	0.477
Same sex	1,372,098	0.496	0.500
Class size	1,372,098	23.054	4.053
Unit of observation: teacher-year			
Experience	91,243	12.798	9.949
Experience 0	92,511	0.063	0.242
Experience 1–3	92,511	0.165	0.371
Experience 4–9	92,511	0.230	0.421
Experience 10–24	92,511	0.365	0.481
Experience 25+	92,511	0.164	0.371
Teacher exam score	92,511	–0.012	0.812
Advanced degree	92,511	0.197	0.398
Regular licensure	92,511	0.670	0.470
Certified	92,511	0.039	0.194
Peer experience 0	85,490	0.064	0.164
Peer experience 1–3	85,490	0.166	0.255
Peer experience 4–9	85,490	0.230	0.289
Peer experience 10–24	85,490	0.364	0.334
Peer experience 25+	85,490	0.164	0.256
Peer teacher exam score	85,490	–0.009	0.578
Peer advanced degree	85,490	0.198	0.274
Peer regular licensure	85,490	0.676	0.426
Peer certification	85,490	0.039	0.140

Notes: The few teachers with more than 50 years of experience are coded as having 50 years of experience. Sample size differs between the years of experience variable and the experience category variables because teachers with missing experience data are not included in the years of experience mean.

gives us roughly 1.37 million student-year observations matched to teachers we are confident taught the students the material being tested. Summary statistics for our data are presented in Table 1.

The students are roughly 62 percent white and 29.5 percent black, and are evenly divided between boys and girls (similar to the full state sample). About 65 percent of students are the same race as their teacher, and about 50 percent are the same sex as their teacher. The average class size is 23, with a standard deviation of 4. About 11 percent of students' parents did not finish high school, 43 percent received a high school diploma, roughly 30 percent had post-high school education but no four-year college degree, and roughly 14 percent of the students had parents who have a four-year college degree or graduate degree as their highest level of education. The test scores for reading and math have been standardized to have a mean of zero and unit variance, based on *all* students in that grade in that year. The average year-to-year test score growth is 0, with standard deviation of 0.583 for math and 0.613 for reading. Students in our sample attend a total of 1,545 schools. Schools, on average, had 101 students and 6.6 teachers.

About 92 percent of teachers we successfully match to students are female, 83 percent are white, and 15 percent are black. The average teacher in our data has 13 years of experience, and roughly 6 percent of the teachers have no experience.[6] About 20 percent of teachers have advanced degrees. The variable "regular licensure" refers to whether the teacher has received a regular state license or, instead, is working under a provisional, temporary, emergency, or lateral entry license. About 67 percent of the teachers in our sample have regular licensure. We normalize scores on the elementary education or the early childhood education tests that all North Carolina elementary school teachers are required to take, so that these scores have a mean of zero and unit variance for each year in the data. Teachers in our sample perform near the mean, with a standard deviation of 0.81. Lastly, about 4 percent of teachers have national board certification.

For part of our analysis, we use the mean characteristics of the other teachers in the same school and grade to indicate peer quality. Table 1 includes summary statistics for these measures. The variation we exploit comes from the movement of peers into or out of a school grade, so we look at several summary statistics to get a better understanding of this process in our data. First, we consider the distribution of peer group size. The average teacher in our data has about three other teachers in the same school grade and year that appear in our data. More than 90 percent of teachers have six or fewer colleagues in our data. The small number of teachers per school grade suggests that the relevant quality of peers in a teacher's own grade may change substantially with the introduction or exit of just one or two good or bad teachers. During the years 2001–2006, 65.8 percent of teachers are in the same school and grade as the most recent previous year in which they appeared in the data (going as far back as 1996), 6 percent are in the same school but teaching a different grade, 7.4 percent have moved from another school in our data since the most recent previous year, and 20.9 percent do not appear previously in our data. These high levels of mobility aid our identification.

We are also interested in which teachers are moving between grades and schools, so we compared the observable characteristics of teachers who are in the same grade and school as the previous year, the same school but a different grade as the previous year, a different school than in the previous year, and new to the data. The characteristics of these groups of teachers are quite similar (with the obvious exception of experience for teachers new to the data) suggesting that teachers who move between schools or grades are similar to teachers who do not. To see if mobile teachers moved to schools and grades with systematically better or worse peers, we

[6] Teacher experience is based on the amount of experience credited to the teacher for the purposes of determining salary; therefore, it should reflect total teaching experience in any school district.

computed the difference between each teacher's own characteristics and the average of her new peers' characteristics. Teachers who move from a different grade in the same school differ from their peers only in that they are slightly more likely to have regular licensure. Teachers moving between schools are more likely to have advanced degrees and regular licensure than their new peers. However, both these differences are economically small. These comparisons suggest that teachers who change schools or grades are similar to their new colleagues.[7]

IV. Measures of Teacher Peer Quality

A naïve empirical strategy to test whether teachers exert spillover effects on each other's students would be to estimate standard student value-added regressions with the inclusion of the mean test score growth of students of a teacher's peers. We do not pursue this strategy because the performance of students of a teacher's peers is, itself, a function of the teacher's own attributes. We address this problem with two different measures of peer quality that are not codetermined with a teacher's own performance. The first approach is to use the observable characteristics of peer teachers, and the second is to use the value-added of peer teachers estimated in an *out-of-sample* pre-period. The two different approaches complement and provide a robustness check on, each other. In both approaches, our models identify the social interaction effect, which is a combination of the effect of group characteristics on individual outcomes and the effect of group behavior on individual behavior (Manski 1993).

A. *Observable Characteristics as a Measure of Quality*

For the first proxy for peer quality, we compute the average characteristics for the peers of each teacher. For each school-year-grade cell, we compute the mean attributes of all other teachers in that cell, so that peer quality for teacher j in grade g at school s in year t, $\overline{\mathbf{W}}_{j'gst}$, is the mean characteristic of all *other* teachers j' in grade g at school s in year t. These peer averages are summarized in Table 1. We include these peer averages as a measure of peer quality $\overline{P}_{j't}$ in equation (2). Changes in this measure of peer quality occur when the characteristics of a teacher's peers change (e.g., becoming more experienced or obtaining regular licensure) or when the identity of a teacher's peers change. Because observable teacher characteristics, such as experience, vary exogenously with time, and because teachers are unlikely to obtain certification as a result of their peers, this approach is unlikely to be subject to the reflection problem. Our second measure of peer quality, however, relies solely on changes in the identity of a teacher's peers. The first approach has the advantage of being straightforward and allowing us to include data on almost all teachers, but, as in previous research, we find these characteristics are weak predictors of teacher quality. For this reason, we prefer our second approach for most of our analysis.

[7] The characteristics of mobile teachers and their new peers are summarized in Table A2 in the Web Appendix.

B. *Estimated Value-Added as a Measure of Quality*

Our main proxy for teacher peer quality is the historical estimated value-added of a teacher's peers. Because a teacher's value-added could be due to exposure to high-ability peers, it is important to identify variation in peer quality (as measured by value-added) that is not subject to spillover bias in the estimation equation. We address this problem by using *out-of-sample* estimates of teacher value-added based on data between 1995 and 2000, while estimating the effect of *changes* in estimated peer value-added on *changes* in own-student outcomes using data from 2001 through 2006. Using *changes* in peer quality addresses the concern that the *level* of a teacher's peer quality could have been affected by her own quality in the pre-sample period.

Using pre-sample (1995–2000) data, we estimate teacher value-added by estimating a student achievement model of the form (1) with the inclusion of indicator variables denoting if the student i is in class with teacher j (for each teacher). A detailed description of the value-added estimation, including the estimation equation and the results, is included in the Web Appendix. The coefficients on the teacher indicator variables, the θ_j's, are standardized and normalized, and are used as measures of teacher quality in the estimation sample (2001–2006 data). As with the observable teacher characteristics, peer quality for teacher j in grade g at school s in year t,$\overline{\theta}_{j'gst}$, is the mean estimated value-added of all *other* teachers j' in grade g at school s in year t. These estimated teacher value-added effects do not vary over time, so all of the variation in mean peer value-added comes from changes in the identity of a teacher's peers and, as such, is not subject to the reflection problem.

This value-added approach has the disadvantage that teachers who are not in-sample between 1996 and 2000 will have no estimated value-added. New teachers, teachers from out of state, and nonelementary school teachers will have no estimated value-added in our estimation sample (2001–2006). Because we would like to include all teachers in our estimation sample, and would like to use all of a teacher's peers, we use the full sample of teachers, and we assign the mean of the distribution to teachers with no estimated teacher value-added, as well as including control variables for the proportion of a teacher's peers with no estimated value-added. The proportion of teachers in a teacher's peer group with no estimated value-added serves as a proxy for the characteristics of teachers with missing peers. To ensure that our treatment of teachers with missing value-added does not drive our results, we estimated models that include dummies for having missing peers, use imputed teacher value-added based on observable characteristics for those teachers with missing teacher effects, include the number of new teachers to the grade in a given year, and use the mean only of those teachers with estimated value-added. Across all these models, the results are virtually unchanged.[8]

[8] Using the mean only for those teachers with estimated value-added results in peer effects that are 14 percent smaller in math and 4 percent smaller in reading. Because ignoring teachers without value-added introduces additional measurement error, a reduction in the estimated effect is expected. In practice, the reductions are small.

Table 2—Effect of Observable Teacher Peer Quality on Student Test Scores

	School fixed effects	Student-school fixed effects	Teacher-school and school-year fixed effects	School fixed effects	Student-school fixed effects	Teacher-school and school-year fixed effects
	Dependent variable: math test score			Dependent variable: reading test score		
	OLS	OLS	OLS	OLS	OLS	OLS
Model	(1)	(2)	(3)	(4)	(5)	(6)
Lagged score	0.7674	—	0.7658	0.739	—	0.7332
	(0.0021)***	—	(0.0018)***	(0.0016)***	—	(0.0016)***
Experience 1–3	0.0651	0.1005	0.0547	0.0408	0.0616	0.0324
	(0.0045)***	(0.0082)***	(0.0041)***	(0.0035)***	(0.0070)***	(0.0038)***
Experience 4–9	0.0816	0.1215	0.0683	0.0547	0.0743	0.0323
	(0.0046)***	(0.0080)***	(0.0054)***	(0.0035)***	(0.0069)***	(0.0050)***
Experience 10–24	0.0997	0.1383	0.0747	0.0754	0.0967	0.0377
	(0.0045)***	(0.0078)***	(0.0070)***	(0.0035)***	(0.0066)***	(0.0064)***
Experience 25+	0.1025	0.1368	0.0616	0.0835	0.1008	0.0295
	(0.0048)***	(0.0084)***	(0.0088)***	(0.0037)***	(0.0071)***	(0.0080)***
Peer experience 1–3	0.0248	0.042	0.0288	0.0071	0.0204	0.017
	(0.0071)***	(0.0142)***	(0.0069)***	(0.0055)	(0.0117)*	(0.0064)***
Peer experience 4–9	0.0193	0.0363	0.0264	0.006	0.0153	0.0132
	(0.0073)***	(0.0145)**	(0.0074)***	(0.0056)	(0.0120)	(0.0068)*
Peer experience 10–24	0.0247	0.0442	0.0218	0.0161	0.0303	0.0294
	(0.0072)***	(0.0142)***	(0.0075)***	(0.0055)***	(0.0117)***	(0.0069)***
Peer experience 25+	0.0238	0.0383	0.0209	0.0145	0.0259	0.0154
	(0.0078)***	(0.0152)**	(0.0083)**	(0.0059)**	(0.0125)**	(0.0075)**
Licensure score	0.0172	0.0179	—	0.0043	0.0018	—
	(0.0012)***	(0.0031)***	—	(0.0009)***	(0.0022)	—
Advanced degree	–0.0057	–0.0018	—	–0.004	–0.0015	—
	(0.0024)**	(0.0073)	—	(0.0019)**	(0.0050)	—
Regular licensure	0.0375	0.0583	—	0.0215	0.0324	—
	(0.0041)***	(0.0084)***	—	(0.0032)***	(0.0068)***	—
Certified	0.0347	0.0477	—	0.0156	0.0207	—
	(0.0046)***	(0.0111)***	—	(0.0035)***	(0.0081)**	—
Peer licensure score	0.0007	0.0027	0.0034	–0.0008	0.0013	–0.0022
	(0.0020)	(0.0037)	(0.0024)	(0.0015)	(0.0030)	(0.0022)
Peer advanced degree share	0.0031	0.0016	0.0049	–0.0038	–0.0025	–0.0099
	(0.0038)	(0.0074)	(0.0043)	(0.0029)	(0.0061)	(0.0040)**
Peer regular licensure share	0.0092	0.0112	–0.0096	0.0113	0.0128	–0.0024
	(0.0064)	(0.0124)	(0.0066)	(0.0050)**	(0.0103)	(0.0060)
Peer certification share	0.0126	0.0355	0.0025	0.0017	0.0191	–0.008
	(0.0069)*	(0.0133)***	(0.0076)	(0.0054)	(0.0111)*	(0.0068)
Observations	1,200,633	1,200,633	1,200,633	1,192,896	1,192,896	1,192,896
R^2	0.16	0.5		0.16	0.49	

Notes: All models include indicator variables for the gender and racial matches between the teacher and the students, class size, and grade-by-year fixed effects. All regressions include student demographic control variables, except models that include student fixed effects. All regressions include teacher control variables, except models that include teacher fixed effects (teacher experience included in all models). The omitted teacher experience group is teachers with zero years of experience. All regressions include an indicator variable for having missing experience data, and control for the proportion of peers with missing experience data. Robust standard errors clustered by teacher-school in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

V. Results

First, we consider the effect of the average of teachers' peers' observable characteristics on teachers' own performance (i.e., estimating equation (2) while using observable peer characteristics as our measure of peer quality). Table 2 presents these results. We report the results for math and reading test scores in the left and right panels, respectively. Although we focus on our preferred models, we present results obtained with school fixed effects, student fixed effects, and including teacher-by-school and school-by-year effects. The effects of own-teacher characteristics across all models are reasonable for math and reading. Students have higher test scores in both subjects when their own teacher has a regular teaching license, has higher scores on her license exam, is national board certified, and has more years of experience. Having a teacher with no previous experience is particularly detrimental, and having a teacher with an advanced degree appears to be *negatively* correlated with test scores, conditional on the other covariates.

We now turn our attention to the effect of peer characteristics. We focus on the results for our preferred model in columns 3 and 6. In this specification, for math and reading, the coefficients on all the peer experience categories are positive and statistically significant. Because the omitted variable is the share of peers with no years of experience, this indicates that having more peers with more than one year of teaching experience has a statistically significant positive effect. The differences between other experience categories are smaller and generally not statistically significant. Average peer licensure score and the share of peers with advanced degrees have small and statistically insignificant coefficients for math and reading. One can reject the joint hypothesis that the coefficients of the teacher-peer characteristics are equal to 0 at the 1 percent level for math and reading. Looking to specific characteristics, one can reject the joint hypothesis that teacher-peer experience coefficients are equal to 0 at the 10 percent level for math and reading. One cannot reject, however, the joint hypothesis that coefficients for teacher peer characteristics, other than peer experience, are equal to zero at traditional levels for either math or reading.

To summarize the effect of observable peer quality, we compute the value-added associated with a teacher's own observable characteristics (from columns 1 and 4 in Table 2), and then use this crude estimate of value-added as a summary statistic for all of a teacher's observed characteristics. We then reestimate the models replacing teacher characteristics with these summary statistics and mean teacher peer characteristics with the mean summary statistics of her peers. We find that a one standard deviation increase in own-teacher value-added due to *observable* characteristics is associated with a 3.6 and 2.6 percent of a standard deviation increase in math and reading test scores, respectively. Also, a one standard deviation increase in peer value-added due to *observable* peer characteristics is associated with a 0.8 and 0.6 percent of a standard deviation increase in math and reading test scores, respectively. These peer effects are statistically significant at the 1 percent level and yield coefficients that are about one-quarter of the size of the own-teacher effect. As previously noted, observable teacher characteristics are relatively weak predictors of a teacher's own quality, so these results are likely to be a lower bound on the true peer effects.

Table 3—Effect of Mean Peer Value-Added on Student Test Scores

	School fixed effects	Student fixed effects	Teacher-school and school-year effects	School fixed effects	Student fixed effects	Teacher-school and school-year effects
	Math test score			Reading test score		
	OLS	OLS	OLS	OLS	OLS	OLS
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Lagged score	0.7728	—	0.7712	0.7293	—	0.7233
	(0.0009)***	—	(0.0009)***	(0.0010)***	—	(0.0010)***
Teacher effect	0.1268	0.1689	—	0.0547	0.0785	—
	(0.0031)***	(0.0062)***	—	(0.0027)***	(0.006)***	—
Mean teacher peer effect	0.0522	0.0604	0.0398	0.0262	0.0346	0.026
	(0.0037)***	(0.0076)***	(0.0049)***	(0.0035)***	(0.0044)**	(0.0050)***
Observations	684,696	689,387	684,696	679,262	683,850	679,262
R^2	0.18	0.88		0.17	0.87	

Notes: Estimated using data from 2001 through 2006. All models include indicator variables for the gender and racial matches between the teacher and the students, class size, and year-by-grade fixed effects. All regressions include student demographic control variables except models that include student fixed effects. All regressions include teacher control variables except models that include teacher fixed effects (note that teacher experience is included in all models). All models include indicators for missing estimated value-added as well as the proportion of peers with no estimated value-added. The omitted teacher experience group is teachers with missing experience data. Robust standard errors are clustered by teacher-school in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

We now examine the results that use pre-period value-added as a potentially more powerful indicator of peer quality.

Peer Value-Added Results—Table 3 shows the effect of a teacher's peers' estimated value-added (estimated out-of-sample using 1995–2000 data) in math and reading on her own students' math and reading test score growth (using data from 2001 through 2006). To ensure that the teacher value-added results are not driven simply by the observable teacher characteristics, we estimated models that included estimated peer quality and observable peer characteristics. The coefficients on the observable peer characteristics are not statistically significant when estimated peer quality is included, and the inclusion of observable teacher peer characteristics has very little effect on the peer value-added estimates.[9] This suggests that the peer value-added estimates *are not* driven by any of the observable peer characteristics summarized in the previous section. Because observable teacher peer characteristics have little predictive power conditional on estimated teacher value-added, and because including them does not change the results in any meaningful way, we omit

[9] In models that include teacher peer experience and teacher peer value-added, one cannot reject the null hypothesis of no teacher-peer experience effects (conditional on peer value-added) at the 20 percent level for either math or reading. In contrast, for math and reading, the hypothesis that peers' value-added is equal to 0 (conditional on peer experience) is rejected at the 1 percent level.

observable teacher peer characteristics from this part of the analysis. Note that all models include the full set of controls from Table 2.

The results for both math and reading are robust across specifications that include school fixed effects, school and student fixed effects, and our preferred model, so we focus on the preferred specifications. The preferred model uses only within teacher and school variation to remove any selection of teachers to better peers, and it includes school-by-year effects to account for any school policies or school-specific shocks that could affect both peer quality and student test scores.[10] Columns 1–3 show the effects on math test scores. The coefficient on peer value-added for math in column 3 is 0.0398, suggesting that a 1 standard deviation increase in the mean estimated value-added of a teacher's peers is associated with a 3.98 percent of a standard deviation increase in math test scores. This is more than twice the size of the effect estimated using observable peer characteristics. For the average teacher with three peers, replacing 1 peer with another that has 1 standard deviation higher value-added will increase her students' math test scores by 1.3 percent of a standard deviation. This corresponds to between one-tenth and one-fifth of the own-teacher effect. Columns 4–6 show the effects on reading test scores. The effects are qualitatively similar to those for math. However, the magnitudes are smaller (a consistent finding in the teacher quality literature). The preferred model, in column 6, includes teacher-by-school and school-by-year effects. It shows that a 1 standard deviation increase in mean peer value-added is associated with a statistically significant 0.026 standard deviation increase in student reading test scores. For the average teacher with three peers, replacing one peer with another that has one standard deviation higher value-added will increase her students' test scores by 0.86 percent of a standard deviation. As for math, this corresponds to between one tenth and one fifth of the own-teacher effect.[11]

One implication of significant teacher peer effects is that failing to take teacher peer inputs into account when estimating own-teacher value-added could lead to inconsistent estimates. Although peer effects are important in explaining variation in student test scores, the amount explained by teacher quality is virtually identical in models that include or do not include peer value-added. In math, the proportion of the variance in test scores associated with the teacher fixed effects, $\mathrm{Cov}(A_{ij}\theta_j)/\mathrm{Var}(A_{ij})$, is 0.141 when peer attributes are included and 0.1432 when they are not included. In reading, $\mathrm{Cov}(A_{ij}\theta_j)/\mathrm{Var}(A_{ij})$ is 0.067 when peer attributes are included and 0.069 when they are not included. This suggests that the explanatory power of teacher effects is very slightly reduced when contemporaneous peer value-added is included.[12]

[10] This was implemented using the "felsdvreg" command written by Thomas Cornelissen, described in Cornelissen (2006), based on the three-way error model proposed by John M. Abowd, Robert H. Creecy, and Francis Kramarz (2002).

[11] A model that includes student fixed effects and teacher-school effects yields a coefficient on math peers of 0.026 with a standard error of 0.008, along with a coefficient on reading peers of 0.0196 with a standard error of 0.01.

[12] If some students were taught by their homeroom teacher's peers but were wrongly classified as being taught by the homeroom teacher, the explanatory power of the own teacher would be lower when peer attributes are included. The fact that the explanatory power of the own teacher is unchanged when peers are also included is consistent with our contention that the spillovers are not due to the actual teacher being misclassified as the teacher's peer.

VI. Specifications and Falsification Tests

Because students are not randomly assigned to teachers and teachers are not randomly assigned to schools or classrooms, it is important that we isolate variation that is not confounded by student selection, teacher self-selection, or correlated with confounding factors that also affect student achievement. Although including teacher-by-school effects credibly addresses the self-selection of teachers to peers, and although the inclusion of school-by-year effects credibly addresses the concern that schools that see improvements in peer quality may be improving in other areas, a few endogeneity concerns remain. We address these below.

A. *Dynamic Sorting Could Bias the Estimated Teacher Effects and Lead to Spurious Peer Effects*

It is possible that our results are confounded by dynamic sorting (or tracking) that would not be fully controlled for with a time-invariant student fixed effect or time-changing lagged test scores for two reasons. First, one of the identifying assumptions required to obtain unbiased teacher fixed effects on average, is that unobserved student characteristics are uncorrelated with true teacher ability, conditional on the included covariates. Dynamic sorting not captured by lagged test scores and other observable student characteristics could lead to bias in the estimated teacher value-added. Second, one of the identifying assumptions in the peer value-added models is that changes in a teacher's peers are uncorrelated with unobserved student characteristics. This may not be true with dynamic student sorting. For example, suppose principals assign "difficult" students to teachers with the highest value-added and assign the "easiest" students to less experienced or less able teachers. In such a scenario, when a strong older teacher retires and is replaced by a weaker and less experienced teacher, incumbent teachers will be more likely to receive the "difficult" students. Sorting of students across classrooms in such a manner would make it look as though having weaker peers hurts the incumbent teacher if the econometrician is unable to control sufficiently for student ability. This particular dynamic sorting story would be problematic because it would generate a *negative* correlation between true teacher quality and unobserved student ability.

Since these are important potential threats to the validity of our results, we present a falsification test of our identifying assumption that the unobserved student error term is not correlated with teacher value-added (i.e., $E[\varepsilon|\theta] = 0$) in Table 4. Because a student's future teacher should have no causal effect on that student's current test score performance, any nonzero effect would indicate bias. We contend that if there is positive/negative selection, the estimated value-added of the teacher the student will have in the following year will be positively/negatively correlated with the student's achievement in the current year. If there is positive/negative selection, those teachers who are systematically associated with contemporaneous gains should be predictive of test score gains/losses for their *future* students.

Columns 1 and 2 show the coefficients of the estimated value-added (estimated out of sample) of a student's current teacher and her future teacher for math and reading, respectively. These models include all student and teacher characteristics, school

Table 4—Effect of Future Teachers, Historical Peers, and Future Peers on Student Test Scores

	Math (1)	Reading (2)	Math (3)	Reading (4)	Math (5)	Reading (6)	Math (7)	Reading (8)
Teacher effect	0.12	0.052	—	—	—	—	—	—
(own subject)	(0.006)***	(0.005)***	—	—	—	—	—	—
Future teacher effect	0.002	–0.002	—	—	—	—	—	—
(own subject)	(0.003)	(0.004)	—	—	—	—	—	—
Peer effect	—	—	0.038	0.028	0.031	0.033	0.019	0.035
(own subject)	—	—	(0.005)***	(0.005)***	(0.0081)***	(0.0086)**	(0.0102)**	(0.0123)***
Lagged peer effect	—	—	—	—	0.037	0.021	0.051	0.025
(own subject)	—	—	—	—	(0.0074)***	(0.0084)**	(0.0093)***	(0.0112)**
Second lag of peer	—	—	—	—	0.01	0.018	0.011	0.0308
effect (own subject)	—	—	—	—	(0.0066)*	(0.0073)**	(0.0090)	(0.0089)***
Lead of peer effect	—	—	—	—	—	—	0.009	0.007
(own subject)	—	—	—	—	—	—	(0.0094)	(0.0103)
Teacher-school effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
School-year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	231,390	229,507	680,479	678,389	374,478	371,640	252,538	250,652
$\mathrm{Cov}(A_{ij}, \theta_j) / \mathrm{Var}(A_{ij})$	—	—	0.141	0.067	0.117	0.052	0.117	0.052
$\Pr(P>\|t\|)$ future effect = current effect	<0.001	<0.001	—	—	—	—	—	—
$\Pr(P>\|t\|)$ future teacher effect = 0	0.38	0.64	—	—	—	—	—	—
$\Pr(P>\|t\|)$ future = current	—	—	—	—	—	—	0.4	0.1
$\Pr(P>\|t\|)$ future = lag	—	—	—	—	—	—	<000	0.29
$\Pr(P>\|t\|)$ future = second lag	—	—	—	—	—	—	0.87	0.11

Notes: Estimated using data from 2001 to 2006. The variable "peer effect" is the mean estimated value-added of a teacher's peers (all other teachers at the same school in the same grade during the same year). All models include indicator variables for the gender and racial matches between the teacher and the students, class size, student demographic control variables, teacher experience, indicators for missing estimated value-added, the proportion of peers with no estimated value-added, and year-by-grade fixed effects. The omitted teacher experience group is teachers with missing experience data. Robust standard errors clustered by school-teacher in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

fixed effects, and grade-by-year fixed effects. (The main conclusions are invariant to the specification chosen.) Although the coefficient on the current teacher effect is 0.12 for math and 0.052 for reading (both significant at the 1 percent level), the coefficients for the *future* teacher's effect are only 0.002 and −0.002 for math and reading, respectively (both have p-values greater than 0.3). Teachers that we identify as effective have a strong positive effect on their current students' test scores but no effect on their future students' test scores. The null hypothesis of equality of the current teacher effect and the future teacher effect is rejected at the 1 percent level, the future teacher's coefficients are both less than one-tenth the size of those for contemporaneous teachers, and the future teacher effects for reading and math have opposite signs. This last fact is inconsistent with systematic selection because the math and reading teachers are, in fact, the same teacher. Furthermore, the standard errors on the future teacher value-added are small, indicating that the true values, if not actually zero, are very close to zero. These results suggest that no systematic student sorting occurs.

While we show that the estimated value-added of a student's future teacher does not predict their current test scores, it is helpful to show that future peers of a student's teacher do not affect the student's current test scores. Because teachers cannot learn from their future peers, future peers should have no effect on current student outcomes. As such, if our results were picking up some spurious correlation due to dynamic sorting, one might expect to find similar effects for a teacher's future peers as that of her current or past peers. We test this hypothesis by estimating the peer value-added model while including a teacher's current peers, her peers in the two previous years, (lagged peers), and her peers in the following year (future peer). These results are presented in columns 7 and 8 of Table 4.

The coefficient on mean peer quality the following year is not statistically significant for reading and math scores. In contrast, the effect of lagged peers is large and statistically significant for math, and the second lag is large and statistically significant for reading. Also, the point estimates are smaller for the lead than those of either of the lags or the contemporaneous effects for both math and reading. This test supports the validity of our strategy for three reasons. First, the future teacher effects are not even marginally statistically significant. Second, the future teacher effects are smaller than the current effects and the lag effects for both math and reading, which, under the null hypothesis of no causal effect and independence, would happen only with probability $((0.5)^3)^2 = 0.0156$. Third, one can reject the null hypothesis that the future teacher effect is the same as at least one of the lags or contemporaneous effects for both subjects at the 10 percent level. These results suggest that the identification strategy is valid.

B. *Changes in Peer Quality within Schools May Be Endogenous, so that Peer Quality Improvements Coincide with Other Grade-Specific Interventions*

The remaining endogeneity concern is that schools may be more likely to shift good teachers across grades or to hire better new teachers into a grade (at the same time as the schools shift other resources) when particular grades are performing poorly relative to other grades in the school. In this scenario, even with controls for school-by-year effects, some peer effects could be confounded by other resources and efforts in that grade. We believe that bias resulting from new hiring being correlated with other grade-specific changes is unlikely because schools do not have much control over when teachers leave and because new hiring is likely to take place due to changes in class size or vacancies occasioned by voluntary turnover. It is possible, however, that principals shift teachers across grades in response to poor grade performance at the same time that they implement other grade-specific improvements. We empirically test the possibility of endogenous peer changes both from outside the school and from within the school into a grade.

Specifically, we test for whether, conditional on school-by-year effects, current student performance, lagged students performance, or current peer quality affect the likelihood that a given grade in a given school receives a new peer. We test separately for receiving a peer from the same school, having a new peer in the grade from a different school in North Carolina, or having a new peer in the grade from outside the data. We present the results in Web Appendix Table A3. Across various specifications, one cannot reject the null hypothesis that the arrival of a new teacher to a particular grade within a school is unrelated to the historical level and growth of test outcomes, or to the estimated value-added of incumbent teachers in the grade, in both reading and math. However, teacher experience variables do have predictive power, as one would expect given that experience is a strong predictor of retirement. In sum, we find no evidence of endogenous peer changes.

VII. Suggestive Evidence of Learning

We show, in Section V, that teachers perform better when their peers are better, in terms of both observed and unobserved quality. As discussed in Section I, peer spillovers could exist for a variety of reasons. In this section, we test the four empirical predictions described in Section I that would be consistent with peer learning or peer-induced learning. All results from this point on include teacher-school and school-year fixed effects.

As discussed in Section I, since learning requires some investment on the part of a teacher, we might expect those teachers with the most to gain from these investments to do so. As such, if the results we observe are the result of teachers investing in job-specific human capital, we might expect the effects of peers to be largest among teachers who, ex ante, would experience larger benefits from job specific human capital investments. Most models of job-specific human capital suggest that workers with longer time horizons benefit from learning. Younger workers and workers with greater labor force attachment will be most likely to invest in job-specific human capital (Gary S. Becker 1962; Boyan Jovanovic 1979). We test for these patterns in our data by testing if the marginal effect of peers is larger for teachers with fewer years of experience, teachers who are national board certified, and teachers who are fully licensed. The results are presented in Web Appendix Table A4.

For both math and reading, first-year teachers are more responsive to peer quality changes than teachers with one or more years of experience. The null hypothesis of equality of effects is rejected at traditional levels for math (p-value = 0.001) but not for reading (p-value = 0.54). However, for neither (math or reading) are the marginal effects monotonic in experience. Given that the first year of teaching is the one when teachers acquire the most on-the-job knowledge (as evidenced by the very steep experience value-added profile), these findings support a learning interpretation. The results by national board certification status and license status also support a learning interpretation. Specifically, the interactions between peer quality and being a fully licensed teacher are positive and statistically significant at the 10 percent level for both math and reading. The interactions between peer quality and being certified yield positive, albeit not statistically significant, point estimates for both math and reading. In sum, while not conclusive, the results are consistent with the notion that teachers with greater labor market attachment are more sensitive to peer quality changes.

One of the principal differences between the alternative explanations for the peer effects observed in our main results is that if these peer effects are caused by learning, they should be persistent over time, and have the same sign. Although differences in resources or motivation caused by having better peers should have little effect once the teacher's peers change, any learning that has occurred should stay with the teacher. We test whether peer quality continues to affect a teacher in future years by including the first and second lag of the teacher's average peer quality, along with the contemporaneous measure of the quality of peers in that year. Results are presented in columns 5 and 6 of Table 4. Consistent with a learning interpretation, the peer effects are persistent. For math (column 5), the coefficient on the contemporaneous effect is 0.031, the coefficient on the first lag is 0.037, and the coefficient on the second lag of the peer effect is 0.01. The contemporaneous effect and the lagged effect are both statistically significant at the 1 percent level, while the second lag is statistically significant at the 10 percent level. For reading (column 6), the coefficient on the contemporaneous effect is 0.033, the coefficient on the first lag is 0.021, and the coefficient on the second lag of the peer effect is 0.018. All effects are statistically significant at the 5 percent level. For both subjects, the null hypothesis that the historical effects are equal to 0 is rejected at the 5 percent level. These findings support a learning interpretation. These results suggest that the total effect of peers is larger than what we estimate in the main baseline regression. After two years, a one standard deviation increase in teacher-peer quality that persists over time is associated with a 0.078 standard deviation increase in student test scores in math, and a 0.072 standard deviation increase in student test scores in reading.[13]

Because learning is a cumulative process, another prediction of the learning model is that historical peer quality should "explain away" some of the predictive power of teacher fixed effects. If teachers learn from their peers (or as a result of exposure to their peers), and if learning becomes part of teacher ability, then there should be less variation attributable to the time-invariant teacher indicator variables conditional on the history of their peers. We test this hypothesis by comparing the fraction of the variance in test scores explained by the individual teacher effects in models that do and do not control for lagged peer quality. In models that include only contemporaneous peer quality, $\mathrm{Cov}(A_{ij}, \theta_j) / \mathrm{Var}(A_{ij})$ is 0.141 for math and 0.067 for reading. In models that include the first and second lags of peer quality, $\mathrm{Cov}(A_{ij}, \theta_j) / \mathrm{Var}(A_{ij})$ is 0.117 for math and 0.052 for reading. These differences suggest that between 18 percent and 24 percent of the contemporaneous own-teacher effect can be attributed to her peers in the two previous years. This suggests learning and indicates that the observed spillovers are not due to transient changes in motivation or the allocation of nonclassroom tasks as a result of contemporaneous peer quality changes.[14]

[13] The persistence of peer quality over time also provides compelling evidence that our results are *not* driven by direct contact between students and their teacher's current peers, and is further evidence that our central findings are not driven by dynamic student sorting.

[14] As a test of whether teachers acquire grade-specific knowledge, such as how to teach fourth-grade math or more general teaching skills that would apply to all grades, we interacted the lagged peer value-added with indicators for whether the teacher moved to a new grade at the same school. If teachers acquired grade-specific skills, one would expect there to be greater persistence of peer effect for teachers who remain in the same grade. For math, the interactions are all positive and statistically insignificant, indicating that grade-specific knowledge may drive the spillovers for math. The joint hypothesis that all the interactions are equal to 0 is rejected at the 10

In sum, the empirical predictions that suggest peer learning are supported by the data. Although not all of the interaction effects yield statistically significant estimates, all the point estimates are consistent with a peer learning interpretation of the spillovers. Although we cannot prove that the spillovers are due to peer learning, the evidence, taken in its entirety, suggests that teachers either learn directly from their peers (direct peer learning) or make the decision to invest in the learning as a direct result of exposure to better peers (peer-induced learning).

VIII. Conclusions

We document that a teacher's own performance is affected by the quality of her peers. In particular, *changes* in the quality of a teacher's colleagues (all other teachers in the same school who teach students in the same grade) are associated with *changes* in her students' test score gains. Using two separate measures of peer quality, one based on observable teacher qualifications and the other on estimated peer effectiveness, we find that teachers perform better when the quality of their peers improves within the same school over time. This within-teacher relationship is robust to including school-by-year fixed effects to account for changes in school attributes over time that could be correlated with changes in the make-up of the teacher population. Findings are also robust to including student fixed-effects. In our preferred model, a one standard deviation improvement in *observable* teacher-peer quality is associated with a 0.008 and 0.006 standard deviation increase in math and reading scores, respectively. Using estimated value-added (estimated out-of-sample to avoid simultaneity bias), which is a much better predictor of subsequent student achievement, we find that a one standard deviation improvement in *estimated* teacher-peer quality is associated with a 0.0398 standard deviation increase in math scores and a 0.026 standard deviation increase in reading scores. Across both these measures of teacher quality and different specifications, for the average teacher with three peers, replacing one peer with another that has one standard deviation higher value-added corresponds to between one-fifth and one-tenth of the effect of replacing the own teacher with another that has one standard deviation higher value-added. We present a variety of falsification tests showing that our results are probably not driven by nonrandom dynamic student sorting across classrooms, or by the endogenous movement of teachers across grades or schools.

In an attempt to determine the mechanisms behind these spillovers, we test for empirical patterns that are consistent with peer-related learning. First, we show that less experienced teachers are generally more responsive to changes in peer quality than more experienced teachers. We also find that teachers who are certified and have regular licensure are generally more responsive to peer quality. The most compelling piece of evidence supporting the learning hypothesis is that the effect of teacher peer quality is persistent over time. Most peer effects that operate either through the education production function or through peer monitoring/pressure will have a

percent level. For reading, however, the results are mixed. The second lag is less persistent, while the first lag is more persistent for mobile teachers. The joint hypothesis that all the interactions are equal to 0 is rejected at the 5 percent level. In sum, the results of this test are mixed and inconclusive.

contemporaneous effect. We show that for both math and reading, the quality of a teacher's peers the year before, and even two years before, affects the achievement of her current students. For both subjects, the importance of a teacher's previous peers is as great as, or greater than, that of her current peers. The cumulative effect over three years of having peers with one standard deviation higher effectiveness is 0.078 standard deviations in math and 0.072 standard deviations in reading. Because teachers have about three peers on average, this is about one-third of the size of the own-teacher effect, suggesting that over time, teacher-peer quality is very important. Lastly, we find that peer quality in the previous two years "explains away" about one-fifth of the explanatory power of individual teachers. This suggests that a sizable part of the own-teacher effect is learned as a result of exposure to her previous peers. Although we acknowledge that we cannot prove peer-related learning, we believe these pieces of evidence lend themselves most naturally to a peer-related learning interpretation (either learning directly from peers or peer-induced learning).

As a theoretical matter, knowledge spillovers are tremendously important in canonical models of economic growth, despite relatively little empirical support. Our findings provide important micro evidence of this type of productivity spillover. From a policy perspective, the finding that teachers learn as a result of their peers is important because it has direct implications for how teachers should be placed in schools and how they should be compensated. For example, compensation schemes that reward a teacher's performance relative to her peers may be detrimental to fostering peer learning. Also, the fact that weaker and less experienced teachers are more responsive to peer quality than stronger and more experienced teachers suggests that novice teachers should be exposed to effective, experienced teachers. This would imply that the high concentration of novice teachers in inner-city schools could be particularly detrimental to student performance at these schools in both the long and the short run.

Although we find little evidence in our data that a failure to account for contemporaneous peers leads to biased estimates of the effect of own-teachers on student test scores, we do show that the assumption of no spillovers across teachers is not valid. Although our results are particularly relevant for the education setting, they add to the broader literature on peer effects. They highlight the type of data necessary to find evidence of peer effects and some of the features that may distinguish peer related learning from other types of peer spillovers. Although teachers in elementary school may be a somewhat unique group, the existence of peer effects and the suggestion of peer learning in this environment are suggestive that such spillovers may exist in other settings.

REFERENCES

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics*, 25(1): 95–135.

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." U.S. Census Bureau Technical Paper TP-2002-06.

Anderson, William, and Martin T. Wells. 2008. "Numerical Analysis in Least Squares Regression with an Application to the Abortion-Crime Debate." *Journal of Empirical Legal Studies*, 5(4): 647–81.

Anderson, T. W., and Cheng Hsiao. 1981. "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association*, 76(375): 598–606.

Angrist, Joshua D., and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94(5): 1613–34.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang. 2008. "Superstar Extinction." National Bureau of Economic Research Working Paper 14577.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. Forthcoming. "Social Incentives in the Workplace." *Review of Economic Studies.*

Battu, Harminder, Clive R. Belfield, and Peter J. Sloane. 2003. "Human Capital Spillovers within the Workplace: Evidence for Great Britain." *Oxford Bulletin of Economics and Statistics*, 65(5): 575–94.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy*, 70(S5): 9–49.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2005. "Who Teaches Whom? Race and the Distribution of Novice Teachers." *Economics of Education Review*, 24(4): 377–92.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2007. "How and Why Do Teacher Credentials Matter for Student Achievement?" National Bureau of Economic Research Working Paper 12828.

Clotfelter, Charles T., Helen F. Ladd, Jacob L. Vigdor, and Justin Wheeler. 2007. "High Poverty Schools and the Distribution of Teachers and Principals." *North Carolina Law Review*, 85(5): 1345–80.

Cornelissen, Thomas. 2006. "Using Stata for a Memory Saving Fixed Effects Estimation of the Three-Way Error Component Model." FDZ Methodenreport 03/2006.

Costa, Dora L., and Matthew E. Kahn. 2007. "Surviving Andersonville: The Benefits of Social Networks in POW Camps." *American Economic Review*, 97(4): 1467–87.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, 118(3): 815–42.

Guryan, Jonathan E., Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1(4): 34–68.

Hastings, Justine S., Richard Van Weelden, and Jeffrey Weinstein. 2007. "Preferences, Information, and Parental Choice Behavior in Public School Choice." National Bureau of Economic Research Working Paper 12995.

Hastings, Justine S., and Jeffrey M. Weinstein. 2007. "No Child Left Behind: Estimating the Impact on Choices and Student Outcomes." National Bureau of Economic Research Working Paper 13009.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research Working Paper 7867.

Hoxby, Caroline M., and Gretchen Weingarth. 2005. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Unpublished.

Ichino, Andrea, and Giovanni Maggi. 2000. "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *Quarterly Journal of Economics*, 115(3): 1057–90.

Irwin, Douglas A., and Peter J. Klenow. 1994. "Learning-by-Doing Spillovers in the Semiconductor Industry." *Journal of Political Economy*, 102(6): 1200–1227.

Jackson, C. Kirabo. 2009. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation." *Journal of Labor Economics*, 27(2): 213–56.

Jacob, Brian A., and Lars Lefgren. 2008. "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education." *Journal of Labor Economics*, 26(1): 101–36.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(5): 972–90.

Kandel, Eugene, and Edward P. Lazear. 1992. "Peer Pressure and Partnerships." *Journal of Political Economy*, 100(4): 801–17.

Kane, Thomas J., and Douglas O. Staiger. 2008. "Are Teacher-Level Value-Added Estimates Biased? An Experimental Validation of Non-Experimental Estimates." Unpublished.

Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger. 2006. "School Quality, Neighborhoods, and Housing Prices." *American Law and Economics Review*, 8(2): 183–212.

Koedel, Cory. 2008. "An Empirical Analysis of Teacher Spillover Effects in Secondary School." University of Missouri Department of Economics Working Paper 08-08.

Laschever, Ron. 2005. "The Doughboys Network: Social Interactions and Labor Market Outcomes of World War I Veterans." Unpublished.

Lavy, Victor, and Analia Schlosser. 2007. "Mechanisms and Impacts of Gender Peer Effects at School." National Bureau of Economic Research Working Paper 13292.

Lucas, Robert E., Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, 22(1): 3–42.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531–42.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review*, 99(1): 112–45.

Moretti, Enrico. 2004a. "Human Capital Externalities in Cities." In *Handbook of Regional and Urban Economics*, Vol. 4, ed. J. Vernon Henderson and Jacques-François Thisse, 2243–91. Amsterdam: Elsevier.

Moretti, Enrico. 2004b. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review*, 94(3): 656–90.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica*, 73(2): 417–58.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review*, 94(2): 247–52.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5): S71–102.

Rothstein, Jesse. 2007. "Do Value-Added Models Add Value? Tracking, Fixed Effects, and Causal Inference." Princeton University, Department of Economics Center for Economic Policy Studies Working Paper 1036.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116(2): 681–704.

Sorensen, Alan T. 2006. "Social Learning and Health Plan Choice." *RAND Journal of Economics*, 37(4): 929–45.

Thornton, Rebecca Achee, and Peter Thompson. 2001. "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding." *American Economic Review*, 91(5): 1350–68.

Todd, Petra E., and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal*, 113(485): F3–33.

Zimmerman, Martin B. 1982. "Learning Effects and the Commercialization of New Energy Technologies: The Case of Nuclear Power." *Bell Journal of Economics*, 13(2): 297–310.

This article has been cited by:

1. Kevin C. Bastian, Ludmila Janda. 2018. Does Quantity Affect Quality? Teachers' Course Preparations and Effectiveness. *Journal of Research on Educational Effectiveness* 1-24. [Crossref]
2. C. Kirabo Jackson. 2018. What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *Journal of Political Economy* **126**:5, 2072-2107. [Crossref]
3. Victoria Sevcenko, Sendil Ethiraj. 2018. How Do Firms Appropriate Value from Employees with Transferable Skills? A Study of the Appropriation Puzzle in Actively Managed Mutual Funds. *Organization Science* **29**:5, 775-795. [Crossref]
4. James P. Spillane, Matthew Shirrell, Samrachana Adhikari. 2018. Constructing "Experts" Among Peers: Educational Infrastructure, Test Data, and Teachers' Interactions About Teaching. *Educational Evaluation and Policy Analysis* 016237371878576. [Crossref]
5. Rebecca H. Woodland, Rebecca Mazur. 2018. Of Teams and Ties: Examining the Relationship Between Formal and Informal Instructional Support Networks. *Educational Administration Quarterly* **24**, 0013161X1878586. [Crossref]
6. David M. Quinn, James S. Kim. 2018. Experimental Effects of Program Management Approach on Teachers' Professional Ties and Social Capital. *Educational Evaluation and Policy Analysis* **40**:2, 196-218. [Crossref]
7. James P. Spillane, Megan Hopkins, Tracy M. Sweet. 2018. School District Educational Infrastructure and Change at Scale: Teacher Peer Interactions and Their Beliefs About Mathematics Instruction. *American Educational Research Journal* **55**:3, 532-571. [Crossref]
8. Casey Boyd-Swan, Chris M. Herbst. 2018. The demand for teacher characteristics in the market for child care: Evidence from a field experiment. *Journal of Public Economics* **159**, 183-202. [Crossref]
9. Ben Backes, Michael Hansen, Zeyu Xu, Victoria Brady. 2018. Examining Spillover Effects From Teach For America Corps Members in Miami-Dade County Public Schools. *Journal of Teacher Education* **13**, 002248711775230. [Crossref]
10. Douglas Wieczorek, Brandon Clark, George Theoharis. 2018. Principals' Instructional Feedback Practices During Race to the Top. *Leadership and Policy in Schools* **35**, 1-25. [Crossref]
11. David M. Quinn, James S. Kim. 2017. Scaffolding Fidelity and Adaptation in Educational Program Implementation: Experimental Evidence From a Literacy Intervention. *American Educational Research Journal* **54**:6, 1187-1220. [Crossref]
12. Anne S Robertson, Amanda Moore McBride, Saras Chung, Allison Williams. 2017. Sharing the classroom: A professional development opportunity for teachers and social workers. *Power and Education* **9**:3, 161-176. [Crossref]
13. Yincheng Ye, Kusum Singh. 2017. The effect of working condition on math teacher effectiveness: value-added scores and student satisfaction in teaching. *Educational Research for Policy and Practice* **16**:3, 283-295. [Crossref]
14. Elizabeth A. Bettini, Michelle M. Cumming, Kristen L. Merrill, Nelson C. Brunsting, Carl J. Liaupsin. 2017. Working Conditions in Self-Contained Settings for Students With Emotional Disturbance. *The Journal of Special Education* **51**:2, 83-94. [Crossref]
15. Megin Charner-Laird, Monica Ng, Susan Moore Johnson, Matthew A. Kraft, John P. Papay, Stefanie K. Reinhorn. 2017. Gauging Goodness of Fit: Teachers' Responses to Their Instructional Teams in High-Poverty Schools. *American Journal of Education* **123**:4, 553-584. [Crossref]

16. Li Feng, Tim R. Sass. 2017. Teacher Quality and Teacher Mobility. *Education Finance and Policy* **12**:3, 396-418. [Crossref]
17. Arild Aakvik, Frank Hansen, Gaute Torsvik. 2017. Productivity dynamics, performance feedback and group incentives in a sales organization. *Labour Economics* **46**, 110-117. [Crossref]
18. Prachee Sehgal, Ranjeet Nambudiri, Sushanta Kumar Mishra. 2017. Teacher effectiveness through self-efficacy, collaboration and principal leadership. *International Journal of Educational Management* **31**:4, 505-517. [Crossref]
19. James P. Spillane, Matthew Shirrell. 2017. Breaking Up Isn't Hard to Do. *Educational Administration Quarterly* **109**, 0013161X1769655. [Crossref]
20. Allison Atteberry, Susanna Loeb, James Wyckoff. 2017. Teacher Churning. *Educational Evaluation and Policy Analysis* **39**:1, 3-30. [Crossref]
21. Min Sun, Susanna Loeb, Jason A. Grissom. 2017. Building Teacher Teams. *Educational Evaluation and Policy Analysis* **39**:1, 104-125. [Crossref]
22. Thomas Cornelissen, Christian Dustmann, Uta Schönberg. 2017. Peer Effects in the Workplace. *American Economic Review* **107**:2, 425-456. [Abstract] [View PDF article] [PDF with links]
23. Michael D. Siciliano. 2017. Professional Networks and Street-Level Performance. *The American Review of Public Administration* **47**:1, 79-101. [Crossref]
24. Elizabeth A. Bettini, Jean B. Crockett, Mary T. Brownell, Kristen L. Merrill. 2016. Relationships Between Working Conditions and Special Educators' Instruction. *The Journal of Special Education* **50**:3, 178-190. [Crossref]
25. James P. Spillane, Matthew Shirrell, Megan Hopkins. 2016. Designing and Deploying A Professional Learning Community (PLC) Organizational Routine: Bureaucratic and Collegial Arrangements in Tandem. *Les dossiers des sciences de l'éducation* :35, 97-122. [Crossref]
26. Erik Grönqvist, Jonas Vlachos. 2016. One size fits all? The effects of teachers' cognitive and social abilities on student achievement. *Labour Economics* **42**, 138-150. [Crossref]
27. M. A. Kraft, W. H. Marinell, D. Shen-Wei Yee. 2016. School Organizational Contexts, Teacher Turnover, and Student Achievement: Evidence From Panel Data. *American Educational Research Journal* **53**:5, 1411-1449. [Crossref]
28. Semih Tumen, Tugba Zeydanli. 2016. Social interactions in job satisfaction. *International Journal of Manpower* **37**:3, 426-455. [Crossref]
29. M. P. Steinberg, R. Garrett. 2016. Classroom Composition and Measured Teacher Performance: What Do Teacher Observation Scores Really Measure?. *Educational Evaluation and Policy Analysis* **38**:2, 293-317. [Crossref]
30. Julie Cohen, Michelle Brown. 2016. Teaching Quality Across School Settings. *The New Educator* **12**:2, 191-218. [Crossref]
31. Jason B. Cook, Richard K. Mansfield. 2016. Task-specific experience and task-specific talent: Decomposing the productivity of high school teachers. *Journal of Public Economics* . [Crossref]
32. M. D. Siciliano. 2016. Its the Quality Not the Quantity of Ties That Matters: Social Networks and Self-Efficacy Beliefs. *American Educational Research Journal* **53**:2, 227-262. [Crossref]
33. Eric A. Hanushek. 2016. School human capital and teacher salary policies. *Journal of Professional Capital and Community* **1**:1, 23-40. [Crossref]
34. D. Figlio, K. Karbownik, K.G. Salvanes. Education Research and Administrative Data 75-138. [Crossref]

35. Z. Xu, U. O zek, M. Hansen. 2015. Teacher Performance Trajectories in High- and Lower-Poverty Schools. *Educational Evaluation and Policy Analysis* **37**:4, 458-477. [Crossref]

36. James P. Spillane, Megan Hopkins, Tracy M. Sweet. 2015. Intra- and Interschool Interactions about Instruction: Exploring the Conditions for Social Capital Development. *American Journal of Education* **122**:1, 71-110. [Crossref]

37. D. Herbst, A. Mas. 2015. Peer effects on worker output in the laboratory generalize to the field. *Science* **350**:6260, 545-549. [Crossref]

38. John P. Papay, Matthew A. Kraft. 2015. Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics* **130**, 105-119. [Crossref]

39. Matthew Ronfeldt. 2015. Field Placement Schools and Instructional Effectiveness. *Journal of Teacher Education* **66**:4, 304-320. [Crossref]

40. Matthew P. Steinberg, Lauren Sartain. 2015. Does Teacher Evaluation Improve School Performance? Experimental Evidence from Chicago's Excellence in Teaching Project. *Education Finance and Policy* 1-38. [Crossref]

41. Matthew A. Kraft. 2015. Teacher Layoffs, Teacher Quality, and Student Achievement: Evidence from a Discretionary Layoff Policy. *Education Finance and Policy* 1-41. [Crossref]

42. Matthew Ronfeldt, Susanna Owens Farmer, Kiel McQueen, Jason A. Grissom. 2015. Teacher Collaboration in Instructional Teams and Student Achievement. *American Educational Research Journal* **52**:3, 475-514. [Crossref]

43. James Spillane. 2015. Leadership and Learning: Conceptualizing Relations between School Administrative Practice and Instructional Practice. *Societies* **5**:2, 277-294. [Crossref]

44. Scott A. Imberman, Michael F. Lovenheim. 2015. Incentive Strength and Teacher Productivity: Evidence from a Group-Based Teacher Incentive Pay System. *Review of Economics and Statistics* **97**:2, 364-386. [Crossref]

45. D. Blazar. 2015. Grade Assignments and the Teacher Pipeline: A Low-Cost Lever to Improve Student Achievement?. *Educational Researcher* **44**:4, 213-227. [Crossref]

46. S. Gershenson, L. Langbein. 2015. The Effect of Primary School Size on Academic Achievement. *Educational Evaluation and Policy Analysis* **37**:1 Suppl, 135S-155S. [Crossref]

47. S. M. Johnson. 2015. Will VAMS Reinforce the Walls of the Egg-Crate School?. *Educational Researcher* **44**:2, 117-126. [Crossref]

48. Dimitri Van Maele, Mieke Van Houtte. 2015. Trust in school: a pathway to inhibit teacher burnout?. *Journal of Educational Administration* **53**:1, 93-115. [Crossref]

49. Eric Isenberg, Bing-ru Teh, Elias Walsh. 2015. Elementary School Data Issues for Value-Added Models: Implications for Research. *Journal of Research on Educational Effectiveness* **8**:1, 120-129. [Crossref]

50. Dan Goldhaber, Joe Walch, Brian Gabele. 2014. Does the Model Matter? Exploring the Relationship Between Different Student Achievement-Based Teacher Assessments. *Statistics and Public Policy* **1**:1, 28-39. [Crossref]

51. M. A. Kraft, J. P. Papay. 2014. Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Evaluation and Policy Analysis* **36**:4, 476-500. [Crossref]

52. Kamal Hamdan, Jill Aguilar, Patricia Yee, Andrea Nee, Xiomara Benitez, Cindy Medina, Jeff Sapp. Recruitment, Selection, Placement, and Support in the Preparation of Quality of Urban Secondary Teachers 221-237. [Crossref]

53. Quentin Brummet. 2014. The effect of school closings on student achievement. *Journal of Public Economics* **119**, 108-124. [Crossref]

54. C. Kirabo Jackson. 2014. Teacher Quality at the High School Level: The Importance of Accounting for Tracks. *Journal of Labor Economics* **32**:4, 645-684. [Crossref]

55. C. Ferrera, J. Fernández, A.C. Marcos. 2014. The cooperative learning: Understanding and increasing the knowledge of the facilities design without a professor extra effort. *Multidisciplinary Journal for Education, Social and Technological Sciences* **1**:2, 1. [Crossref]

56. Raj Chetty, John N. Friedman, Jonah E. Rockoff. 2014. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review* **104**:9, 2593-2632. [Abstract] [View PDF article] [PDF with links]

57. C. Kirabo Jackson, Jonah E. Rockoff, Douglas O. Staiger. 2014. Teacher Effects and Teacher-Related Policies. *Annual Review of Economics* **6**:1, 801-825. [Crossref]

58. Benjamin Master. 2014. Staffing for Success. *Educational Evaluation and Policy Analysis* **36**:2, 207-227. [Crossref]

59. Amine Ouazad. 2014. Assessed by a Teacher Like Me: Race and Teacher Assessments. *Education Finance and Policy* 1-39. [Crossref]

60. Ben Ost. 2014. How Do Teachers Improve? The Relative Importance of Specific and General Human Capital. *American Economic Journal: Applied Economics* **6**:2, 127-151. [Abstract] [View PDF article] [PDF with links]

61. David S. Lyle, John Z. Smith. 2014. The Effect of High-Performing Mentors on Junior Officer Promotion in the US Army. *Journal of Labor Economics* **32**:2, 229-258. [Crossref]

62. James P. Spillane, Megan Hopkins. 2013. Organizing for instruction in education systems and school organizations: how the subject matters. *Journal of Curriculum Studies* **45**:6, 721-747. [Crossref]

63. C. Kirabo Jackson. 2013. Can higher-achieving peers explain the benefits to attending selective schools? Evidence from Trinidad and Tobago. *Journal of Public Economics* **108**, 63-77. [Crossref]

64. Jason A. Grissom, Susanna Loeb, Nathaniel A. Nakashima. 2013. Strategic Involuntary Teacher Transfers and Teacher Performance: Examining Equity and Efficiency. *Journal of Policy Analysis and Management* n/a-n/a. [Crossref]

65. C. KIRABO JACKSON. 2013. DO COLLEGE-PREPARATORY PROGRAMS IMPROVE LONG-TERM OUTCOMES?. *Economic Inquiry* no-no. [Crossref]

66. C. Kirabo Jackson. 2013. Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *Review of Economics and Statistics* **95**:4, 1096-1116. [Crossref]

67. Michael R. Strain. 2013. Single-sex classes & student outcomes: Evidence from North Carolina. *Economics of Education Review* **36**, 73-87. [Crossref]

68. Brian A. Jacob. 2013. The Effect of Employment Protection on Teacher Effort. *Journal of Labor Economics* **31**:4, 727-761. [Crossref]

69. M. Sun, W. R. Penuel, K. A. Frank, H. A. Gallagher, P. Youngs. 2013. Shaping Professional Development to Promote the Diffusion of Instructional Expertise Among Teachers. *Educational Evaluation and Policy Analysis* **35**:3, 344-369. [Crossref]

70. Dan Goldhaber, Michael Hansen. 2013. Is it Just a Bad Class? Assessing the Long-term Stability of Estimated Teacher Performance. *Economica* **80**:319, 589-612. [Crossref]

71. Sean Corcoran, Dan Goldhaber. 2013. Value Added and Its Uses: Where You Stand Depends on Where You Sit. *Education Finance and Policy* **8**:3, 418-434. [Crossref]

72. Dan Goldhaber, Stephanie Liddle, Roddy Theobald. 2013. The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review* **34**, 29-44. [Crossref]

73. Sarena F. Goodman, Lesley J. Turner. 2013. The Design of Teacher Incentive Pay and Educational Outcomes: Evidence from the New York City Bonus Program. *Journal of Labor Economics* **31**:2, 409-420. [Crossref]

74. Eric S. Taylor,, John H. Tyler. 2012. The Effect of Evaluation on Teacher Performance. *American Economic Review* **102**:7, 3628-3651. [Abstract] [View PDF article] [PDF with links]

75. Dan Goldhaber, Joe Walch. 2012. Strategic pay reform: A student outcomes-based evaluation of Denver's ProComp teacher pay initiative. *Economics of Education Review* **31**:6, 1067-1083. [Crossref]

76. John P. Papay, Martin R. West, Jon B. Fullerton, Thomas J. Kane. 2012. Does an Urban Teacher Residency Increase Student Achievement? Early Evidence From Boston. *Educational Evaluation and Policy Analysis* **34**:4, 413-434. [Crossref]

77. James H. Nehring, Ellen J. O'Brien. 2012. Strong agents and weak systems: University support for school level improvement. *Journal of Educational Change* **13**:4, 449-485. [Crossref]

78. William R. Penuel, Min Sun, Kenneth A. Frank, H. Alix Gallagher. 2012. Using Social Network Analysis to Study How Collegial Interactions Can Augment Teacher Learning from External Professional Development. *American Journal of Education* **119**:1, 103-136. [Crossref]

79. Susan Moore Johnson. 2012. Build the Capacity of Teachers and Their Schools. *Phi Delta Kappan* **94**:2, 62-65. [Crossref]

80. Tim R. Sass, Jane Hannaway, Zeyu Xu, David N. Figlio, Li Feng. 2012. Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics* **72**:2-3, 104-122. [Crossref]

81. John P. Papay, Susan Moore Johnson. 2012. Is PAR a Good Investment? Understanding the Costs and Benefits of Teacher Peer Assistance and Review Programs. *Educational Policy* **26**:5, 696-729. [Crossref]

82. Eric A. Hanushek, Steven G. Rivkin. 2012. The Distribution of Teacher Quality and Implications for Policy. *Annual Review of Economics* **4**:1, 131-157. [Crossref]

83. C. Kirabo Jackson. 2012. School competition and teacher labor markets: Evidence from charter school entry in North Carolina. *Journal of Public Economics* **96**:5-6, 431-448. [Crossref]

84. M. Ronfeldt. 2012. Where Should Student Teachers Learn to Teach?: Effects of Field Placement School Characteristics on Teacher Retention and Effectiveness. *Educational Evaluation and Policy Analysis* **34**:1, 3-26. [Crossref]

85. Dan Goldhaber, Betheny Gross, Daniel Player. 2011. Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best?. *Journal of Policy Analysis and Management* **30**:1, 57-87. [Crossref]

86. C. Kirabo Jackson,, Henry S. Schneider. 2011. Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry. *American Economic Journal: Applied Economics* **3**:3, 244-267. [Abstract] [View PDF article] [PDF with links]

87. Kenneth A. Frank, Yong Zhao, William R. Penuel, Nicole Ellefson, Susan Porter. 2011. Focus, Fiddle, and Friends. *Sociology of Education* **84**:2, 137-156. [Crossref]

88. C. Kirabo Jackson. 2010. Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago*. *The Economic Journal* **120**:549, 1399-1429. [Crossref]

89. Anne Henry Cash. A Call for Mixed Methods in Evaluating Teacher Preparation Programs 547-572. [Crossref]