Teachers

"I don't want an apple, Danny—do you have any money?"

Broad References

Books

Lovenheim, M., Turner, S. E. (2017). *Economics of Education.* Macmillan Higher Education.

Neal, D. A. (2018). *Information, Incentives, and Education Policy.* Harvard University Press.

Papers

– Douglass, H. R. (1958). "What Is a Good Teacher?". *The High School Journal*, 41(4), 110-113.[1]

– Rosen, S. (1987). "Some Economics of Teaching". *Journal of Labor Economics*, 5(4, Part 1), 561-575.

[1] Unfortunately I can't find this reference, which seems very nice.

> Adam Smith's discussion of the payment of teachers is reviewed in terms of industrial organization and agency theory. The impli cit stu-dent fees necessary to support annual salaries average $1.30 per class meeting in primary and secondary schools and rise to $4.00 per lecture and up for college teachers. While salaries in teaching a re much smaller than in the large-scale visual media, implicit valuat ions per contact hour in teaching are at least six hundred times larger than in television. Classroom teaching is expensive because a teac her's scale of operations is sharply constrained by the student-teach er ratio.

Interesting paper, it seems that Rosen was truly a scholar! Interesting discussion of why the market for educational services is typically "anchored" on institutions, but also raises the question of why and how would teachers be compensated in decentralized student-teacher interactions.

– Lazear, E. P. (2001)*** "Educational Production". *The Quarterly Journal of Economics*, 116(3), 777-803.

> Classroom education has public good aspects. The technology is such that when one student disrupts the class, learning is reduced for all other students. A disruption model of educational production is presented. It is shown that optimal class size is larger for better-behaved students, which helps explain why it is difficult to find class size effects in the data. Additionally, the role of discipline is analyzed and applied to differences in performance of Catholic and public schools. An empirical framework is discussed where the importance of sorting students, teacher quality, and other factors can be assessed.

I really like this paper. It starts with a paragraph that couldn't be clearer and most representative on the way economists think about patterns in human behavior: "There exists an enormous empirical literature on the relation of educational attainment to class size. Results in this literature vary from significant class size effects to no (or sometimes perverse) class size effects. The inability to find consistent class size effects is most perplexing. At some basic level*, the failure to observe class size effects makes no sense because observed class class size is generally smaller than the entire number of students at any particular grade level. Why bear the expense of having four kindergarten classes of 30 rather than one class of 120 if class size truly does not matter? Furthermore, observed class size varies with age of the student. Preschool classes are smaller than large lecture classes for college students. How is this to be explained if class size is irrelevant?" (p. 777-8). The basic starting point is to think about* formal *classroom teaching as a public good, where congestion (peer) effects are negative externalities created when one student impede the learning of classmates (see fn. 5, for why this distinguishes formal education from learning in infancy or in the labor market). The model describes the optimizing behavior of a school that selects the number of teachers and classrooms maximizing profits in a competitive market — zero profits — in the presence of these negative externalities. The most important quantity of interest in the model (p) is the probability that any given student does not engage on disrupting behavior that destroys the classroom public good — notice that this is an "O-ring" theory of classroom production, à la Kremer, where problems in one link of the ring can completely destroy output. What are the implications of this model for teacher quality, which is the focus of this section of readings? Lazear discusses this in pages 795-6, where he states that p is "as much a function of the teacher as it is of the student characteristics". He then presents a numerical example on the effects of increasing teacher quality through pay increases that would allow schools to engage in more selective hiring. Also, the model provides a rationale of why class size reductions should provide better results for disadvantaged and special needs children. Given the centrality of this paper, it also yields interesting interpretations of many other patterns in economics of education which are discussed in other parts of this annotated bibliography.*

– Todd, P. E., Wolpin, K. I. (2003)*** "On the Specification and Estimation of the Production Function for Cognitive Achievement". *The Economic Journal*, 113(485), F3-F33.[2], specially section 2.3.4.

[2]In general, this is a paper that seems worth studying much more carefully than I was able to do.

> This paper considers methods for modelling the production function for cognitive achievement in a way that captures theoretical notions that child development is a cumulative process depending on the history of family and school inputs and on innate ability. It develops a general modelling framework that accommodates many of the estimating equations used in the literatures. It considers different ways of addressing data limitations, and it makes precise the identifying assumptions needed to justify alternative approaches. Commonly used specifications are shown to place restrictive assumptions on the production technology. Ways of testing modelling assumptions and of relaxing them are discussed.

I believe this is a canonical paper on the topic. The most important part is the second half, from section 2.2. onwards, equation (3) is the model proposed in its most general form, where cognitive achievement, as measured by test performance at some particular age, is the outcome of a cumulative process of knowledge acquisition that combines the whole history of family inputs, school inputs, genetic endowments and random error. The two basic problems in estimating such a model are omitted variables (genetic endowments) and missing data on the whole history of inputs. Section 2.3. presents an "inventory" of the specifications in the literature, and the way they deal with these problems (see Table 3). Importantly, the author is explicit about assumptions on the production technology and on the input decision rules that would justify its applications. The inventory encompasses the "contemporaneous specificication", "value added", "cumulative". Some takeaways from the basic "value added" model are very interesting, for instance that any model that admits the presence of unobserved endowments must also deal with the fact that this will be most likely correlated with baseline achievement, and this would bias all *estimates in the regression. Importantly, section 2.3.4. is a discussion on teacher and school value added estimates: are they policy effects or production function parameters? This is an important discussion for the interpretation of these estimates. This is an overarching topic of the paper.*

– Staiger D. O., Rockoff J.E. (2010a)***. "Searching for Effective Teachers with Imperfect Information". *Journal of Economic Perspectives*. 24:97–117 + Staiger D. O., Rockoff J.E. (2010b)**. "Online Appendix: Searching for Effective Teachers with Imperfect Information". *Journal of Economic Perspectives*. Available Online.

> Over the past four decades, empirical researchers – many of them economists – have accumulated an impressive amount of evidence on teachers. In this paper, we ask what the existing evidence implies for how school leaders might recruit, evaluate, and retain teachers. We begin by summarizing the evidence on five key points, referring to existing work and to evidence we have accumulated from our research with the nation's two largest school districts: Los Angeles and New York City. First, teachers display considerable heterogeneity in their effects on student achievement gains. Second, estimates of teacher effectiveness based on student achievement data are noisy measures. Third, teachers' effectiveness rises rapidly in the first year or two of their teaching careers but then quickly levels out. Fourth, the primary cost of teacher turnover is not the direct cost of hiring and firing, but rather is the loss to students who will be taught by a novice teacher rather than one with several years of experience. Fifth, it is difficult to identify at the time of hire those teachers who will prove more effective. As a result, better teachers can only be identified after some evidence on their actual job performance has accumulated. We then explore what these facts imply for how principals and school districts should act, using a simple model in which schools must search for teachers using noisy signals of teacher effectiveness. The implications of our analysis are strikingly different from current practice. Rather than screening at the time of hire, the evidence on heterogeneity of teacher performance suggests a better strategy would be identifying large differences between teachers by observing the first few years of teaching performance and retaining only the highest-performing teachers.

This paper is interesting in general, but the approach from p. 105 onwards and the online appendix is probably what should be used as distinctive and incorporated into a course contemplating teachers.

Vegas, E., Ganimian, A. (2013). "Theory and Evidence on Teacher Policies in Developed and Developing Countries". IADB Working Paper.

The past decade has seen the emergence of numerous rigorous impact evaluations of teacher policies. This paper reviews the economic theory and empirical evidence on eight teacher policy goals: (1) setting clear expectations for teachers; (2) attracting the best into teaching; (3) preparing teachers with useful training and experience; (4) matching teachers' skills with students' needs; (5) leading teachers with strong principals; (6) monitoring teaching and learning; (7) supporting teachers to improve instruction; and (8) motivating teachers to perform. The paper also discusses key concepts and methods in econometrics to understand existing studies and offers some directions for future research.

Survey that takes both a theoretical and abstract approach and discusses specific references. This paper also cites and discusses references from "classical" economics that are very useful for thinking about teachers. I was inspired by this and incorporated some of them in this bibliography.

Jackson, C. K., Rockoff, J. E., Staiger, D. O. (2014)***. "Teacher Effects and Teacher-Related Policies". *Annual Review of Economics*, 6(1), 801-825.

The emergence of large longitudinal data sets linking students to teachers has led to rapid growth in the study of teacher effects on student outcomes by economists over the past decade. One large literature has documented wide variation in teacher effectiveness that is not well explained by observable student or teacher characteristics. A second literature has investigated how educational outcomes might be improved by leveraging teacher effectiveness through processes of recruitment, assignment, compensation, evaluation, promotion, and retention. These two lines of inquiry are closely tied; the first tells us about the importance of individual teachers, and the second tells us how this information can be used in policy and practice. We review the most recent findings in economics on the importance of teachers and on teacher-related policies aimed at improving educational production.

Mbiti, I. M. (2016). "The Need for Accountability in Education in Developing Countries". *Journal of Economic Perspectives*, 30(3), 109-32.

Despite the rapid growth in enrollment rates across the developing world, there are major concerns about the quality of education that children receive. Across numerous developing countries, recent learning assessments have revealed that children are not able to develop basic numeracy and literary skills. These low levels of learning are the result of a number of interrelated factors, many of which reflect the low levels of accountability across multiple levels of the education system. In this paper, I document the main education challenges facing developing countries, including the lack of accountability among teachers and school management. I also review recent literature that documents the effectiveness of interventions aimed at addressing these accountability issues. Finally, I assess the potential for the market to improve accountability in the education sector in developing countries.

Gershenson, S. (2021). "Identifying and Producing Effective Teachers". *IZA Discussion Papers*, (14096).

Teachers are among the most important school-provided determinants of student success. Effective teachers improve students' test scores as well as their attendance, behavior, and earnings as adults. However, students do not enjoy equal access to effective teachers. This article reviews some of the key challenges associated with teacher policy confronted by school leaders and education policymakers, and how the tools of applied economics can help address those challenges. The first challenge is that identifying effective teachers is difficult. Economists use value-added models to estimate teacher effectiveness, which works well in certain circumstances, but should be just one piece of a multi-measure strategy for identifying effective teachers. We also discuss how different policies, incentives, school characteristics, and professional-development interventions can increase teacher effectiveness; this is important, as schools face the daunting challenge of hiring effective teachers, helping teachers to improve, and removing ineffective teachers from the classroom. Finally, we discuss the supply and mobility of teachers, including the consequences of teacher absenteeism, the distribution of initial teaching placements, and the characteristics and preferences of those who enter the profession.

Cross-Country Perspectives

Bruns, B., De Gregorio, S., Taut, S. (2016). "Measures of Effective Teaching in Developing Countries". *Research on Improving Systems of Education Working Paper*, 16(009).

Bietenbeck, J., Piopiunik, M., Wiederhold, S. (2018). "Africa's Skill Tragedy: Does Teachers' Lack of Knowledge Lead to Low Student Performance?". *Journal of Human Resources*, 53(3), 553-578.

Hanushek, E. A., Piopiunik, M., Wiederhold, S. (2019). The value of smarter teachers international evidence on teacher cognitive skills and student performance. Journal of Human Resources, 54(4), 857-899.

HANUSHEK, E., AND L. WOESSMANN (2008): "The Role of Cognitive Skills in Economic Development," Journal of Economic Literature, 46, 607–668. [927]

International Differences in Educational Achievement Hanushek, E. A., Woessmann, L. (2011). The economics of international differences in educational achievement. In Handbook of the Economics of Education (Vol. 3, pp. 89-200). Elsevier.

Pritchett, L. (2006). Does learning to add up add up? The returns to schooling in aggregate data. Handbook of the Economics of Education, 1, 635-695.

The Labor Market of Teachers

Who Becomes a Teacher?

Biasi, B. (2018). "The labor market for teachers under different pay schemes" (No. w24813). National Bureau of Economic Research.

Dinerstein, M., Megalokonomou, R., Yannelis, C. (2020). "Human Capital Depreciation" (No. w27925). National Bureau of Economic Research.

Rothstein, J. (2015). Teacher quality policy when supply matters. American Economic Review, 105(1), 100-130.

Loeb and Page. "Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary", The Review of Economics and Statistics, 82 (3), 2000.

Corcoran, S. P., W. N. Evans, and R. M. Schwab. 2004. "Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957-2000," American Economic Review 94(2), 230-235.

Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. Review of Economics and Statistics, 95(4), 1096-1116.

Nagler, M., Piopiunik, M., West, M. R. (2020). Weak markets, strong teachers: Recession at career start and teacher effectiveness. *Journal of Labor Economics*, 38(2), 453-500.

Effects of Teachers' Observables and Human Capital

Ost, B. (2014). How do teachers improve? The relative importance of specific and general human capital. American Economic Journal: Applied Economics, 6(2), 127-51.

Grönqvist, E., Vlachos, J. (2016). One size fits all? The effects of teachers' cognitive and social abilities on student achievement. Labour Economics, 42, 138-150.

Hanushek, E. A., Piopiunik, M., Wiederhold, S. (2014). The Value of Smarter Teachers. THE JOURNAL OF HUMAN RESOURCES, 5(4), 4.

Kraft, M. A., Papay, J. P., Chi, O. L. (2020). Teacher skill development: Evidence from performance ratings by principals. Journal of Policy Analysis and Management, 39(2), 315-347.

Marioni, L. D. S., Freguglia, R. D. S., Menezes-Filho, N. A. (2020). The impacts of teacher working conditions and human capital on student achievement: evidence from brazilian longitudinal data. Applied Economics, 52(6), 568-582.

Kane TJ, Rockoff JE, Staiger DO. 2006. What does certification tell us about teacher effectiveness? Evidence from New York City. NBER Work. Pap. 12155

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. American Economic Review, 95(2), 158-165.

Lim, J., Meer, J. (2017). The impact of teacher–student gender matches random assignment evidence from South Korea. Journal of Human Resources, 52(4), 979-997.

Gershenson, S., Hart, C. M., Hyman, J., Lindsay, C., Papageorge, N. W. (2018). The long-run impacts of same-race teachers (No. w25254). National Bureau of Economic Research.

Qureshi, J. A., Ost, B. (2020). The Role of Families in Student Sorting to Teachers. Journal of Human Resources, 55(2), 470-503.

Teacher Effects: Quality as Value-Added

Comments

A **teacher effect** is a number given to systematic variation in output across students assigned to the same teacher, however that was achieved — better communication to students, classroom management, encouragement of greater effort by students or parents, etc.

For policy decisions, teacher effects have been used to identify high-quality teachers. There are three questions that have been central to this research program. First, to define a teacher effect, one has to pin down dimensions of the human capital stock of students and a time horizon to define what "output" means. Do these measures bring reliable information that could not be obtained using other teachers' characteristics?

□

Do teachers count? Are all teachers perfectly substitutable? An old question!

– Hanushek, E. (1971). "Teacher Characteristics and Gains in Student Achievement: Estimation using Micro Data". *The American Economic Review*, 61(2), 280-288.

> Educational research has been slow in providing definite answers to public policy questions for several understandable reasons: the subject of the educational process is extremely complex especially as regards to physiological and psychological aspects; any theoretical development of a learning theory amenable to analysis for policy purposes is absent; and the required data traditionally have not been collected. This analysis represents a next step of statistical inquiry into the educational process from a public policy point of view.

Comment: I really like this paper, because it shows how much the basic questions in this literature have evolved — unfortunately, maybe not so much with the "theoretical development of a learning theory amenable to analysis for policy purposes", I feel that the model in Todd and Wolpin (EJ, 2003) is very similar, albeit identification is more thoroughly discussed.

Quality as Value-Added

– Rivkin, S. G., Hanushek, E. A., Kain, J. F. (2005). "Teachers, Schools, and Academic Achievement". *Econometrica*, 73(2), 417-458.

This paper disentangles the impact of schools and teachers in influencing achievement with special attention given to the potential problems of omitted or mismeasured variables and of student and school selection. Unique matched panel data from the UTD Texas Schools Project permit the identification of teacher quality based on student performance along with the impact of specific, measured components of teachers and schools. Semiparametric lower bound estimates of the variance in teacher quality based entirely on within-school heterogeneity indicate that teachers have powerful effects on reading and mathematics achievement, though little of the variation in teacher quality is explained by observable characteristics such as education or experience. The results suggest that the effects of a costly ten student reduction in class size are smaller than the benefit of moving one standard deviation up the teacher quality distribution, highlighting the importance of teacher effectiveness in the determination of school quality.

Comment: .

Carrell, S. E., West, J. E. (2010). Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy*, 118(3), 409-432.

In primary and secondary education, measures of teacher quality are often based on contemporaneous student performance on standardized achievement tests. In the postsecondary environment, scores on student evaluations of professors are typically used to measure teaching quality. We possess unique data that allow us to measure relative student performance in mandatory follow-on classes. We compare metrics that capture these three different notions of instructional quality and present evidence that professors who excel at promoting contemporaneous student achievement teach in ways that improve their student evaluations but harm the follow-on achievement of their students in more advanced classes.

Staiger D. O., Rockoff J.E. (2010). "Searching for Effective Teachers with Imperfect Information". *Journal of Economic Perspectives*. 24:97–117

Jackson, C. K. (2014). "Teacher Quality at the High School Level: The Importance of Accounting for Tracks". *Journal of Labor Economics*, 32(4), 645-684.

Chetty, R., Friedman, J. N., Rockoff, J. E. (2014). "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates". *American Economic Review*, 104(9), 2593-2632.

Chetty, R., Friedman, J. N., Rockoff, J. E. (2014). "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood". *American Economic Review*, 104(9), 2633-79.

Jackson, C. K., Rockoff, J. E., Staiger, D. O. (2014). "Teacher Effects and Teacher-Related Policies". *Annual Review of Economics*, 6(1), 801-825.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. The Quarterly Journal of Economics, 131(3), 1415-1453.

Koedel, C., Mihaly, K., Rockoff, J. E. (2015). "Value-added Modeling: A Review." Economics of Education Review, 47, 180-195.

Kraft, M. A., Grace, S. (2016). Teaching for tomorrow's economy? Teacher effects on complex cognitive skills and social-emotional competencies. NBER Working Paper. Brown University, Providence, RI)

Jackson, C. K. (2018). "What do test scores miss? The importance of teacher effects on non–test score outcomes". *Journal of Political Economy*, 126(5), 2072-2107.

Teachers affect a variety of student outcomes through their influence on both cognitive and noncognitive skill. I proxy for students' noncognitive skill using non–test score behaviors. These behaviors include absences, suspensions, course grades, and grade repetition in ninth grade. Teacher effects on test scores and those on behaviors are weakly correlated. Teacher effects on behaviors predict larger impacts on high school completion and other longer-run outcomes than their effects on test scores. Relative to using only test score measures, using effects on both test score and noncognitive measures more than doubles the variance of predictable teacher impacts on longer-run outcomes.

Jackson, C. K. (2019). "The Full Measure of a Teacher: Using Value-Added to Assess Effects on Student Behavior". *Education Next*, 19(1), 62-69. [link]

> When students look back on their most important teachers, the social aspects of their education are often what they recall. Learning to set goals, take risks and responsibility, or simply believe in oneself are often fodder for fond thanks–alongside mastering precalculus, becoming a critical reader, or remembering the capital of Turkmenistan. The author argues while test scores are often the best available measure of student progress, they do not capture every skill needed in adulthood.

Opper, I. M. (2019). "Does Helping John help Sue? Evidence of Spillovers in Education". *American Economic Review*, 109(3), 1080-1115.

Bitler, M., Corcoran, S., Domina, T., Penner, E. (2019). "Teacher Effects on Student Achievement and Height: A Cautionary Tale" (No. w26480). National Bureau of Economic Research.

Lattimore, J. (2020). "A Critique of Value-Added Modeling: A Mixed Methods Approach".

2.2 Identifying Good Classroom Practices

Kane, Thomas J., Eric S. Taylor, John H. Tyler, and Amy L. Wooten. 2011. "Identifying Effective Classroom Practices Using Student Achievement Data." Journal of Human Resources 46 (3): 587-613.

2.3 Policy Strategies I: Improving Teacher Quality

Do these measures bring reliable information that could not be obtained using other teachers' characteristics?

Rules for Teacher Dismissal using Value-Added Distribution: Simulation Evidence

Hanushek E. A. (2011). "The economic value of higher teacher quality". *Economics of Education Review*. 30: 466–79.

Staiger D. O., Rockoff J.E. (2010). "Searching for Effective Teachers with Imperfect Information". *Journal of Economic Perspectives*. 24:97–117

Certification Programs

Kusumawardhani, P. N. (2017). Does teacher certification program lead to better quality teachers? Evidence from Indonesia. Education Economics, 25(6), 590-618.

Kraft, M. A., Brunner, E. J., Dougherty, S. M., Schwegman, D. (2018). Teacher accountability reforms and the supply of new teachers. Unpublished manuscript.

Providing Teachers' Value-Added to School Managers

Rockoff J.E., Staiger D.O., Kane T.J., Taylor E.S. (2012). "Information and Employee Evaluation: Evidence from a Randomized Intervention in Public Schools". *American Economic Review*, 102: 3184–3213

> We examine how employers learn about worker productivity in a randomized pilot experiment which provided objective estimates of teacher performance to school principals. We test several hypotheses that support a simple Bayesian learning model with imperfect information. First, the correlation between performance estimates and prior beliefs rises with more precise objective estimates and more precise subjective priors. Second, new information exerts greater influence on posterior beliefs when it is more precise and when priors are less precise. Employer learning affects job separation and productivity in schools, increasing turnover for teachers with low performance estimates and producing small test score improvements.

Professional Development Through In-Service Formal Training

Angrist J, Lavy V. (2001). "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools". *Journal of Labor Economics* 19(2): 343-69

Harris, Douglas N., and Tim R. Sass. 2011. "Teacher Training, Teacher Quality, and Student Achievement." Journal of Public Economics 95: 798-812.

Pianta RC. 2011. Teaching Children Well: New Evidence-Based Approaches to Teacher Professional Development and Training. Report, Center for American Progress

Rockoff J. 2008. "Does Mentoring Reduce Turnover and Improve Skills of New Employees? Evidence from Teachers in New York City." Manuscript.

Professional Development Through Online Off-the-Shelf Material

Jackson, K., Makarin, A. (2018). "Can Online Off-the-shelf Lessons Improve Student Outcomes? Evidence from a Field Experiment". *American Economic Journal: Economic Policy*, 10(3), 226-54.

> Many websites now warehouse instructional materials designed to be taught by teachers in a traditional classroom. What are the potential benefits of the new resources? We analyze an experiment in which we randomly give middle school math teachers access to existing high-quality, off-the-shelf lessons, and in some cases, support to promote their use. Teachers receiving access alone increased students' math achievement by a marginally significant 0.06 σ. Teachers who received access and support increased students' math achievement by 0.09 σ. Weaker teachers experience larger gains, suggesting that these lessons substitute for teacher skill or efforts. The online materials are more scalable and cost effective than most policies aimed at improving teacher quality, suggesting that, if search costs can be overcome, there is a real benefit to making high-quality instructional materials available to teachers on the Internet.

Professional Development Through Mentoring/Coaching of Young Teachers

The typical mentoring relationship is between (i) an experienced teacher who is high-performing broadly defined and (ii) a novice (newly hired) teacher. The mentor is sometimes a coworker, but often a former classroom teacher specializing in and trained in mentoring novices.

Kraft, M. A., Blazar, D., Hogan, D. (2018). "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-analysis of the Causal Evidence". *Review of Educational Research.* 88(4), 547-588.

> Teacher coaching has emerged as a promising alternative to traditional models of professional development. We review the empirical literature on teacher coaching and conduct meta-analyses to estimate the mean effect of coaching programs on teachers' instructional practice and students' academic achievement. Combining results across 60 studies that employ causal research designs, we find pooled effect sizes of 0.49 σ on instruction and 0.18 σ on achievement. Much of this evidence comes from literacy coaching programs for prekindergarten and elementary school teachers in the United States. Although these findings affirm the potential of coaching as a development tool, further analyses illustrate the challenges of taking coaching programs to scale while maintaining effectiveness. Average effects from effectiveness trials of larger programs are only a fraction of the effects found in efficacy trials of smaller programs. We conclude by discussing ways to address scale-up implementation challenges and providing guidance for future causal studies.

Professional Development Through Peer Observation

These policies use colleagues as observers to improve instruction. Our pairing design draws more teachers into the "mentor" role in a less-formal, perhaps lower-stakes, way. The mechanisms behind effects on student achivement can be: (i) skill accumulation of the observee; (ii) other changes in behavior and effort of the observee. Overall, the rationale of the policy is based on the first mechanism: teacher observers are suited to support the teacher observervees professional development — teachers may learn from coworkers. Notice that the costs of such a policy — assuming that there is no direct costs, i.e., teachers would not be paid to be observers — would involve the opportunity costs of using observers' time.

The steps of implementation and some interesting design choices documented by this literature seem to be:

1. How would observers and observers be chosen?
2.

Burgess, S., Rawal, S., Taylor, E. S. (2019). "Teacher peer observation and student test scores: Evidence from a field experiment in English secondary schools". *Journal of Labor Economics.*

> This paper reports on a field experiment in 82 high schools trialing a low-cost intervention in schools' operations: teachers working in the same school observed and scored each other's teaching. Students in treatment schools scored 0.07σ higher on math and English exams. Teachers were further randomly assigned to roles —observer and observee— and students of both types benefited, observers' students perhaps more so. Doubling the number of observations produced no difference in student outcomes. Treatment effects were larger for otherwise low-performing teachers.

Murphy, R., Weinhardt, F., Wyness, G. (2020). Who teaches the teachers? A RCT of peer-to-peer observation and feedback in 181 schools. *Unpublished Manuscript.*

Papay, J. P., Taylor, E. S., Tyler, J. H., Laski, M. E. (2020). "Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment using Teacher Performance Data. *American Economic Journal: Economic Policy*, 12(1), 359-88.

> We study a program designed to encourage learning from coworkers among school teachers. In an experiment, we document gains in job performance when high- and low-skilled teachers are paired and asked to work together on improving their skills. Pairs are matched on specific skills measured in prior evaluations. Each pair includes a target teacher who scores low in one or more of nineteen skills, and a partner who scores high in (many of) the target's deficient skills. Student achievement improved 0.12σ in low-skilled teachers' classrooms. Improvements are likely the result of target teachers learning skills from their partner.

Educational Technologies

Taylor, E. S. (2018). "New Technology and Teacher Productivity". *Unpublished Manuscript.*

> I study the effects of a labor-replacing computer technology on the productivity of classroom teachers. In a series of field-experiments, teachers were provided computer-aided instruction (CAI) software for use in their classrooms; CAI provides individualized tutoring and practice to students one-on-one with the computer acting as the teacher. In mathematics, CAI reduces by one-quarter the variance of teacher productivity, as measured by student test score gains. The reduction comes both from improvements for otherwise low-performing teachers, but also losses among high-performers. The change in productivity partly reflects changes in teachers' decisions about how to allocate class time and teachers' effort.

Ferman, B. Lima, Lycia Riva, Flavio (2020). "Artificial Intelligence, Teacher Tasks and Individualized Pedagogy". *Unpublished Manuscrip.*

This paper investigates how educational technologies that use different combinations of artificial and human intelligence are incorporated into classroom instruction, and how they ultimately affect learning. We conducted a field experiment to study two technologies that allow teachers to outsource grading and feedback tasks on writing practices of high school seniors. The first technology is a fully automated evaluation system that provides instantaneous scores and feedback. The second one uses human graders as an additional resource to enhance grading and feedback quality in aspects in which the automated system arguably falls short. Both technologies significantly improved students' essay scores in a large college admission exam, and the addition of human graders did not improve effectiveness in spite of increasing perceived feedback quality. Both technologies also similarly helped teachers engage more frequently on personal discussions on essay quality with their students. Taken together, these results indicate that teachers' task composition shifted toward nonroutine activities and this helped circumvent some of the limitations of artificial intelligence. More generally, our results illustrate how the most recent wave of technological change may relocate labor to analytical and interactive tasks that still remain a challenge to automation.

Unconditional Salary Increase

De Ree, J., Muralidharan, K., Pradhan, M., Rogers, H. (2018). "Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia." *The Quarterly Journal of Economics*, 133(2), 993-1039.

Pugatch, T., Schroeder, E. (2018). Teacher pay and student performance: evidence from the Gambian hardship allowance. Journal of Development Effectiveness, 10(2), 249-276.

Publicizing VA?

Bergman, P., Hill, M. J. (2018). The effects of making performance information public: Regression discontinuity evidence from Los Angeles teachers. Economics of Education Review, 66, 104-113.

Flexible Pay under Bargaining Replaces Seniority

Biasi, B. (2018). The labor market for teachers under different pay schemes (No. w24813). National Bureau of Economic Research.

Biasi, B., Sarsons, H. (2020). Flexible Wages, Bargaining, and the Gender Gap (No. w27894). National Bureau of Economic Research.

Teacher Specialization

– Fryer Jr, R. G. (2018). "The 'Pupil' Factory: Specialization and the Production of Human Capital in Schools. *American Economic Review*, 108(3), 616-56.

I conducted a randomized field experiment in traditional public elementary schools in Houston, Texas designed to test the potential productivity benefits of teacher specialization. The average impact of encouraging schools to specialize their teachers on student achievement is 0.11 standard deviations per year on a combined index of math and reading test scores. I argue that the results are consistent with a model in which the benefits of specialization driven by sorting teachers into a subset of subjects based on comparative advantage is outweighed by inefficient pedagogy due to having fewer interactions with each student, though other mechanisms are possible.

Teacher Pay for Performance

CAWLEY, J., J. J. HECKMAN, AND E. J. VYTLACIL (1999): "On Policies to Reward the Value Added by Educators," Review of Economics and Statistics, 81, 720–727. [885]

Goldhaber, D., Bignell, W., Farley, A., Walch, J., Cowan, J. (2016). Who chooses incentivized pay structures? Exploring the link between performance and preferences for compensation reform in the teacher labor market. Educational Evaluation and Policy Analysis, 38(2), 245-271.

Grönqvist, E., Hensvik, L., Thoresson, A. (2020). Teacher career opportunities and school quality (No.

2020: 2). Working Paper.

Hill, A. J., Jones, D. B. (2020). The Impacts of Performance Pay on Teacher Effectiveness and Retention Does Teacher Gender Matter?. Journal of Human Resources, 55(1), 349-385.

Imberman, S. A., Lovenheim, M. F. (2015). Incentive strength and teacher productivity: Evidence from a group-based teacher incentive pay system. Review of Economics and Statistics, 97(2), 364-386.

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. American Economic Review, 99(5), 1979-2011.

Leaver, C., Ozier, O., Serneels, P., Zeitlin, A. (2021). Recruitment, effort, and retention effects of performance contracts for civil servants: Experimental evidence from Rwandan primary schools. arXiv preprint arXiv:2102.00444.

Mbiti, I., Romero, M., Schipper, Y. (2019). Designing effective teacher performance pay programs: Experimental evidence from tanzania (No. w25903). National Bureau of Economic Research.

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., Rajani, R. (2019). Inputs, incentives, and complementarities in education: Experimental evidence from Tanzania. The Quarterly Journal of Economics, 134(3), 1627-1673.

Liebowitza, D. D. (2019). High-Stakes Teacher Evaluation for Accountability and Growth: Should Policy Treat them as Complements or Substitutes?.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." American Economic Review, 102(4): 1241- 78. Dee, T. and Wyckoff, J. "Incentives, Selection, and Teacher Performance: Evidence from IMPACT". NBER 19529, 2014 Figlio and Kenny "Individual Teacher Incentives and Student Performance", NBER working paper 12627, 2006. Glewwe, Ilias and Kremer "Teacher Incentives", NBER working paper 967, 2003. Lavy, V. "Evaluating the effect of teacher s group performance incentives on pupil achievement". The Journal of Political Economy, 110(6), 2002. Lavy, Victor. 2009. "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics."American Economic Review, 99(5): 1979-2011. Karthik Muralidharan and Venkatesh Sundararaman. Teacher performance pay: Experimental evidence from india. Journal of Political Economy, 119(1):pp. 39–77, 2011. ISSN 00223808. Jackson, C. K., Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. American Economic Journal: Applied Economics, 1(4), 85-108. Duflo, A., Kiessel, J., Lucas, A. (2020). External Validity: Four Models of Improving Student Achievement (No. w27298). National Bureau of Economic Research.

Sun, M., Loeb, S., Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues. Educational Evaluation and Policy Analysis, 39(1), 104-125.

Taylor, E. S. (2018). Skills, job tasks, and productivity in teaching: Evidence from a randomized trial of instruction practices. Journal of Labor Economics, 36(3), 711-742.

Kraft, M. A., Brunner, E. J., Dougherty, S. M., Schwegman, D. J. (2019). Teacher evaluation reforms and the supply and quality of new teachers. Unpublished paper, Brown University.

Class Size

2.4 Policy Strategies II: Reducing Discrimination

Do these measures bring reliable information that could not be obtained using other teachers' characteristics?