NBER WORKING PAPER SERIES

THE LABOR MARKET FOR TEACHERS UNDER DIFFERENT PAY SCHEMES

Barbara Biasi

Working Paper 24813
http://www.nber.org/papers/w24813

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2018, Revised January 2019

An earlier version of this paper, titled “Unions, Salaries, and The Market for Teachers: Evidence fromWisconsin,” was the first chapter of my dissertation at Stanford University: thanks to Caroline Hoxby, Luigi Pistaferri, Raj Chetty, and Petra Moser for their patient advising and support. I also thank Jason Abaluck, Jaime Arellano-Bover, Vittorio Bassi, Leah Boustan, Elizabeth Cascio, Judy Chevalier, Zo¨e Cullen, Will Dobbie, Mark Duggan, Liran Einav, Chao Fu, Robert Jensen, Andrew Johnston, Lester Lusher, Robert MacMillan, Davide Malacrino, Hani Mansour, Costas Meghir, Christopher Neilson, Petra Persson, Nicola Pierri, Dina Pomeranz, David Price, Juan Rios, Fiona Scott Morton, Isaac Sorkin, Edoardo Teso, Pietro Tebaldi, Caio Waisman, and seminar participants at various conferences and institutions for useful comments and conversations. Thanks to Sean P. Cottrell and Carl B. Frederick of the WDPI for their help in accessing and understanding the data. Joshua Ryan Petersen has provided excellent research assistance. Financial support from the National Academy of Education through a Spencer Dissertation Fellowship and from the Stanford Institute for Economics and Policy Research through a Leonard W. Ely and Shirley R. Ely Fellowship is gratefully acknowledged. All mistakes are my own. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Barbara Biasi. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Labor Market for Teachers Under Different Pay Schemes
Barbara Biasi
NBER Working Paper No. 24813
July 2018, Revised January 2019
JEL No. I20,J31,J45,J51,J61,J63

ABSTRACT

Compensation of most US public school teachers is rigid and solely based on seniority. This paper studies the labor market effects of a reform that gave school districts in Wisconsin full autonomy to redesign teacher pay schemes. Following the reform, some districts switched to flexible compensation and started paying high-quality teachers more. Teacher quality increased in these districts relative to those with seniority pay due to a change in workforce composition and an increase in effort. I estimate a structural model of this labor market to investigate the effects of counterfactual pay schemes on the composition of the teaching workforce.

Barbara Biasi
Yale School of Management
165 Whitney Avenue
New Haven, CT 06520
and NBER
barbara.biasi@yale.edu

A online appendix is available at http://www.nber.org/data-appendix/w24813

1 Introduction

Teachers are a central input in the production of student achievement (Rockoff, 2004; Rivkin et al., 2005), and their impact persists into adulthood (Chetty et al., 2011, 2014b). Therefore, policies that attract and retain high-quality teachers to the profession have far-reaching consequences.[1] More attractive compensation packages are sometimes proposed as a tool to attract effective teachers. However, in most US public school districts, the setting of teacher pay does not allow for the payment of financial rewards for effectiveness. If allowed to determine pay in a more flexible way, could school districts improve the quality of their teaching workforce? This paper addresses this question by taking advantage of a reform to the collective-bargaining process for teachers in Wisconsin, and offers a comprehensive study of this labor market.

Studying the effects of changes in the structure of teacher pay on labor supply and demand is challenging due to a dearth in variation in pay practices among public school districts.[2] The vast majority of districts pay teachers according to similar lock-step schedules. This implies that all teachers with the same education degree and years of experience are paid identically, regardless of their effectiveness or the demand for their labor (Podgursky, 2006). In addition, these schedules are often very similar across all districts within a state, owing to pattern bargaining facilitated by the state's teachers' union. With salaries set in this rigid way, identifying the effects of changes in pay schemes is very difficult.

To investigate these questions, I exploit a rare source of variation in teacher pay. In 2011, the Wisconsin legislature unexpectedly passed Act 10, a law that discontinued collective bargaining requirements over teachers' salary schedules. Previously, teacher pay was determined solely by seniority and education, using schedules negotiated between each school district and its union. Act 10 allowed districts full autonomy to determine compensation without union consent and, importantly, allowed them to negotiate salaries with individual teachers.

Districts used the flexibility introduced by Act 10 in different ways. Hand-collected information from 2015 employee handbooks (documents listing district-specific workplace policies and procedures) reveals that approximately half of the districts quickly took advantage of their

[1]The monetary cost of recruiting, hiring and training replacement teachers has been estimated to be around $15,000 in 2000, which translates into a $2.25 billion cost for the US as a whole. Turnover can negatively affect student achievement both directly, by disrupting instruction (Boyd et al., 2008; Ronfeldt et al., 2013), and indirectly, by affecting the composition of the teaching workforce (Adnot et al., 2017).

[2]Other works have used other changes in teachers' labor markets to learn about teacher supply and demand. For example, Hensvik (2012) studies the effect of private vouchers, Jackson (2009) examines the influence of school de-desegregation, and Fitzpatrick and Lovenheim (2014) study the impact of an early teacher retirement policy.

new-found discretion and replaced seniority-based schedules with flexible salary schemes that allowed for pay differences among teachers with similar seniority. For ease of exposition, I refer to these districts as "flexible pay" (FP). The other half, which I call "seniority pay" (SP), maintained the use of seniority-based schedules as of the publication of their 2015 handbook.

Act 10 triggered significant differences in salaries among teachers in FP districts who would have been paid exactly the same amount under the pre-Act 10 regime. To shed light on what drove these differences, I test whether a teacher's pay is correlated with her value-added (VA), a widely used measure of effectiveness. Standard VA models are designed for datasets that allow to link teachers and students through classroom identifiers. As is true of most administrative datasets, however, data from Wisconsin only allow teachers and students to be linked up to the grade level. To estimate teacher effects in the presence of this data limitation, I develop a new estimator based on teacher turnover as in Rivkin et al. (2005). I validate this estimator using data from New York City teachers and students (which include classroom identifiers) and I show that, while less precise than standard VA, this estimator is an unbiased predictor of a teacher's effectiveness.

VA estimates and information on teachers' salaries indicate that, after Act 10, salaries of high-VA teachers rose more than did those of low-VA teachers in FP districts, but not in SP districts. This is an interesting finding in itself; school districts in Wisconsin do not use VA to evaluate teachers nor do they even calculate it. This suggests that districts can identify highly effective teachers and, when allowed, chose to reward them more.

The differences in teacher salaries that arose among districts after Act 10 could have changed teachers' incentives to work in a given district, and in turn, could have affected each district's workforce composition. To see this, consider for example the school districts of Appleton (FP) and Oshkosh (SP), which belong to the same commuting zone in Eastern Wisconsin. In 2010, prior to Act 10, a teacher with a Master's degree and 4 years of experience would earn \$47,000 per year in Appleton (with a minimum of \$46,500 and a maximum of \$49,000) and \$45,800 in Oshkosh (with a minimum of \$42,000 and a maximum of \$49,000), according to each district's pay schedule. After the reform, the highest salary a teacher with those attributes could earn in Appleton rose to \$68,000, while the lowest stayed at \$44,000. In Oshkosh, on the other hand, the same teacher would earn between \$39,000 and \$43,000. To the extent that salary was performance-based in Appleton, a simple Roy model (Roy, 1951) predicts that high-quality teachers would move from Oshkosh to Appleton, whereas low-quality teachers would move from Appleton to Oshkosh or leave the market altogether. Indeed, although in the four years prior to the reform only four teachers moved from Oshkosh to Appleton, in the four years after Act 10 ten teachers moved. Further, all of the movers ranked above median for VA, and experienced an average salary increase of 16 percent ($6,500) per year. By contrast, the five teachers who moved from Appleton to Oshkosh after Act 10 all ranked below median for VA, and did not experience any significant increase in salary.

Extending this analysis to all districts confirm these patterns: After Act 10, teachers with *ex ante* higher VA (measured using pre-Act 10 test scores) were 1.13 times more likely to move from SP to FP districts compared with lower VA teachers, and 44 percent less likely to exit. These labor-supply responses produced a 0.05-0.07 standard deviations increase in average teacher quality in FP districts, relative to SP districts. These results demonstrate, partly in contrast with previous studies (such as Hanushek et al., 2004), that higher pay *does* attract teachers in some contexts.

In addition to affecting the composition of teachers, rewarding their effectiveness may also influence teachers' effort in the classroom. To test this hypothesis I allow VA to vary before and after Act 10 for each teacher, and I estimate the FP-SP difference in this time-varying measure after Act 10 compared with before. I find that, overall, VA increased by 0.11 standard deviations in FP districts relative to SP. Approximately 40 percent of this increase is due to changes in teacher effort, whereas the remaining 60 percent is due to changes in composition.

Although the differences in pay schemes between FP and SP districts after Act 10 create a unique setting to learn about their effects on teachers' labor markets, this setting involves several challenges that bear mentioning. First, post-Act 10 pay schemes were chosen by district administrators and could therefore be related to a range of district-level observable and unobservable characteristics that directly affect outcomes. Second, Act 10 did not just change teachers' salaries; it also increased employees' contributions to pensions and health care, and it reduced the power of unions. Interpreting the post-Act changes in outcomes between FP and SP districts as the effect of changes in pay requires assuming that any pre-existing differences between these two groups of districts did not change after Act 10 and that the other provisions of the Act (which applied to all districts) affected FP and SP districts in the same way. In support of these assumptions, I perform a series of checks: 1) I show that FP and SP districts are on parallel pre-trends with respect to all my outcome variables; 2) I show that FP and SP districts also exhibit parallel trends in a variety of attributes, including funding, union influence, and student attributes; 3) I demonstrate that the results are not affected by controlling for a wide range of time-varying observable factors, or by estimation on a matched sample; and 4) I perform the bounding exercise of Altonji et al. (2005) to account for selection into FP on the basis of unobservables. This exercise shows that, at worst, time-varying unobservables explain 7 percent of the main effect, suggesting that the scope for bias is limited.

My findings indicate that the introduction of flexible pay in a subset of Wisconsin districts led to an improvement in the composition of the teaching workforce in these districts compared with the rest of the state. This compositional change is fairly small (0.05 standard deviations of VA), as movements and exits are rare events even after Act 10. The change could become more pronounced over time as more low-quality teachers leave the market from FP districts and more high-quality teachers enter the profession, especially if pay becomes more strongly correlated with teacher effectiveness. This scenario, however, assumes that SP districts stick with seniority pay in the medium and long run. What would happen if the same flexible pay scheme were introduced in all districts instead? The sorting and exiting patterns outlined so far are the combination of both demand and supply forces; it is therefore difficult to answer this question by simply extrapolating from these partial-equilibrium results.

To address the limitations of a reduced-form approach, I build and estimate a structural model of the teachers' labor market. Districts (the demand side) extend job offers to teachers (the supply side). These offers are characterized by salaries, modeled as an exogenous, district-specific function of seniority, education, and VA (I relax the binary FP-SP assumption and estimate salary functions at the district level). Teachers have preferences over a job's attributes (including salary). They review all the offers they receive and choose the one that maximizes their utility (or choose to exit the labor market). Districts decide which job offers to extend in order to maximize a payoff that depends on teachers' attributes, subject to a budget constraint (on the total wage bill) and a capacity constraint (on the total number of teachers they need to hire). Importantly, when making hiring decisions, districts take into account the fact that the probability that a given offer is accepted depends on both teachers' preferences *and* the offers made by all the other districts. This feature of the model allows supply and demand to match in equilibrium.

To identify the parameters of teacher supply, I exploit the differences in pay triggered by Act 10 combined with teachers' movements and exits. Demand parameters are identified by cross-district differences in budget and capacity constraints (which arise when teachers move out of or exit from the district), combined with a district's decision on how to fill its vacancies and how to allocate its budget.[3]

The model model permits separate identification of the impact of supply and demand on job matches in equilibrium. Supply estimates are particularly useful for policy: They can be used to compute the monetary value of non-wage job characteristics valued by teachers and to quantify, for example, how much more a certain teacher would have to be paid to be induced to teach in a different district. Estimates of the supply parameters show that teachers are attracted by higher salaries, dislike moving to far-away districts, and face significant moving costs.

I use the model to simulate how the composition of the teaching workforce would change under two alternative pay schemes. The first consists of one district increasing its salary/quality correlation (which I use as a proxy for "merit" or "quality" pay) and confirms the reduced-form findings. The second analyzes the introduction of quality pay in *all* districts, a more challenging question to study due to general equilibrium effects. Simulations show that this second scheme is associated with a much smaller increase in workforce quality compared with the first: When all districts reward seniority at the same rate, teachers have lower incentives to move across districts, and any compositional improvement is entirely driven by exits of low-quality teachers.

This exercise is useful to understand what would happen if all districts switched to flexible pay, a scenario that could easily arise in Wisconsin when SP districts realize that they are losing good teachers. It also suggests that the observed improvement in the composition of the teaching workforce experienced by FP districts might be short-lived, resulting in smaller long-term effects of a statewide change in pay schemes.

A few limitations of the model should be kept in mind when interpreting these results. First, the model does not explicitly allow for workers' decisions to *enter* the teaching profession, and it implicitly assumes that the quality of new teachers is constant over time and unaffected by the reform. In the medium run, a change in teacher pay could fundamentally alter the selection of new teachers in FP and SP districts (Dolton, 1990; Hoxby and Leigh, 2004; Rothstein, 2014; Kraft et al., 2018).[4] Second, the model does not incorporate changes in effort. Even under the second counterfactual, where compositional gains are small, the benefits to schools and students could

[3]Section 7.3 discusses identification of the model's parameters more in depth.

[4]Dolton (1990) emphasizes the importance of earnings growth in the decision to become a teacher. Hoxby and Leigh (2004) show that the increase in wage compression that followed the rise in unionization explains most of the decline in entry of high-performing teachers in the US since 1960. Rothstein (2014) suggests that higher salaries and lower tenure rates can improve the supply of new teachers. Lastly, Kraft et al. (2018) suggest that the introduction of teacher evaluation policies led to a decline in the supply of new teachers. A simple analysis of the selectivity of college degrees for new teachers (as a proxy for teaching quality) does not show evidence of changes in the composition of new teachers after 2011.

be large if all teachers respond to the change in pay scheme by exerting more effort.

This paper makes three main contributions. First, it exploits newly-available, large-scale variation in pay schemes to estimate teachers' labor supply and demand. Previous studies have been limited to small bonuses awarded on top of regular pay (Hanushek et al., 2004; Clotfelter et al., 2008; Dee and Wyckoff, 2015), limited cross-sectional variation in salaries (Stinebrickner, 2001; Boyd et al., 2013), and across-the-board salary increases (Figlio, 2002).[5]

This paper can also be seen as an exploration, in the personnel economics tradition, of how pay affects selection and incentives of a particularly important class of workers (Lazear, 2000a,b; Bandiera et al., 2005; Abramitzky, 2009; Khan et al., 2015). Financial incentives for teachers have been shown to have a significant impact on student achievement outside the US.[6] Plans implemented in the US, however, have yielded mixed results (see Jackson et al., 2014; Neal et al., 2011, for a review).[7] In addition, this paper provides new evidence that school districts are willing to compensate high VA teachers when given the opportunity to do so and that teachers respond to these incentives by exerting more effort in the classroom (Imberman and Lovenheim, 2015; Brehm et al., 2017).

Lastly, this paper is one of the first to study the effects of a recent *decline* in union powers. Most studies of teachers' unions have analyzed increases in unionization (Eberts and Stone, 1987; Hoxby, 1996; Lovenheim, 2009).[8] Reducing union powers, however, does not necessarily have symmetric effects. The effects of a decline in unionization on teachers' labor markets are particularly interesting in the aftermath of *Janus v. AFSCME*, as more states could be affected in the future.

The rest of the article is organized as follows. Section 2 outlines the institutional framework and describes Act 10. Section 3 presents the data. Sections 4, 5, and 6 describe the empirical findings on salaries, the composition of the workforce, and teachers' effort respectively. Section 7 describes the structural model, and Section 8 illustrates the results from simulations of counterfactual pay schemes. Section 9 concludes.

[5] Willén (2018) studies the introduction of individual wage bargaining for Swedish public school teachers. Unlike Act 10, however, this policy change did not lead to pay differences among high-quality and low-quality teachers. As a result, it had no effects on workforce composition (in terms of demographics) or student outcomes.

[6] This literature includes studies conducted in India (Muralidharan and Sundararaman, 2011; Duflo et al., 2012), Israel (Lavy, 2002), England (Atkinson et al., 2009), and Kenya (Glewwe et al., 2010).

[7] Although some studies have found that teacher performance pay has positive effects on student test scores in the US (Ladd, 1999; Figlio and Kenny, 2007; Sojourner et al., 2014; Imberman and Lovenheim, 2015; Dee and Wyckoff, 2015; Brehm et al., 2017), others have shown that such incentives are ineffective at boosting achievement (Dee and Keys, 2004; Figlio and Kenny, 2007; Springer et al., 2011; Goodman and Turner, 2013; Fryer, 2013).

[8] Notable exceptions are Han (2016), Litten (2016), and Roth (2017), who study the effects of recent episodes of de-unionization on outcomes such as teacher turnover, teacher salaries, retirement, and student achievement.

2 Teacher Compensation Before and After Act 10

In most US public school districts, teacher salaries are determined using "steps-and-lanes" salary schedules that express pay as a function of years of experience and highest education degree (Podgursky, 2006). Appendix Figure A1 shows an example of a salary schedule. Movements along its "steps" (rows, which correspond to experience levels) and "lanes" (columns, which correspond to education degrees) are associated with an increase in pay.

In states with collective bargaining (CB) for public sector employees, these schedules are negotiated between school districts and teachers' unions, while in states without CB they are typically determined at the state level (e.g. Georgia).[9] CB agreements usually prevent districts from adjusting pay at the individual level: Experience and education are the only determinants of salaries and pay is unrelated (at least directly) to teacher effectiveness (Podgursky, 2006).

2.1 Wisconsin's Act 10

In 1959, Wisconsin became the first state to introduce CB for public sector employees (Moe, 2013). Since then, teachers' unions have gained considerable power and have been involved in negotiations with school districts over key aspects of a teaching job.[10] Until 2011 teacher salaries were set using a schedule, which was part of each district's CB agreement.

On June 29, 2011, the state legislature passed the Wisconsin Budget Repair Bill, which came to be known as Act 10. Intended to address a projected $3.6 billion budget deficit through cuts in public sector spending, Act 10 introduced a number of provisions, enforced onto all school districts and their employees. First and most importantly, Act 10 limits the scope of salary negotiations to base salaries (whose annual growth rate is limited to the rate of inflation), hence preventing unions from negotiating the salary schedule. Second, it requires unions to hold yearly recertification elections, limits the validity of CB agreements to one year, and prohibits automatic collection of union dues from employees' paychecks.[11] Third, it raised employee

[9] As of 2014, 49 percent of public school teachers belonged to at least one union (Bureau of Labor Statistics, 2014). Schools are unionized on a district-by-district basis. While CB is not a constitutional right for public sector teachers, at the time of writing this right was granted by all states except Georgia, North Carolina, South Carolina, Texas, and Virginia.

[10] 424 public school districts in Wisconsin typically serve either one city or one or more towns and villages. They operate public schools, hire teachers, and allocate teachers to schools. Each district enrolls an average of 1,900 students. Sixteen urban districts enroll 15,000 students on average (with Milwaukee Public Schools enrolling 67,000 students and the Madison Metropolitan School District enrolling 26,500 students), 63 suburban districts enroll 3,000 students, and 344 rural districts enroll 1,000.

[11] In 26 right-to-work states, including Texas, Florida, and Wisconsin, teachers who choose not to join the union are not required to pay monthly dues, despite being covered by collective-bargaining (CB) agreements. In all other states (including California, New York, and Illinois) non-union teachers are also required to pay a fee to the union as

contributions to the pension fund (from 0 to 5.8 percent of wages) and to health insurance premia (from 0 to 12.6 percent), and it required districts to actively search for the most cost-effective health care plans. Lastly, it reduced state aid to school districts and decreased their revenue limit.[12]

2.1.1 Act 10 and Teacher Salaries: Flexible Pay vs. Seniority Pay

With salary schedules no longer allowed in CB agreements, Wisconsin districts have gained the possibility to determine teacher pay more flexibly. In particular, they are now allowed to reward teachers for attributes not directly compensated by standard schedules and to adjust salaries on an individual basis without union consent.

Although the provisions of Act 10 applied to all school districts, the way in which districts reacted to their new freedom over teacher pay-setting varied: As of 2015, approximately half of all districts were still setting pay using a schedule only based on experience and education, whereas the remaining half had discontinued the use of such a schedule. To characterize each district's post-Act 10 pay regime I collected districts' employee handbooks, documents listing the duties and rights of all teachers. Before Act 10, all handbooks contained a schedule; after Act 10, only some of them do. I classify all districts whose 2015 handbooks included a schedule (and did not mention any other types of bonuses or increments) as seniority-pay (SP) and all the remaining districts as flexible-pay (FP). More information on the handbooks is contained in Section 3.

The Racine Metropolitan School District, one of the state's largest urban districts, is an example of a SP district: Its 2015 handbook contains a seniority-based schedule (Appendix Figure A1).[13] The handbook specifies that both a teacher's initial placement on the schedule and movements along steps and lanes are to be determined solely on the basis of seniority and academic credentials.

The Green Bay Area Public School District, the fifth largest in Wisconsin, is an example of a FP district. Its 2015 handbook does not contain a schedule, and it explicitly states that "[t]he District will determine the starting salary for a new employee."[14] While the handbook mentions

a condition of employment. Union membership dropped by nearly 50 percent in Wisconsin in the 5 years after the passage of Act 10. See D. Belkin and K. Maher, *Wisconsin Unions See Ranks Drop Ahead of Recall Vote*, The Wall Street Journal. Retrieved from https://www.wsj.com/articles/SB10001424052702304821304577436462413999718.

[12]This last provision was included in Act 32 of July 1, 2011, which amended some provisions of Act 10. Revenue limits are the maximum level of revenues a district can raise through general state aid and local property taxes.

[13]See the Racine School District website for the most recent version of its teacher salary schedule.

[14]See the Green Bay Area Public School District website for the most recent version of its employee handbook.

the possibility that teachers' salaries might increase in steps over time, no indication is given that such steps will be solely linked to seniority and education. The handbook also specifies that "[a]n employee may be held to the previous year's step for less than satisfactory performance." This language, common among FP districts, indicates that the district retains the autonomy to set teacher salaries on an individual basis and to adjust them every year as it sees fit.

2.1.2 Comparing FP and SP Districts: Identification Challenges

Decisions over post-Act 10 pay schemes were made by school district administrators (such as superintendents and school board members). Possible drivers of this decision include fiscal concerns, the desire to compensate high-quality teachers or to preserve teachers' morale, and the increased pressure to compete with other districts for talented teachers (Kimball et al., 2016). This decision could be correlated with time-varying observable and unobservable characteristics of the school districts. For example, districts with an *ex ante* lower-quality workforce or lower student achievement could choose to switch to FP in an attempt to improve, or wealthier districts could switch because they can afford to spend more on salaries. Differences in observable and unobservable characteristics pose a challenge for the identification of a causal effect of changes in pay structures across districts if they vary over time; interpreting any post-Act 10 comparison in outcomes between FP and SP districts as the causal effect of pay structures requires assuming that these differences remained constant after 2011.

A simple comparison of observable characteristics across FP and SP districts before Act 10 shows that FP districts served a lower share of low-SES students, paid higher salaries to teachers and principals, had less experienced teachers, and are more likely to be located in suburban areas with higher property values (Table 1 and Appendix Table A1). These differences, however, did not change after Act 10 (Appendix Figure A2). Appendix Table A1 estimates difference-in-differences models that compare these observables across FP and SP districts over time.[15] Estimates for *FP * post*, which capture the change in the FP-SP difference after Act 10, are all small and insignificant.[16] While not necessarily informative of changes in unobservables, these tests rule out the possibility that any differences in outcomes between FP and SP districts after Act 10 could be driven by changes in observables over time.

[15] Appendix Table A1 shows estimates of α and β in $X_{dt} = \alpha FP_d + \beta FP_d * post_t + \gamma post_t + \varepsilon_{dt}$, where X_{dt} includes a range of district attributes, FP_d equals 1 for FP districts, and $post_t$ equals 1 for years after Act 10. Estimates of α capture *ex ante* differences in attributes between FP and SP districts, while estimates of β capture the change in this difference after Act 10.

[16] Trends in these observables are shown in Appendix Figures A2, A3, A4, and A5.

A possible confounder that deserves mention is districts' management. A switch to FP could correlate with better managerial capabilities, as it requires quantifying hard-to-measure teacher attributes (such as performance). More generally, superintendents and principals of FP districts may have different perceptions of what constitutes a "fair" compensation scheme or may differentially value teachers' characteristics (such as experience or quality).[17] This could translate into different managerial practices and could have a direct effect on the outcomes of interest. Superintendents and principals, however, have a very limited scope for differentiating their practices, as most teachers' duties and rights are strictly regulated by CB agreements (both before and after Act 2011). Even if managerial practices were different between FP and SP districts, these differences would only be problematic if they arose at the same time as Act 10. Furthermore, the data indicate that superintendents, principals, and other managerial staff are observationally similar across FP and SP districts (Appendix Tables A1 and A2). This suggests that the decision to switch to FP was likely influenced by idiosyncratic preferences of local leaders rather than by systematic differences in management.

Other Provisions of Act 10. In addition to changing teachers' pay schemes, Act 10 included a number of other provisions, some of which directly affected teachers. In theory, these provisions were applied uniformly across the state. In practice, however, they could have differentially impacted districts depending on districts' characteristics. For example, the weakening of teachers' unions could have affected a district's pay scheme and teacher sorting differently depending on *ex ante* union strength. Similarly, increases in employees' contributions to health care and retirement plans and a reduction in state aid could have led districts to reallocate their budget across different items and to offer health care plans of different quality, and this could have directly affected teacher sorting (D'Andrea, 2013).

Appendix Table A1 shows that expenditure on different budget items and measures of union strength (such as indicators for whether the district had a union election in each year and whether the union managed to recertify) were very similar across FP and SP districts, both before and after Act 10.[18] Furthermore, the bounding exercise of Altonji et al. (2005) indicates only a small role for unobservables (Section 5). This exercise allows me to quantify the role of unobservables under the assumption that the variation in the outcome variable that is driven by

[17] For example, if districts place the highest priority on seniority, they might be more likely to stay SP. If they place a higher value on attributes not rewarded by a standard schedule, they might be more likely to switch to FP.

[18] Appendix Figure A4 shows the 2012–2016 trends in the share of FP and SP districts which held a union election in each year and in which the union successfully re-certified. Trends in various budget items across FP and SP districts are shown in Appendix Figure A3.

unobservables and the variation driven by observables have the same relationship with the type of post-Act 10 pay structure. Overall, these tests suggest that the bulk of the difference in post-Act 10 outcomes can be attributed to differences in pay schemes, as opposed to other provisions of the Act.

Table 1: Summary Statistics, Wisconsin Districts, 2007–2011

	(1) Full sample			(2) Matched sample		
	FP	SP	Difference	FP	SP	Difference
enrollment	2929.7	3108.3	-178.6	2929.7	2789.3	140.4
			(761.8)			(505.5)
Black students	0.028	0.032	-0.0035	0.028	0.030	-0.0017
			(0.0083)			(0.0091)
Hispanic students	0.053	0.049	0.0037	0.053	0.055	-0.0020
			(0.0076)			(0.0096)
disadvantaged students	0.28	0.32	-0.041**	0.28	0.30	-0.023
			(0.019)			(0.023)
math scores (sd)	0.18	0.065	0.12	0.18	0.14	0.043
			(0.080)			(0.10)
teacher salary ($)	52659.5	50517.9	2141.6***	52659.5	51485.6	1173.8
			(709.4)			(911.2)
teacher experience (yrs)	14.9	15.6	-0.67**	14.9	15.3	-0.43
			(0.27)			(0.30)
teachers w/ BA	0.46	0.49	-0.031	0.46	0.48	-0.020
			(0.020)			(0.023)
teachers w/ Master	0.53	0.50	0.028	0.53	0.51	0.016
			(0.020)			(0.023)
teachers w/ PhD	0.0013	0.0010	0.00028	0.0013	0.00066	0.00067
			(0.00048)			(0.00054)
urban district	0.069	0.074	-0.0051	0.069	0.070	-0.0015
			(0.035)			(0.042)
suburban district	0.29	0.20	0.097*	0.29	0.25	0.049
			(0.057)			(0.074)
property values p.p. ($)	803196.9	604887.4	198309.5**	803196.9	630296.6	172900.3
			(91952.3)			(130388.3)
value-added	-0.087	-0.13	0.048	-0.087	-0.071	-0.015
			(0.064)			(0.075)
expenditure p.p ($)	15126.3	15531.1	-404.7	15126.3	14734.9	391.4
			(366.7)			(400.7)
Nr. of districts	102	122	224	122	56	178

Notes: Means, differences in means, and standard errors (in parentheses) of district-level characteristics in FP and SP districts (columns 1) and in matched FP and SP districts (columns 2). The FP subsample includes 102 districts; the SP subsample includes 122 districts. The subsample covers 83 percent of the total student population. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts and includes 56 districts.

Accounting for FP-SP Differences and Other Provisions of Act 10: Controls and a Matched Sample. To check for pre-Act 10 differential trends in outcome variables across FP and SP districts (which could indicate the presence of time-varying unobservables), all my empirical tests are accompanied by time-varying estimates of the parameters of interest. In addition, I control for the *ex ante* differences in observables in a flexible way, interacting their pre-Act 10 (2009–2011) averages with year fixed effects.[19] Lastly, I complement my results with those obtained using a matched sample, constructed to smooth the small *ex ante* differences in observables between FP and SP districts (Table 1).[20] I build the sample using nearest-neighbor matching with replacement (Abadie and Imbens, 2011). I match each FP with a SP district on the basis of 2009–2011 district attributes, including enrollment, the share of low-SES students, average teacher salaries (for all teachers and for those with less than five years of experience), the share of teachers with less than three years of experience, the share of teachers with more than 20 years of experience, the share of teachers with a master's degree, the location of the district (urban, suburban, rural), property values, expenditure, and state aid per pupil.[21] The final sample contains 102 FP and 56 SP districts (Table 1, column 2).[22]

3 Data and Measurement

The main data set contains information on the universe of Wisconsin teachers, linked with student test scores to calculate teacher VA. I combine it with information on post-Act 10 salary structures for each district, drawn from employee handbooks. Lastly, I use school- and district-level characteristics as controls and to construct the matched sample. Data are reported by academic year, referenced using the calendar year of the spring semester (e.g. 2007 for 2006-07).

Teacher Data. I draw information on the population of Wisconsin teachers from the *PI-1202 Fall Staff Report - All Staff Files* for the years 2007–2015, made available by the Wisconsin Department of Public Instruction (WDPI). These files contain information on all individuals employed by the WDPI in each year and include personal and demographic information, education, years of teaching experience, and characteristics of job assignments (including total salary, grades and subject taught, full-time equivalency (FTE) units, and school and district identifiers). I restrict

[19] Appendix Figure A2 shows no evidence of differential pre-trends in these observables across the two groups of districts in the years leading to Act 10.

[20] Despite these differences in means, Appendix Figure A13 shows that the distributions of these characteristics are very similar across the two groups of districts.

[21] Since the sample is with replacement, matched SP districts are counted multiple times if they serve as controls for more than one district. Appendix Table A3 shows estimates of the probit model underlying the matching procedure.

[22] Appendix Figure A14 shows that the distributions of a set of district characteristics across FP and matched SP districts appear very similar; a Kolmogorov-Smirnov test fails to reject the null hypothesis of equality in the distribution of all variables at the 5 percent confidence level.

the sample to non-substitute, tenured teachers working in FP and SP districts.[23] Salaries are expressed in 2015 dollars and in FTE units.[24]

School and District Data. School-level data from the Wisconsin Information System for Education (WISE) include total enrollment and the share of economically disadvantaged, Black, and Hispanic students. District-level covariates include equalized property values from the WDPI (used to calculate property tax levies) and indicators for whether the district is located in an urban, suburban, or rural area.[25] Budget data from the WDPI include revenues by source and expenditures by item, for all districts and for the years 2008–2015. Lastly, information on union election outcomes is from the records of the Wisconsin Employment Relation Commission (WERC).

Student Test Scores and Demographics. Student-level data include math and reading test scores from the Wisconsin Knowledge and Concepts Examination (WKCE, 2007–2014) and Badger test (2015–2016), for all students in grades 3 to 8, as well as demographic characteristics such as gender, race and ethnicity, socio-economic (SES) status, migration status, English-learner status, and disability. The WKCE was administered in November of each school year, whereas the Badger test was administered in the spring. To account for this change, for the years 2007–2014 I assign each student a score equal to the average of the standardized scores for the current and the following year.

Employee Handbooks and Salary Schedules. I collected information on districts' pay schemes from their 2015 employee handbooks, available for 224 out of 422 districts (including seven high school districts), which enroll approximately 83 percent of all students.[26] I classify a district as SP for the entire post-Act 10 period if its 2015 handbook contains a salary schedule and does not mention rewards for performance or merit and as FP otherwise. If a schedule is published but bonuses linked to performance are mentioned, the district is classified as FP.[27]

[23] I exclude long- and short-term substitute teachers, teaching assistants and other support staff, and contracted employees since salaries for these workers are calculated differently from those of permanent teachers.

[24] Due to evident mistakes in the reporting of salary information, I discard information for teachers in the school district of Kenosha, as well as for those in the school district of Milwaukee for the year 2015.

[25] These variables are based on the US Census urban-rural classification.

[26] Handbooks are published on each district's website. Unclassified districts (i.e., those for which handbooks are not available) either do not have a website or do not make their handbook public. Appendix Table A6 compares FP and SP districts with unclassified districts. The latter are smaller, enroll more disadvantaged students, pay lower salaries, and are disproportionately located in rural areas.

[27] It is possible that districts classified as FP did not change pay scheme immediately after the passage of Act 10, but after a few years. By the same token, it is possible that districts classified as SP as of 2015 switched to FP after 2015. For this reason I end my analysis in 2015.

3.1 Measurement: Teacher Value-Added

I measure teacher quality using value-added (VA), defined as the teacher's effect on test scores conditional on other determinants of achievement (such as past test scores, student demographics, and school fixed effects; Hanushek, 1971; Rockoff, 2004; Rivkin et al., 2005; Chetty et al., 2014a). Albeit not a perfect measure of talent (Rothstein, 2010; Kraft, 2017; Jackson, 2018), VA is generally considered a good signal of a teacher's effectiveness (Kane and Staiger, 2008; Chetty et al., 2014a).

VA is usually estimated using datasets that make it possible to link teachers to the pupils they taught through classroom identifiers. Information on students' and teachers' classroom was not maintained by the WDPI before 2017. This implies that I can link a teacher to all the students enrolled in her school and grade in a given year, but not to the specific students she taught.

To obtain a measure of teacher effectiveness in the presence of this data limitation I leverage the identification approach of Rivkin et al. (2005), who face the same issue using data from Texas. I then combine this approach with the estimation method of Chetty et al. (2014a). The starting point is the following model of achievement:

$$A^*_{kt} = \beta X_{kt} + \nu_{kt} \tag{1}$$
$$\text{where } \nu_{kt} = \mu_{i(kt)t} + \theta_{c(kt)} + \varepsilon_{kt},$$

A^*_{kt} is a standardized measure of test scores for student k in year t, X_{kt} is a vector of student and school-specific controls, and $i(kt)$ denotes student k's teacher in t. VA is the estimate of $\mu_{i(kt)t}$, the teacher-specific component of test score residuals. Chetty et al. (2014a) use the following estimator:

$$\hat{\mu}_{it} = \sum_{m=t-l}^{t-1} \hat{\psi}_m \bar{A}_{im} \tag{2}$$
$$\text{where } \bar{A}_{it} = \frac{1}{N_{it}} \sum_{k:i(kt)=i} A_{kt} \text{ and } A_{kt} = A^*_{kt} - \hat{\beta} X_{kt}, \tag{3}$$

$\hat{\beta}$ is the OLS estimate of β in (1), N_{it} is the number of students taught by i, and the coefficients $\hat{\psi} = (\hat{\psi}_{t-l} ... \hat{\psi}_{t-1})'$ are selected to minimize the mean-squared error of the forecast of test scores, with a procedure analogous to OLS:

$$\hat{\psi} = \arg \min_{\{\psi_{t-l},\ldots,\psi_{t-1}\}} \sum_i \left(\bar{A}_{it} - \sum_{s=t-l}^{t-1} \psi_s \bar{A}_{is} \right)^2 \tag{4}$$

Essentially, $\hat{\mu}_{it}$ is the best linear predictor of $\bar{A}_{it}$, the average test score residuals of teacher i's students in year t, given test score residuals from the previous l years. When teachers and students can only be linked up to the grade level, constructing $\bar{A}_{it}$ is not possible. The best approximation is an average of test score residuals for all students in teacher i's grade and school in year t:

$$\bar{A}^g_{it} = \frac{1}{N_{gst}} \sum_{k:g^e(it)=g^p(kt),s^e(it)=s^p(kt)} A_{kt}$$

where $g^e(it)$ ($s^e(it)$) is the grade (school) where teacher i teaches in year t, $g^p(kt)$ ($s^p(kt)$) is the grade (school) attended by student k in year t, and N_{gst} is the number of students in grade g and school s in year t. The estimator then becomes $\hat{\mu}_{it} = \sum_{m=t-l}^{t-1} \hat{\psi}_m \bar{A}^g_{im}$, where $\hat{\psi} = \arg\min_{\{\psi_{t-l},\ldots,\psi_{t-1}\}} \sum_i \left(\bar{A}^g_{it} - \sum_{s=t-l}^{t-1} \psi_s \bar{A}^g_{is} \right)^2$.

Identification. How can we identify a teacher's effect if we cannot precisely link her to the students she taught? The identification argument follows Rivkin et al. (2005) and relies on teacher turnover across grades and/or across schools. In the absence of turnover, all teachers in the same grade and school would be assigned the same average residual every year, and distinguishing their individual effects would be impossible. With multiple years of data and in the presence of turnover, however, teacher switches across schools, or within schools and across grades, allow me to isolate the effect of the individual teacher through the comparison of test score residuals before and after her arrival in a given grade and school. Importantly, teacher turnover makes it possible to more precisely identify not only the effects of the teacher who switches, but also the effects of all other teachers in her same grade and school at any point in time.[28] Appendix B.1 illustrates the identification argument with a simple example.

Limitations of VA Estimates and a Validation Exercise. Because students are linked to teachers at the grade-school level, and because not all teachers switch grade or school, VA of a teacher could also be a function of test scores of students she never taught. This will introduce measurement error in the estimates. In addition, VA of teachers who are always in the same grade-school

[28]The aggregation of test scores at the grade level also overcomes one of the most problematic form of selection, which occurs within schools and grades and across classrooms (Rivkin et al., 2005). The (forced) use of grade-school estimates circumvents this form of selection and is in practice equivalent to an instrumental variable estimator based on grade rather than classroom assignment.

will never be separately identifiable (and each of them would be assigned an average of their true effects).

Classical measurement error would undermine the efficiency of the parameter estimates (when used as the dependent variable) or generate attenuation bias (when used as an explanatory variable). This noise, however, can be even more problematic when it is non-classical, i.e., correlated with other variables in a model. This could happen, for example, if estimates are more precise for teachers who move and if movements are not random (e.g. they could be correlated with a district's pay scheme).

To assess the performance of my measures relative to standard estimates, I use data from New York City (NYC) teachers and students, which include classroom links. I estimate VA using the standard approach (which exploits classroom links, CL hereafter) as well the approach described above, which links students to teachers on the basis of grade and school (GL hereafter). A comparison of these two sets of estimates reveals the following facts.

1. Although less precise than CL, GL still explains a substantial portion of the variance in test scores. In the NYC data, the standard deviation of VA for math teachers is 0.19 for CL and 0.11 for GL (Appendix Figure B1). In the Wisconsin data, the standard deviation of GL is 0.08 for math (Appendix Figure B2).

2. GL is a forecast-unbiased estimator of CL. Forecast bias can be defined as $f = 1 - \gamma$ in the regression $\hat{\mu}_i^{CL} = \alpha + \gamma\hat{\mu}_i^{GL} + \chi_i$. Appendix Figure B3 shows the linear relationship between CL and GL; the estimated bias (one minus the slope of the fit line) is equal to 0.05 and it is indistinguishable from zero.

3. GL is a forecast-unbiased estimator of a teacher's future student achievement. Using teacher switches as a quasi-experiment, as in Chetty et al. (2014a), yields an estimated bias of 0.05 for GL (Appendix Table B1).

4. The difference between CL and GL is uncorrelated with student or teacher observables (including the probability of turnover) in the NYC data (Appendix Table B2), mitigating concerns for non-classical measurement error. Clearly, this test cannot be performed with the Wisconsin data. While this result plausibly applies for VA estimated before Act 10 (when teachers' turnover was arguably more random), it might not hold after 2011 (when sorting could happen in response to a change in pay). I discuss the implications of non-classical measurement error along with the description of my results (Section 6).

VA estimates are available for 20,370 teachers of math and reading in grades 4 to 8, including the final sample of 16,862 tenured teachers in 98 FP and 119 SP districts serving elementary and middle schools (see Appendix Table A4 for a summary).[29]

My empirical analyses use two measures of VA. The first is *ex ante* VA, calculated as the average teacher effect for the years 2007–2011. This measure is constructed to parse out any effects of Act 10 on effort and to focus on selection. The second is a time-varying measure, allowed to vary before and after Act 10 for each teacher and used to study changes in effort. By construction, *ex ante* VA is only available for the subsample of teachers who were already in the system before 2011. While this does not affect the estimation of VA *per se* (which uses information on all teachers in a given grade and year), analyses of teacher selection will be based on this (possibly selected) subset of teachers. Appendix Table A5 compares teachers with and without *ex ante* VA on the basis of observables. Perhaps not surprisingly, teachers with *ex ante* VA have higher experience. Their *ex post* VA, however, is not statistically different from that of teachers without *ex ante* VA.

4 Salary Responses to Act 10

Act 10 gave districts considerable flexibility over the design of teacher pay. I start my empirical analysis by quantifying how salaries changed in FP and SP districts in the aftermath of the reform. I focus on two metrics: the degree of pay dispersion among teachers with similar experience and education and the relationship between salaries and teacher quality, measured with VA. Appendix Figure A6 plots the full distribution of salaries in FP and SP districts between 2007 and 2015.

Dispersion in Salaries. Figure 1 shows median salaries and 90-10 percentile ranges by two-year experience classes and for teachers with a master's degree in two large and comparable urban districts: Racine (top panel), a SP district, and Green Bay (bottom panel), a FP district.[30]

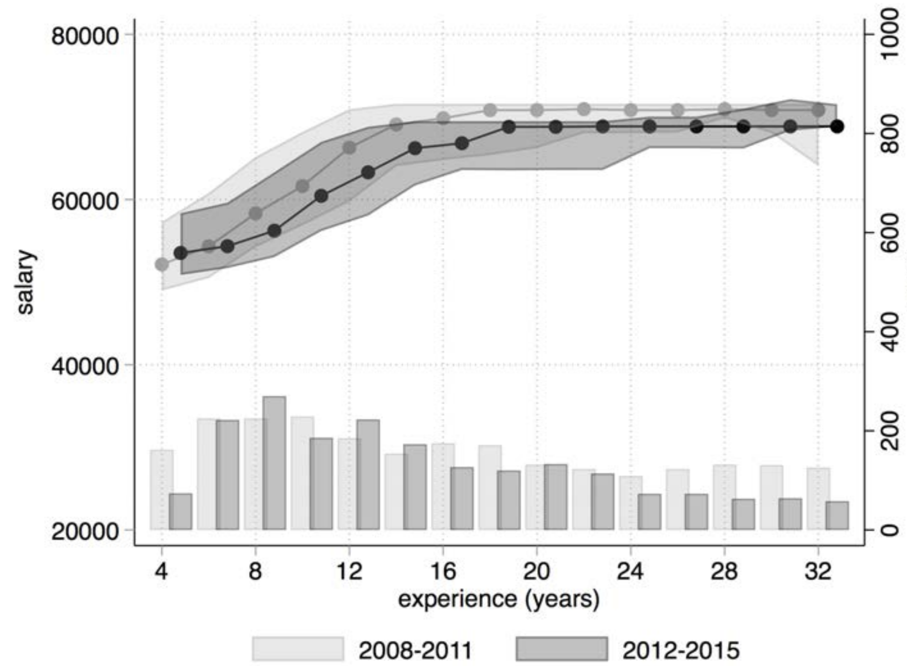
Before Act 10, the salary distribution was very similar across the two districts (although base salaries were lower in Green Bay). The median salary for teachers with five or six years of experience was equal to $54,337 in Racine (with a 90-10 percentile range of $10,308) and to $47,799 in Green Bay (with a range of $5,962). For teachers with 11 or 12 years of experience the median was $66,285 in Racine (with a range of $11,205) and $59,452 in Green Bay (with a range

[29] VA estimates are not available for teachers in seven high school districts, since standardized test scores are not administered in high school. Although most of the empirical analysis is restricted to tenured teachers, VA is calculated for all teachers.

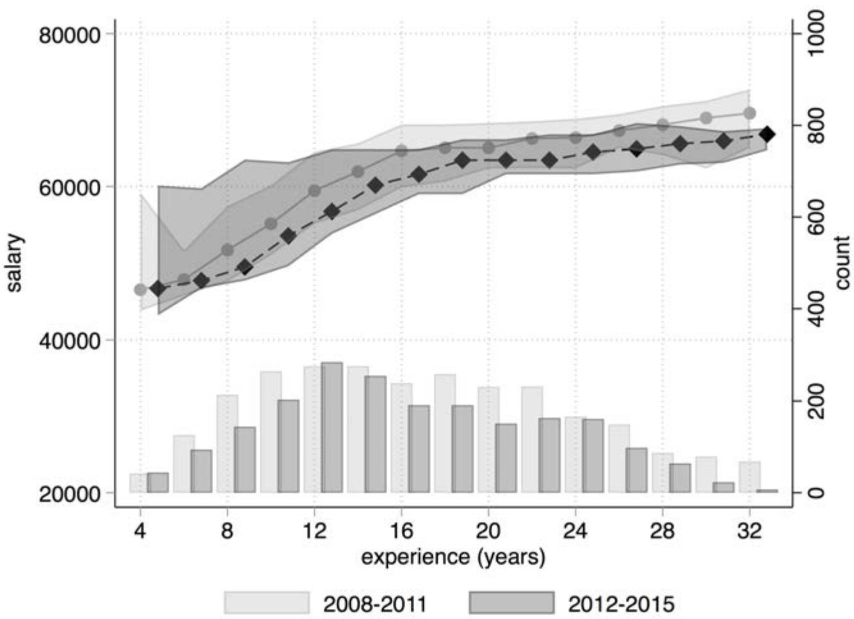
[30] The two districts are comparable in size, enrolling 20,514 and 20,457 students in 2012, respectively.

Figure 1: Empirical Salary Schedule - Median and 90-10 Percentile Ranges Range of Salaries, 2008–2011 and 2012–2015, School Districts of Racine (top) and Green Bay (bottom)

Panel A: Racine Unified School District



Panel B: Green Bay Area School District



Notes: Median and 90-10 percentile range of salaries, by two-year experience classes, for teachers in the school districts of Racine (panel A) and Green Bay (panel B), for the years 2008–2011 (grey line and lighter area) and 2012–2015 (black line and darker area). The bars correspond to counts of teachers in each seniority bin. The sample is restricted to teachers with three to 35 years of experience and with a master's degree.

of $9,426).

After Act 10, the difference in salary dispersion between the two districts becomes striking (bottom panel). The 90-10 percentile range for teachers with five or six years of experience was equal to $7,923 in Racine and $13,127 in Green Bay. For teachers with 11 or 12 years of experience it was $10,739 in Racine and $11,088 in Green Bay. No differences in salary dispersion can be observed for teachers with higher levels of experience.

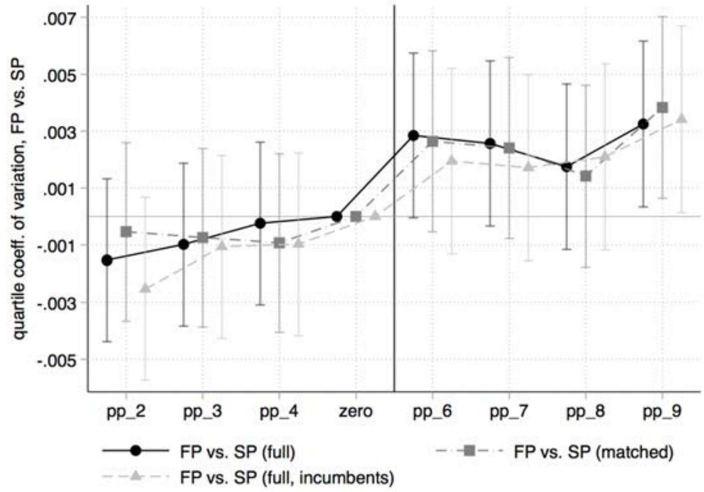
To more systematically quantify the increase in dispersion across all FP and SP districts, Figure 2 shows the trend in the FP-SP difference in the quartile coefficient of dispersion (QCD), defined as the difference between the 75th and 25th percentile of salaries divided by its sum, and calculated within each district and for teachers with similar experience and education. This measure is meant to capture the dispersion in salaries in a way that is insensitive to changes in the salary levels. The FP-SP difference in QCD is flat and indistinguishable from zero between 2007 and 2011; it increases to 0.3 percent in 2012, remaining at this level until 2015.[31] This finding is robust to using other measures of dispersion, such as the coefficient of variation. The increase in pay dispersion indicates that the departure from a salary schedule regime in FP districts led to teachers with the same experience and education earning different salaries. This suggests that FP districts used their newly-acquired flexibility to compensate teachers for attributes not rewarded by a standard lock-step schedule.

To understand the extent to which the observed increase in pay dispersion is driven by changes in salaries of incumbent teachers (i.e. teachers who were already in the district in the previous year) as opposed to changes in salaries offered to new hires, I re-estimate the FP-SP difference in QCD solely on the subsample of incumbents. While imprecisely estimated, the post-Act 10 increase in this difference is smaller than the one in the full sample but greater than zero (Figure 2, dashed line). This indicates that the post-Act 10 increase in pay dispersion is driven by both changes in salaries for new hires and pay renegotiation for incumbent teachers.

Salaries and Teacher Quality. What drove the post-Act 10 increase in salary dispersion in FP districts? To answer this question, the ideal test would estimate the correlation between pay and those teacher attributes, not rewarded under seniority pay, which districts may want to compensate under a FP scheme, including (but not limited to) preparedness, progress, leadership, and professional development. Most of these attributes, however, are only observable to principals and other school administrators and are difficult to measure. I hence settle on a more

[31] Trends in the raw QCD for FP and SP districts are shown in Appendix Figure A7.

Figure 2: Quartile Coefficient of Dispersion in Salaries: FP-SP Difference, 2008–2015



Notes: Point estimates and 90 percent confidence intervals of the FP-SP difference in the median QCD (relative to the difference in 2011). The differences are estimated as δ_s in the equation $k_{ijt} = \alpha FP_j + \sum_{s\neq 2011} \delta_s FP_j{*}\tau_s{+}\tau_t{+}\varepsilon_{ijt}$, where k_{ijt} is the QCD of group i of teachers in district j and year t, FP_j equals one for FP districts, and τ_t are year fixed effects. Each group contains teachers with the same experience and education in each district and year. QCDs are calculated as the ratio between the difference and the sum of the 75th and 25th percentiles of salaries, computed separately for each group of teachers. The sample is restricted to tenured teachers (with more than three years of experience) working in FP, SP, and matched SP districts. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. The sample of incumbents contains teachers already teaching in the district in the previous year.

modest task and study the correlation between salaries and teacher VA, conditional on experience and education. While districts do not observe nor explicitly use VA to evaluate teachers, this measure could be correlated with other attributes that districts can observe and value.

I estimate this correlation using the following model:

$$\log(w_{ijt}) = \delta_0 VA_{it} + \delta VA_{it} * post_t + \beta X^w_{it} + \theta_j + \tau_t + \varepsilon_{ijt} \tag{5}$$

where w_{ijt} is the salary earned by teacher i in district j and year t, VA_{it} is teacher VA (calculated as the average over the years 2007–2011 and 2012–2015, and standardized to have mean 0 and variance 1), and the variable $post_t$ equals one for the years 2012–2015. The vector X^w_{it} controls for experience and education in a flexible way, and it includes a non-parametric function of years of experience, interacted with indicators for the highest education degree and with a dummy for years after 2011 (to allow the gradient between salaries, experience, and education to vary after Act 10). The vector of district fixed effects θ_j controls for district-specific differences in salaries and the vector of year fixed effects τ_t controls for time trends in a non-parametric way. I estimate the equation using OLS, and I calculate bootstrapped standard errors (to account for the fact that

VA is an estimated variable) clustered at the district level. In this specification, the coefficient δ_0 captures the conditional correlation between salaries and VA before 2011, while the parameter δ captures the change in this correlation after 2011.

In the sample of FP districts, the conditional correlation between salary and VA is indistinguishable from zero until 2011 (with an estimate of δ_0 equal to -0.0008, Table 2, column 1, p-value equal to 0.71), and it becomes positive and significant after 2011 (with an estimate of δ equal to 0.005, Table 2, column 1, significant at 5 percent). This implies that a one-standard deviation higher VA is associated with a 0.5 percent higher salary. In the full and matched samples of SP districts, estimates of δ are instead much smaller and indistinguishable from zero (equal to 0.1 and 0.2 percent respectively, Table 2, columns 2 and 3). Estimates of δ for FP and SP districts are also significantly different from each other (Appendix Table A8).[32] Consistent with Figure 1, estimates of δ are larger for teachers with less than 10 years of experience in FP districts (1.4 percent, Table 2, column 4, significant at 1 percent).

Table 2: Teacher Salaries and Value-Added. OLS, Dependent Variable is log(Salary)

	All teachers			Teachers with ≤10 years of experience		
	(1) FP	(2) SP (full)	(3) SP (matched)	(4) FP	(5) SP (full)	(6) SP (matched)
VA	-0.0008	-0.0011	-0.0017	-0.0081*	-0.0014	-0.0020
	(0.0022)	(0.0011)	(0.0026)	(0.0045)	(0.0020)	(0.0036)
VA * post	0.0048**	0.0014	0.0025	0.0137***	0.0015	0.0028
	(0.0024)	(0.0015)	(0.0024)	(0.0050)	(0.0029)	(0.0032)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Edu*exp*post	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39445	48507	19986	12129	16216	6188
# districts	98	119	54	98	118	54

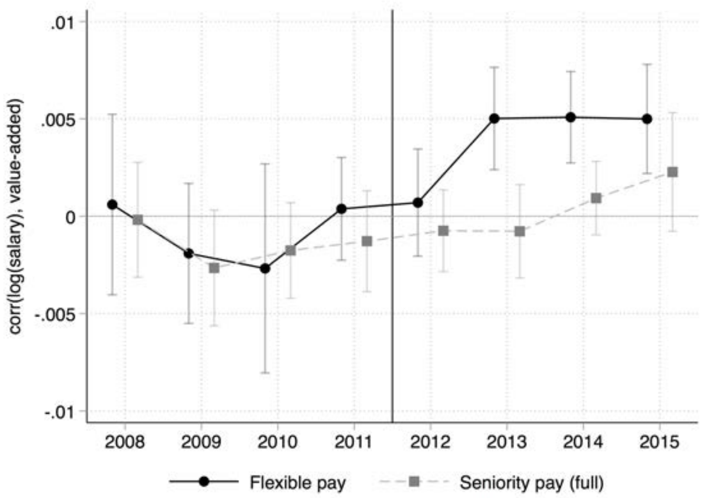
Notes: The dependent variable is the natural logarithm of salaries. The variable *VA* is teacher VA, normalized to have mean 0 and standard deviation 1. The variable *post* equals 1 for years following 2011. All the regressions include year and district fixed effects, as well as indicators for years of experience interacted with indicators for highest education degree interacted with *post*. VA is calculated as the average of a time-varying measure over the years 2007–2011 and 2012–2015. The sample is restricted to tenured teachers (with more than three years of experience) working in FP, SP, and matched SP districts, and it covers years 2007 to 2015. In columns 4-6, the sample is further restricted to include teachers with less than 10 years of experience. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. Bootstrapped standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

To assess how the correlation between salaries and teacher quality changed over time and to

[32] Tests for the differences in the estimates of δ between FP and SP districts are reported in Table A8.

check for the existence of pre-trends, I estimate the parameter δ separately for each year between 2008 and 2015 and for FP and SP districts. These estimates, shown in Figure 3, are indistinguishable from zero and very similar across both groups of districts in the years 2008–2011. In line with Table 2, estimates become positive and statistically significant in FP districts after 2011, reaching 0.5 percent in 2013 (Figure 3, solid line). They instead remain indistinguishable from zero in SP districts until 2015 (Figure 3, dashed line).

Figure 3: Correlation, Salaries and Value-Added: FP and SP Districts, 2008–2015



Notes: OLS estimates and 90 percent confidence intervals of the coefficients δ_s in the regression $log(w_{ijt}) = \sum_{s=2008}^{2015} \delta_s \tau_s * VA_{it} + \beta X_{it}^{w} + \theta_j + \tau_t + \varepsilon_{ijt}$. The variable $log(w_{ijt})$ is the natural logarithm of salary for teacher i working in district j in year t. The variable VA_{it} is teacher VA. The vector X_{it}^{w} includes a non-parametric function of years of experience, interacted with indicators for the highest education degree and with a dummy for years after 2011. The vector θ_j contains district fixed effects and the vector τ_t contains year fixed effects. The coefficients δ_s are estimated separately for FP and SP districts. VA is calculated as the average of a time-varying measure over the years 2007–2011 and 2012–2015. The sample is restricted to tenured teachers (with more than three years of experience) working in FP and SP districts. Bootstrapped standard errors are clustered at the district level.

Appendix Figure A9 shows the semi-parametric relationship between salaries and VA, captured by the pre- vs. post-Act 10 difference in conditional salaries by deciles of VA. In FP districts, teachers in the top decile earn approximately 2.5 percent more than teachers in the bottom decile (significant at 1 percent), whereas in SP districts they earn only 1.1 percent more (significant at 10 percent). The correlation between conditional salaries and VA is highest for teachers with four or five years of experience in FP districts (2.3 percent, significant at 1 percent), whereas it is indistinguishable from zero for teachers with more than 10 years of experience in FP districts and for all teachers in SP districts (Appendix Figure A10).

Although positive, estimates of δ are small in magnitude, and at a first glance it might seem hard to believe that such small salary premia produce any change in teacher behavior at all.

It should be emphasized, however, that districts do *not* use VA when making decisions over teacher pay. Interviews with FP districts' administrators reveal that their post-Act 10 schemes are designed to reward teachers for a number of attributes, including (but not limited to) their preparation, leadership, learning, and professional development.[33] If these characteristics have a positive but small correlation with VA, this could result in low estimates of δ due to attenuation bias.[34] In light of this, the estimates of δ should be interpreted as suggestive evidence that districts use their post-Act 10 pay flexibility to reward teacher characteristics that are, at least to some extent, positively correlated with VA, rather than as true estimates of the actual salary premia enjoyed by teachers under the new payment scheme.

5 Movements, Exits, and Changes in Workforce Composition

The cross-district differences in salaries that arose in the aftermath of Act 10 changed teachers' job prospects. A simple Roy model (Roy, 1951, outlined in Appendix C) predicts that a switch to a FP regime in some (but not all) districts would induce high-quality teachers to move from SP to FP districts and low-quality teachers to either move in the opposite direction or to leave the market. The intuition behind this result is that a SP scheme under-compensates high-quality teachers relative to a FP scheme, whereas a FP scheme penalizes low-quality teachers. I test these predictions by studying teachers' movements across districts and exits from this labor market.

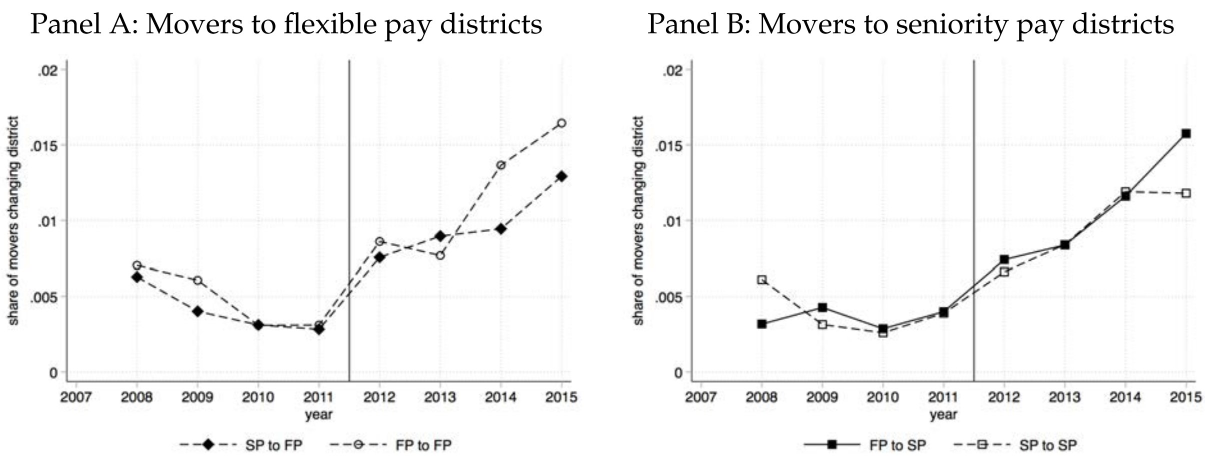
5.1 Movements Across Districts

Teacher movements increased rapidly in the aftermath of Act 10, across districts of different type (i.e. from FP to SP and *vice versa*), as well as within districts of the same type (Figure 4). Moving rates (defined as the ratio between the number of teachers moving to a certain type of district and the total number of teachers in the type of district of origin) increased from 1.8 to 4.3 percent from SP to FP, and from 2.0 to 3.9 percent from FP to SP. Similarly, movements between SP districts increased from 2.0 to 3.9 percent and movements between FP districts increased

[33] From interviews with superintendents of a subset of 12 FP and SP districts, conducted in December 2017.

[34] Papay and Kraft (2015) show that professional development is associated with improvements in teacher quality. Dobbie (2011) demonstrates that teacher leadership is a good predictor of future student test scores among Teach for America corps members. Jackson et al. (2014) provide a review of the literature on teacher attributed associated with VA. In Appendix Table A9, I also test whether FP districts pay higher salaries to teachers in subjects that usually experience teacher shortages, such as math and science, conditional on experience, education, and VA. These estimates do not show significant premia for teachers in these subjects.

Figure 4: Moving Rates, by District of Origin and Destination



Notes: Shares of teachers changing district, by type of district of origin and destination. Shares are defined with respect to the district of origin.

from 2.4 to 4.7 percent.[35]

Although the post-Act 10 increase in moving rates is fairly similar across types of districts, the characteristics of movers could be different. For example, the introduction of a FP regime (which rewards quality) could have induced higher VA teachers to move from SP to FP districts and lower VA teachers to move in the opposite direction. To test this hypothesis I study whether the probability of moving to a district of a given type (conditional on the district of origin) differs between high- and low-quality teachers. I use the following models:

$$MoveFP_{ikjt} = \beta_0^{FP} highVA_i + \beta^{FP} highVA_i * post_t + \gamma_1^{FP} X_{it} + \gamma_2^{FP} Z_{jt} + \theta_k + \tau_t + \varepsilon_{ikjt} \quad (6)$$

$$MoveSP_{ikjt} = \beta_0^{SP} highVA_i + \beta^{SP} highVA_i * post_t + \gamma_1^{SP} X_{it} + \gamma_2^{SP} Z_{jt} + \theta_k + \tau_t + \varepsilon_{ikjt} \quad (7)$$

where $MoveFP_{ikjt}$ equals one if teacher i moves from a district k to a district j of type FP in year t, and $MoveSP_{ikjt}$ equals one if teacher i moves from a district k to a district j of type SP in year t.[36] The variable $highVA_i$ equals one for teachers with *ex ante* value added above the median. The vector X_{it} includes indicators for the number of years of experience interacted with indicators for the highest education degree. The vector Z_{jt} controls for characteristics of the district of destination, such as an array of pre-Act 10 teacher, student, and district attributes

[35]Such a large increase in movements within districts of the same type might appear surprising. Consider, however, that the overall increase in movements and exits after Act 10 led to a surge in vacancies. This could have induced some teachers to move between districts of the same type for reasons not strictly related to salaries. The empirical evidence on movements across districts of the same type does not show clear patterns of sorting with respect to quality.

[36]If district j is SP, $MoveFP_{ikjt} = 0$; similarly, if district j is FP, $MoveSP_{ikjt} = 0$.

interacted with an indicator for year t, as well as indicators for whether the district had a union recertification election in year t and whether the election was successful.[37] The vector θ_k includes fixed effects for the district of origin. I estimate the model via OLS separately for teachers working in FP and SP districts in the previous year, and I cluster standard errors at the level of the district of origin. The coefficients β^{FP} and β^{SP} estimate the post-Act 10 change in the probability of moving to a FP or SP district, respectively, for high-quality teachers relative to low-quality ones.

Table 3: Teacher Sorting. OLS, Dependent Variable Equals 1 for Teachers Moving to or Exiting from a District

	Moving to FP		Moving to SP		Exiting from	
	(1) from SP	(2) from FP	(3) from FP	(4) from SP	(5) FP	(6) SP
high VA	0.0001	-0.0017	0.0009	0.0005	0.0033	0.0033
	(0.0010)	(0.0011)	(0.0013)	(0.0010)	(0.0048)	(0.0044)
high VA * post	0.0034*	0.0013	-0.0050***	-0.0030*	-0.0179*	-0.0055
	(0.0018)	(0.0016)	(0.0019)	(0.0016)	(0.0093)	(0.0062)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes	Yes
CB controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33734	30172	30172	33734	17881	22537
# districts	121	100	100	121	98	119
Y-mean	0.003	0.004	0.003	0.003	0.041	0.041

Notes: The dependent variable equals 1 for teachers who move to a FP district (columns 1-2), move to a SP district (column 3-4), exit from a FP district (column 5), and exit from a SP district (column 6). In columns 1 and 4 the sample is restricted to teachers already working in a SP district; in columns 2 and 3 it is restricted to those already working in a FP district. The variable *high VA* equals one for teachers with *ex ante* VA above the median. The variable *post* equals one for years after Act 10. All the regressions include year and district fixed effects. *District controls* include interactions between 2009–2011 averages of district characteristics interacted with year fixed effects. *CB controls* include an indicator for whether the district had a union recertification election in year *t* and whether the election was successful. *Teacher controls* include indicators for the number of years of experience and for the highest education degree. Columns 5 and 6 control non-parametrically for age. *Ex ante* VA is calculated as the average of a time-varying measure over the years 2007–2011. The sample is restricted to tenured teachers (with more than three years of experience) working in FP and SP districts, and covers years 2008 to 2015 (columns 1-4) and years 2008 to 2012 (columns 5-6). Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

[37] Pre-Act 10 characteristics include the 2009–2011 averages of enrollment, share of low-SES students, salary for all teachers and for teachers with less than five years of experience, property values per pupil, indicators for urban and suburban districts, total expenditure and state aid per pupil, share of teachers with a master's degree, share of teachers with less than three years of experience, and share of teachers with more than 20 years of experience.

OLS estimates indicate sorting of high-quality teachers from SP to FP districts and sorting of low-quality teachers from FP to SP districts. Teachers in SP districts with VA above the median were 0.34 percentage points more likely to move to a FP district after Act 10 compared with teachers with VA below the median (estimate of *high VA * post*, Table 3, column 1, significant at 10 percent). Compared with an average moving rate from SP to FP of 0.27 percent in 2008-2011, this corresponds to a 113 percent increase. In contrast, high-quality teachers in FP districts are -0.50 percentage points *less* likely to move to a SP district after Act 10 compared with low-VA teachers, or -167 percent (Table 3, column 3, significant at 1 percent). Higher VA teachers are only 0.13 percentage points more likely to move across FP districts (Table 3, column 1, p-value equal to 0.42) and-0.30 percentage points less likely to move across SP districts (Table 3, column 4, significant at 10 percent). Estimates are robust to excluding Milwaukee and Madison (Appendix Table A10) and to using the matched sample (Appendix Table A11).

To investigate the presence of pre-trends, in Figure 5 I allow β^{FP} and β^{SP} to vary over time between 2008 and 2015, normalizing them to zero in 2011. Time-specific estimates of β^{FP} on the subsample of teachers working in SP districts are very close to zero between 2008 and 2010, confirming the absence of pre-trends; they become positive and significant after 2011, reaching 0.4 percentage points in 2012 (significant at 5 percent). Estimates of β^{SP} on the subsample of teachers working in FP districts are also indistinguishable from zero between 2008 and 2010, and they become negative after 2011, dropping to -0.9 percentage points in 2014 (significant at 5 percent).

As an additional test of sorting, in columns 1 and 2 of Table 4 I estimate the post-Act 10 difference in VA between movers to FP districts and movers to SP districts. I use the following empirical model on the subsample of teachers who move across districts in each year:

$$\begin{aligned} VA_i^{m(kjt)} &= \beta_0 FP_j + \beta FP_j * post_t + \gamma_1 X_{it} + \gamma_2 Z_{jt} \\ &\quad + \eta_1 FP_k + \eta_2 FP_k * post_t + \eta_3 FP_k * FP_j + \eta_4 FP_k * FP_j * post_t + \tau_t + \varepsilon_{ijkt} \end{aligned} \tag{8}$$

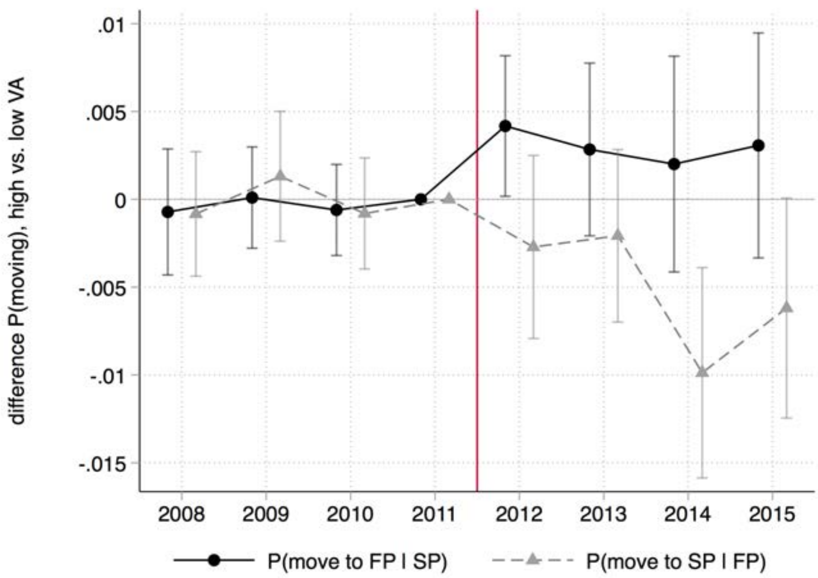
where $VA_i^{m(kjt)}$ is average *ex ante* VA of teacher i, who moves from district k to district j in year t. The coefficient β captures the post-Act 10 change in VA of movers to FP districts after Act 10, conditional on the district of origin and relative to movers to SP districts.

OLS estimates of β on the full sample of FP and SP districts indicate that, after Act 10 and conditional on the district of origin (captured by FP_k and its interactions with FP_j and $post_t$), movers to FP districts have a 0.36 standard deviations higher VA compared with movers to SP districts (Table 4, column 1, significant at 10 percent). The estimate is robust to controlling for various school district budget items, including per-teacher expenditure on salaries, retirement, and health and other insurance, as well as total per-student expenditure and state aid (0.37 standard deviations, Table 4, column 2, significant at 10 percent). Estimates on the matched sample are similar in magnitude, although less precise (0.27 and 0.31 standard deviations, Appendix Table A11, columns 1-2, p-values equal to 0.40 and 0.40, respectively).

The findings presented so far are in line with the theoretical predictions of a simple Roy model: After Act 10, high-quality teachers sort into FP districts and low-quality teachers sort into SP districts.

Salaries of Movers. The Roy model implies that the observed cross-district sorting patterns are driven by higher salaries in the district of destination. To provide evidence in line with this prediction, I conduct an event study of post-Act 10 changes in salaries (conditional on experience

Figure 5: Difference in Moving Rates, High VA vs. Low VA Teachers, by District of Origin and Destination



Notes: Estimates and 90% confidence intervals of β in the regression $MovetoW_{ikjt} = \sum_{s=2008}^{2015} \beta_s highVA_i * \tau_s + \gamma_1 X_{it} + \gamma_2 Z_{jt} + \theta_k + \tau_t + \varepsilon_{ikjt}$, where $MovetoW_{ikjt}$ equals one if teacher i moves from district k to district j of type W in year t, and $W = \{FP, SP\}$. $highVA_i$ equals one for teachers with ex ante VA above the median, τ_t are year fixed effects, X_{it} is a vector of teacher controls (including indicators for the number of years of experience and for the highest education degree), Z_{jt} are controls for the district of destination (including interactions between the 2009–2011 averages of district characteristics interacted with year fixed effects and indicator for whether the district had a union recertification election in year t and whether the election was successful), and θ_k are district-of-origin fixed effects. The parameter β_{2011} is normalized to zero. The solid line includes teachers moving out of SP, and the dashed line includes teachers moving out of FP. Standard errors are clustered at the district level.

Table 4: Changes in the Composition of Movers and Exiters. OLS, Dependent Variable is Ex Ante Teacher Value-Added

	(1) Movers	(2) Movers	(3) Exiters	(4) Exiters
IS	-0.0336	-0.0566	0.0023	0.0024
	(0.1618)	(0.1572)	(0.0776)	(0.0880)
FP * post	0.3630*	0.3721*	-0.1837**	-0.1879**
	(0.2095)	(0.1973)	(0.0846)	(0.0892)
Year FE	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes
CB controls	Yes	Yes	Yes	Yes
Teacher controls	Yes	Yes	Yes	Yes
Budget controls	Yes	No	No	Yes
Observations	628	630	2516	2075
# districts	147	149	213	206

Notes: The dependent variable is *ex ante* teacher VA. Columns 1-2 and 5-6 are estimated on the subsample of movers to a district; columns 3-4 are estimated on the subsample of leavers from a district (defined as teachers who leave Wisconsin's teaching workforce). The variable *FP* equals one for FP districts. The variable *post* equals one for years following 2011. All the regressions include year fixed effects. *District controls* include interactions between the 2009–2011 averages of district characteristics interacted with year fixed effects. *CB controls* include an indicator for whether the district had a union recertification election in year t and whether the election was successful. *Teacher controls* include indicators for the number of years of experience and for the highest education degree. *Budget controls* are district-year-level controls for the level of state aid as a share of total revenues, as well as per-teacher expenditure on salaries, retirement, health, life, and other insurance, and other employee benefits. Columns 1-2 include indicators for the type district of origin (FP or SP), interacted with *FP* and with *post*. Columns 3-4 control non-parametrically for age. *Ex ante* VA is calculated as the average of a time-varying measure over the years 2007–2011. The sample is restricted to tenured teachers (with more than three years of experience) working in FP and SP districts and covers years 2008 to 2015. Standard errors in parentheses are clustered at the district level. *** p < .01, ** p < .05, * p < .1.

and education) for high and low VA movers across different types of districts. I estimate the following model on the subsample of teachers who move at least once between 2007 and 2015:

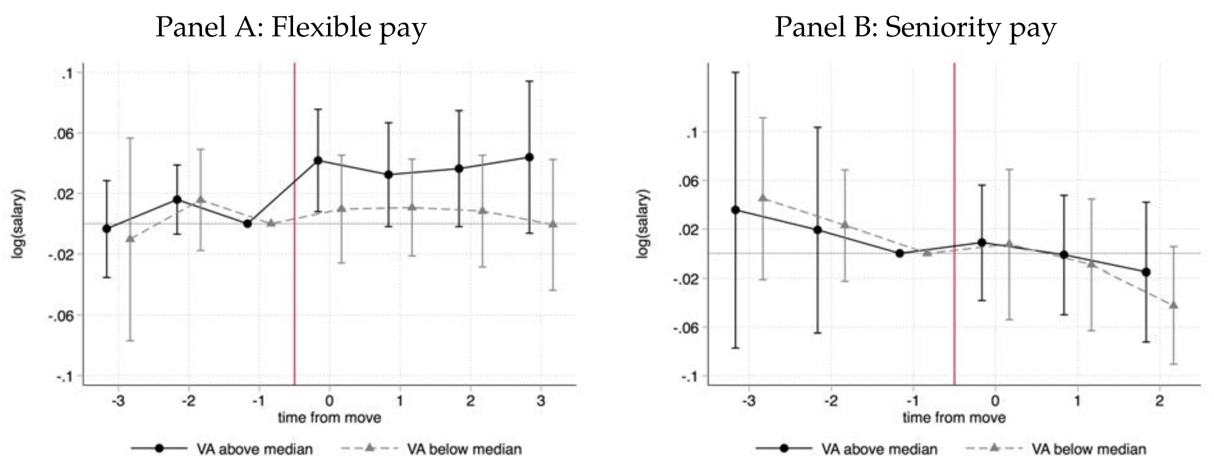
$$\log(w_{ijt}) = \sum_{k=-3}^{3} \beta_k^0 \mathbb{1}(t - Y_i^{m(j)} = k) + \sum_{k=-3}^{3} \beta_k \mathbb{1}(t - Y_i^{m(j)} = k) * \mathbb{1}(Y_i^{m(j)} > 2011) + \gamma X_{it}^w + \theta_j + \tau_t + \varepsilon_{ijt} \tag{9}$$

where the variable $Y_i^{m(j)}$ denotes the year in which teacher i moves to district j.[38] Normalizing β_{-1}^0 and β_{-1} to be zero, the parameter vector β^0 estimates the salary premium (or loss) in the three years before and after a teacher moves (relative to the year preceding a move), whereas the parameter β captures the change in this premium after Act 10. I estimate this model on the

[38]For teachers who move more than once between 2007 and 2015, I consider only the earliest move. The results are robust to using the latest move.

subsample of teachers who move at least once between 2007 and 2015, separately for teachers with VA above and below the median (equal to -0.013 for this subsample) and for teachers in FP and SP districts.

Figure 6: Salaries of Movers Around a Move



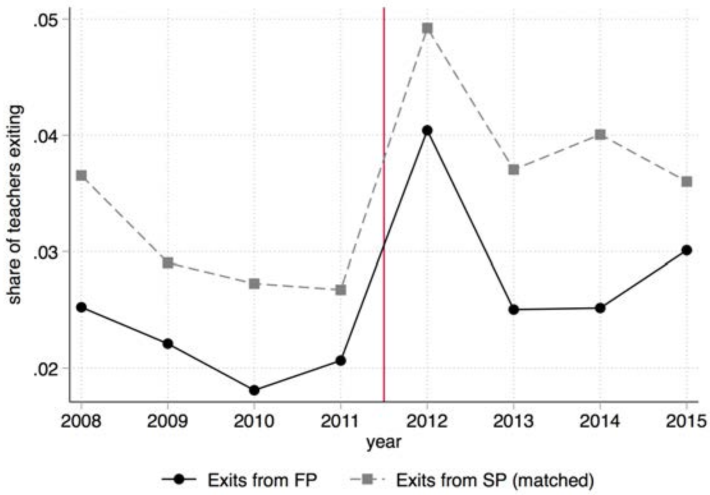
Notes: OLS estimates and 90 percent confidence intervals of the coefficients β_s in the regression $\log(w_{ijt}) = \sum_{k=-3}^{3} \beta_k^0 \mathbb{1}(t - Y_i^{m(j)} = k) + \sum_{k=-3}^{3} \beta_k \mathbb{1}(t - Y_i^{m(j)} = k) * \mathbb{1}(Y_i^{m(j)} > 2011) + \gamma X_{it}^w + \theta_j + \tau_t + \varepsilon_{ijt}$. The variable $log(w_{ijt})$ is the natural logarithm of salary for teacher i working in district j in year t. The variable $Y_i^{m(j)}$ denotes the year in which teacher i moves to district j, $\mathbb{1}(.)$ is an indicator function, and the vector X_{it}^w includes a non-parametric function of years of experience, interacted with indicators for the highest education degree and with a dummy for years after 2011. θ_j are district fixed effects and τ_t are year fixed effects. The coefficient β_{-1} is normalized to zero. The parameters are estimated separately for teachers in FP and in SP districts and with *ex ante* VA above and below the median. The sample is restricted to tenured teachers (with more than three years of experience) working in FP and SP districts. Standard errors are clustered at the district level.

OLS estimates of the vector β_k, shown in Figure 6, indicate that high-quality movers to FP districts experienced a significant 4.5 percent conditional salary increase in the year after a move, compared with similar teachers who moved before Act 10 (Figure 6, Panel A, solid line, significant at 5 percent). Notably, no trends in salary differences can be observed in the years leading to a move. This premium persists up to three years following a move. Low-quality teachers, on the other hand, did not experience any significant change in salaries after moving to a FP district after Act 10 (Figure 6, Panel A, dashed line). Similarly, high-quality and low-quality movers to SP districts experienced no differential change in post-move salaries after Act 10 (Figure 6, Panel B). Estimates on the matched sample are very similar (Appendix Figure A15). These findings provide suggestive evidence that, in the aftermath of Act 10, high-quality teachers were attracted to FP districts by the prospect of higher salaries.

5.2 Exit from Public Schools

The increase in movements of teachers across districts after Act 10 was accompanied by a surge in exits (Figure 7).[39] In 2011, 2.1 percent and 2.7 percent of teachers left FP and SP districts in each year, respectively. In 2012, these rates increased to 4.0 and 4.9 percent.[40]

Figure 7: Exit Rates, by District of Origin



Notes: Share of teachers leaving Wisconsin public schools, by type of district of origin.

Although trends in exit rates appear similar across the two groups of districts, the characteristics of the teachers who left could be different. For example, the introduction of quality pay in FP districts could have induced low-quality teachers to exit at a higher rate compared with high-quality teachers. To test this hypothesis I estimate the following equation:

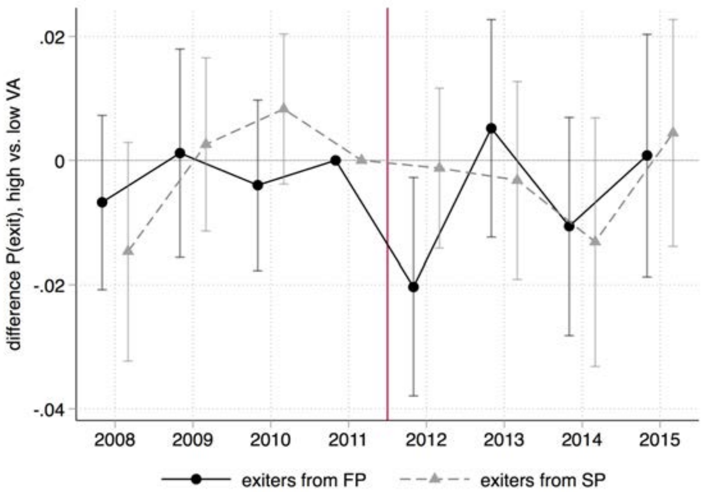
$$e_{ijt} = \beta_0 highVA_i + \beta highVA_i * post_t + \gamma_1 X_{it}^e + \gamma_2 Z_{jt} + \theta_j + \tau_t + \varepsilon_{ijt} \tag{10}$$

where e_{ijt} equals one if teacher i leaves the market from district j in year t. The vector X_{it}^e includes indicators for the highest education degree, as well as a non-parametric control for age and experience interacted with an indicator for years after 2011, to account for a differential propensity to retire after Act 10 (Roth, 2017; Biasi, 2017). Since the bulk of retirement occurred in 2012 I estimate this equation on the years 2008–2012, separately for teachers in FP and SP districts.

[39]Exit rates are defined as the share of individuals who disappear from the records of employees in Wisconsin public schools. Reasons for exiting include retirement, dropping out of the labor force, or a move to a to private school or to another industry/occupation. The staff data does not allow me me to observe a teacher after she leaves, and I am thus unable to distinguish among these reasons.

[40]The spike in exits is partly due to a surge in retirement (Roth, 2017; Biasi, 2017): exit rates of teachers above age 55 increased from 6.2 to 10.2 percent in FP districts and from 7.8 to 13.0 percent in SP districts. However, they also increased for teachers below age 55, from 0.8 to 1.5 percent in FP districts and from 1.1 to 1.8 percent in SP districts.

Figure 8: Difference in Exit Rates, High Value-Added vs. Low Value-Added Teachers, by District of Origin



Notes: Estimates and 90% confidence intervals of β in the regression $e_{ijt} = \sum_{s=2008}^{2015} \beta_s highVA_i * \tau_s + \gamma_1 X_{it}^e + \gamma_2 Z_{jt} + \theta_j + \tau_t + \varepsilon_{ijt}$, where e_{ijt} equals one if teacher i leaves the market from district j in year t, $highVA_i$ equals one for teachers with ex ante VA above the median, τ_t are year fixed effects, X_{it}^e is a vector of teacher controls (including indicators for years of experience, age, and highest education degree), Z_{jt} are district controls, including the 2009–2011 averages of district characteristics interacted with year fixed effects, and indicator for whether the district had a union recertification election in year t and whether the election was successful, and θ_j are district fixed effects. The parameter β_{2011} is normalized to equal zero. Standard errors are clustered at the district level.

OLS estimates of β (shown in Table 3) indicate that, after Act 10, teachers with VA *below* the median were 1.8 percentage points more likely to exit from a FP district compared with teachers with VA above the median (with an estimate of *high VA * post* equal to -0.018, Table 3, column 5, significant at 10 percent). Compared with an average exit rate of 4.1 percent percent for FP districts in 2007-2011, this corresponds to a 44 percent increase in this probability. By comparison, this estimate is indistinguishable from zero in the full sample of SP districts (-0.0055, Table 3, column 5, p-value equal to 0.38) and in matched SP districts (-0.0093, Table A11, column 6, p-value equal to 0.33).[41] Year-specific estimates of β for the years 2008–2011, shown in Figure 8, are very close to zero for both FP and SP districts, confirming the absence of pre-trends. They become negative and significant in 2012 in FP districts (with an estimate of -2.0 percentage points, significant at 10 percent).

To quantify the overall change in VA for teachers who exit from SP and FP districts, I estimate the following model on the subsample of leavers:

$$VA_i^{e(jt)} = \beta_0 FP_j + \beta FP_j * post_t + \gamma_1 X_{it}^e + \gamma_2 Z_{jt} + \varepsilon_{ijt} \tag{11}$$

[41]Estimates are robust to excluding Madison and Milwaukee (Appendix Table A10).

The coefficient β captures the post-Act 10 difference in VA of leavers from FP districts relative to SP, conditional on the district of origin. Estimates on the full sample of FP and SP districts indicate that, after Act 10, VA of leavers from FP districts was 0.18 standard deviations smaller than VA of leavers from SP districts (Table 4, column 4, significant at 5 percent). Estimates on the subsample of matched FP and SP districts are even larger in magnitude, with -0.26 standard deviations (Appendix Table A12, column 8, significant at 1 percent). Taken together, these results indicate a disproportionate exit of low-quality teachers from FP districts compared with SP districts after Act 10.

Salaries of Exiters. Next, I test whether this exit flow is related to a decline in salaries. I estimate the following model:

$$\log(w_{ijt}) = \beta^0 e_{ijt+1} + \beta e_{ijt+1} * post_t + \beta X^w_{it} + \theta_j + \tau_t + \varepsilon_{ijt} \tag{12}$$

I estimate this model separately for teachers in FP and SP districts and with VA above and below the median. Estimates of β capture the post-Act 10 difference in salaries of leavers in the year immediately preceding their exit.

OLS estimates of β, shown in Table 5, indicate that teachers with VA above the median who left FP districts after Act 10 experienced a small and insignificant change in salary right before leaving, compared with similar teachers who exited before Act 10 (0.0019, Table 5, column 1, p-value equal to 0.95). Teachers with VA below the median, on the other hand, experienced a large 2.7 percent decline in salary (Table 5, column 2, significant at 1 percent). In SP districts, high VA leavers experienced a 1.3 percent salary decline after Act 10 (Table 5, column 3, significant at 1 percent), whereas low VA teachers experienced no significant change (Table 5, column 4). These estimates are robust to using the matched sample (columns 5-6). These findings are consistent with the Roy model, which predicts that the disproportionate exits of low-quality teachers from FP districts and of high-quality teachers from SP districts are driven by lower salaries.

5.3 Composition of the Teaching Workforce

Movements of teachers across districts and exits from the market directly affect the composition of the teaching workforce. To quantify this change, I compare *ex ante* teacher VA in FP and SP districts before and after the passage of Act 10. I estimate:

$$VA_i = \beta_0 FP_j + \beta FP_j * post_t + \gamma_1 X_{it} + \gamma_2 Z_{jt} + \tau_t + \varepsilon_{ijt} \tag{13}$$

Table 5: Salaries and Exit. OLS, Dependent Variable is log(Salary)

	Flexible pay		Seniority pay (full)		Seniority pay (matched)	
	(1)	(2)	(3)	(4)	(5)	(6)
	VA$\geq$ med	VA$<$ med	VA$\geq$ med	VA$<$ med	VA$\geq$ med	VA$<$ med
exit	-0.0105**	-0.0038	0.0063**	-0.0068	0.0099*	0.0019
	(0.0041)	(0.0054)	(0.0028)	(0.0044)	(0.0053)	(0.0051)
exit * post	0.0019	-0.0272***	-0.0126***	0.0050	-0.0138**	0.0026
	(0.0125)	(0.0101)	(0.0045)	(0.0057)	(0.0067)	(0.0069)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Edu*exp*post	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20133	19350	23857	24698	9533	10478
# districts	98	97	118	119	54	54

Notes: The dependent variable is the natural logarithm of salaries. The variable *exit* equals one for teachers exiting from a district in the following year. The variable *post* equals one for years following 2011. All the regressions include controls for a non-parametric function of years of experience, interacted with indicators for the highest education degree and with *post*, as well as district and year fixed effects. The sample is restricted to tenured teachers (with more than three years of experience) working in FP (columns 1-2), SP (columns 3-4), and matched SP districts (columns 5-6) and covers years 2008 to 2015. Columns 1, 3, and 5 refer to teachers with *ex ante* VA above the median; columns 2, 4, and 6 refer to teachers with VA below the median. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

The parameter β captures the change in VA in FP relative to SP districts after Act 10. Estimates of β, shown in Table 6, indicate that *ex ante* teacher VA increased by 0.048 standard deviations in FP districts compared with SP after Act 10 (Table 6, column 1, significant at 10 percent). This estimate increases to 0.069 standard deviations with the inclusion of controls for teacher experience and education, for the district's budget composition, and for measures of union strength (column 3), and they are slightly smaller when controlling for district fixed effects (column 4).

To gauge the extent to which the difference in teacher VA between FP and SP districts is due to unobservables, I estimate an upper bound for the bias in β driven by unobservables using the methodology of Altonji et al. (2005). This method relies on the condition that the portion of a district's average VA that is related to unobservables and the portion that is related to the observables included as controls in column (3) have the same relationship with the type of post-Act 10 pay structure. This exercise yields an upper bound for this bias equal to 0.005 standard deviations, which implies that at most seven percent of the estimated compositional change can be attributed to time-varying unobservable differences between FP and SP districts.[42]

Estimates on the matched sample of FP and SP districts yield similar results (Table 6, columns 5-8). Time-varying estimates of β (normalizing the estimate for 2011 to zero), shown in Figure 9, are indistinguishable from zero between 2008 and 2010, and show no evidence of pre-trends. They become positive after 2011, with 0.058 in 2012 (significant at 10 percent) and remain high at this level through 2015.[43]

Table 6: Changes in the Composition of the Teaching Workforce. OLS, Dependent Variable is Ex Ante Teacher Value-Added

	Full sample				Matched sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FP	-0.046	-0.047	-0.041		-0.051	-0.053	-0.070	
	(0.088)	(0.088)	(0.086)		(0.095)	(0.095)	(0.088)	
FP * post	0.048*	0.049*	0.069**	0.029**	0.047	0.047	0.080**	0.041**
	(0.027)	(0.027)	(0.033)	(0.014)	(0.030)	(0.030)	(0.040)	(0.016)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CB controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Budget controls	No	No	Yes	Yes	No	No	Yes	Yes
District FE	No	No	No	Yes	No	No	No	Yes
Observations	65845	65764	65464	65464	45901	45825	45525	45525
# districts	217	217	214	214	152	152	149	149

Notes: The dependent variable is *ex ante* teacher VA. The variable *FP* equals 1 for FP districts. The variable *post* equals one for years following 2011. All the regressions include year fixed effects. *District controls* include interactions between the 2009–2011 averages of district characteristics interacted with year fixed effects. *CB controls* include an indicator for whether the district had a union recertification election in year t and whether the election was successful. *Teacher controls* include indicators for the number of years of experience and for the highest education degree. *Budget controls* are district-year-level controls for the level of state aid as a share of total revenues, as well as per-teacher expenditure on salaries, retirement, health, life, and other insurance, and other employee benefits. Ex ante VA is calculated as the average of a time-varying measure over the years 2007–2011. The sample is restricted to tenured teachers (with more than three years of experience) working in FP, SP, and matched SP districts, and it covers years 2008 to 2015. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

[42] This approach relies on estimating the parameter ρ as defined by Altonji et al. (2005), which represents the correlation between the unobservable component of the independent variable of interest (in this case, *FP*post*) and the unobservable component of the dependent variable. This parameter can be estimated using the estimates of a regression of Equation (13) (including all the controls) and a regression of *FP*post* on the controls, as in Altonji et al. (2005). I estimate ρ to be equal to 0.13.

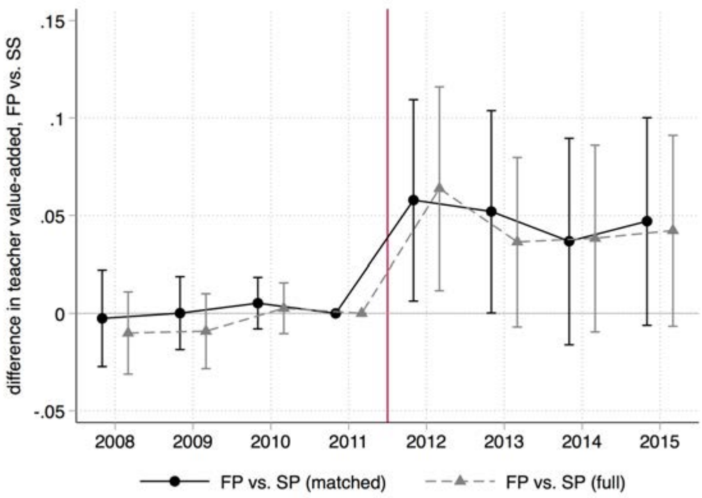
[43] Estimates of the difference in linear pre-trends between FP and SP districts are small and statistically indistinguishable from zero (the results are available upon request). Trends in raw VA are shown in Appendix Figure A11.

It should be noted, at this point, that the above analysis does not include new teachers, for whom VA cannot be calculated. Appendix Figure A16 (top panel) shows that entry rates (defined as the share of new teachers in the population) declined between 2008 and 2011 and increased after Act 10 in both types of districts, possibly due to an increase in the number of vacancies to be filled. If Act 10 induced better or more motivated teachers to enter the market in FP districts, the estimates described so far would represent a lower bound of the true compositional change (Hoxby and Leigh, 2004; Rothstein, 2014).[44] If instead the Act discouraged these teachers from entering, the true compositional change would be smaller. It is also possible that, as of 2015, the supply of new teacher had still not reacted to the policy change. Becoming a teacher requires an education investment of at least two years (the length of a master's degree); the supply of new teachers could therefore respond with a lag.

In an attempt to distinguish between these hypotheses, Appendix Figure A16 (bottom panel) shows trends in the average selectivity of the institution where new teachers obtained their most

[44] Hoxby and Leigh (2004) show that the decline in the entry rates of high-quality teachers in US public schools since 1960 can be attributed to increased compression in wages caused by the rise in unionization. Similarly, Rothstein (2014) demonstrates that higher salaries and lower tenure rates can improve the supply of new teachers.

Figure 9: Changes in Teaching Workforce Composition. Ex Ante Value-Added, FP vs. SP, 2008–2015



Notes: OLS estimates and 90 percent confidence intervals of the coefficients β_s in the regression $VA_i = \alpha FP_j + \sum_{s\neq 2011} \beta_s FP_j * \tau_s + \gamma_1 X_{it} + \gamma_2 Z_{jt} + \tau_t + \varepsilon_{it}$, where VA_i is *ex ante* VA of teacher i employed in district j in time t, FP_j equals one for FP districts, X_{it} includes indicators for the number of years of experience and the highest education degree, Z_{jt} are district controls (including the 2009–2011 averages of district-level characteristics interacted with year fixed effects, and indicator for whether the district had a union recertification election in year t and whether the election was successful), and τ_t are year fixed effects. The sample is restricted to tenured teachers (with more than three years of experience) working in FP, SP, and matched SP districts. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. Standard errors are clustered at the district level.

recent degree, an attribute shown to be correlated with quality (Ballou and Podgursky, 1997; Clotfelter et al., 2010; Hoxby and Leigh, 2004). These trends do not show any change after 2011. This suggests that the characteristics of new teachers did not vary much between 2012 and 2015. This, however, does not eliminate the possibility that composition of the pool of entrants could change over a longer time span. One should interpret and generalize the above findings with this caveat in mind.

6 Effects on Teachers' Effort

The pay scheme adopted by FP districts after Act 10 attracted higher VA teachers from other districts and led lower VA teachers to leave. As movers and leavers represent only a small share of the teachers' population in each year, the resulting compositional change five years after the policy change was rather modest in size. A pay scheme that rewards quality, however, could affect *all* teachers (not only those who move or exit) through changes in the incentives to exert more effort, with potentially larger effects on students.

To test this hypothesis I allow the VA of each teacher to vary between the pre- and post-reform periods. I then estimate the following model:

$$VA_{it} = \beta_0 FP_j + \beta FP_j * post_t + \gamma_1 X_{it} + \gamma_2 Z_{jt} + \tau_t + \varepsilon_{ijt} \tag{14}$$

where VA_{it} is the time-varying VA of teacher i, working in district j in year t. In this equation, the coefficient β captures the *overall* change in teacher quality after Act 10 in FP districts relative to SP, driven by both changes in composition and changes in effort.

OLS estimates of β indicate that the VA of teachers in FP districts increased by 0.11 standard deviations after Act 10 compared with the VA of teachers in SP districts (Table 7, column 1, significant at 10 percent). Assuming that this overall change is simply the sum of a compositional change (estimated in column 3 of Table 6) and a change in effort, approximately 35 percent of the overall increase in VA is due to changes in effort (0.107 - 0.069 divided by 0.107), whereas 65 percent is driven by changes in composition. Time-varying estimates of β in Equation (14), shown in Figure 10 (solid thick line), show no evidence of pre-trends and indicate that this increase happened in 2012 and persisted through 2015.

To isolate changes in effort from changes in composition more directly, I re-estimate Equation (14) using the subsample of incumbent teachers, i.e. those who did not move or exit between

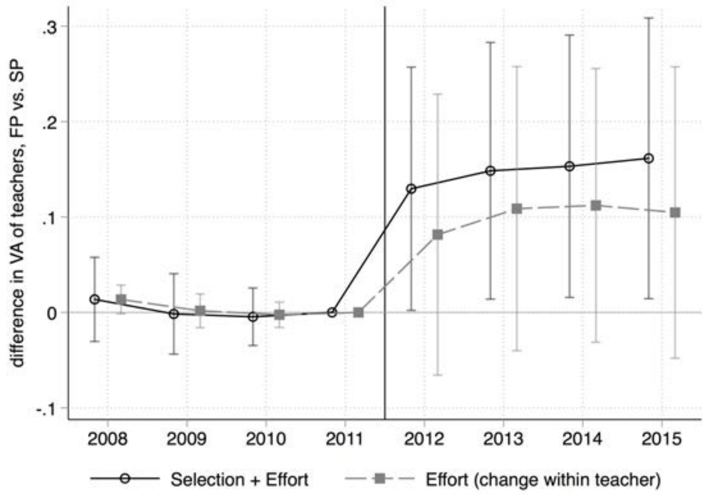
Table 7: Changes in Teacher Effort. OLS and 2SLS, Dependent Variable is Teacher Value-Added

	Full sample			Matched sample		
	(1) Selection + Effort	(2) Effort (Inc.)	(3) Effort (Teacher FE)	(4) Selection + Effort	(5) Effort (Inc.)	(6) Effort (Teacher FE)
FP	-0.0399	-0.0192	-0.0274	-0.0750	-0.1762*	-0.0469
	(0.0846)	(0.0878)	(0.0405)	(0.0872)	(0.0980)	(0.0648)
FP * post	0.1071*	0.0998	0.0523	0.1485*	0.1743**	0.0982
	(0.0636)	(0.0665)	(0.0629)	(0.0784)	(0.0732)	(0.0782)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes	Yes
CB controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher controls	Yes	Yes	Yes	Yes	Yes	Yes
Budget controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE	No	No	Yes	No	No	Yes
Observations	64554	46547	63791	44858	32719	44292
# districts	214	214	214	149	149	149

Notes: The dependent variable is the VA of all teachers (columns 1, 3, 4, and 6) and incumbent teachers (columns 2 and 5). Incumbent teachers are defined as those who do not change district nor exit Wisconsin public schools after Act 10. The variable *FP* equals one for FP districts. In columns 3 and 6, the variable *FP* is instrumented with an indicator for whether a teacher has taught at least once in a FP district between 2007 and 2011. The variable *post* equals one for years following 2011. All the regressions include year fixed effects. *District controls* include interactions between 2009–2011 averages of district characteristics and year dummies. *CB controls* include an indicator for whether the district had a union recertification election in year t and whether the election was successful. *Teacher controls* include indicators for each number of years of experience and for the highest education title (bachelor's degree, master's degree, Ph.D.). VA is calculated as the average of a time-varying measure over the years 2007–2011 and 2012–2015. The sample is restricted to tenured teachers (with more than three years of experience) working in FP, SP, and matched SP districts, and it covers years 2008 to 2015. The matched sample of SP districts is obtained via nearest-neighbor matching on observable characteristics of the school districts. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

2007 and 2015. Although the effects are imprecisely estimated, their magnitude suggests that the VA of incumbent teachers increased by approximately 0.1 standard deviations in FP districts compared with SP districts after Act 10 (Table 7, column 2, p-value equal to 0.13). The sample of incumbent teachers, however, could be endogenous: The decision not to move nor exit from a given type of district could be correlated with unobservable teacher characteristics also related to VA. To address this issue I re-estimate equation (14) with teacher fixed effects in order to capture the change in VA within each teacher. This estimate is estimated imprecisely; its magnitude,

Figure 10: Selection vs. Effort. Value-Added, FP vs. SP, 2008–2015



Notes: OLS/IV estimates and 90 percent confidence intervals of the coefficients β_s in the regression $VA_{it} = \alpha FP_j + \sum_{s \neq 2011} \beta_s FP_j * \tau_s + \gamma_1 X_{it} + \gamma_2 Z_{jt} + \tau_t + \varepsilon_{it}$, where VA_{it} is VA of teacher i employed in district j in time t, FP_j equals one for individual-salary districts, X_{it} includes indicators for the number of years of experience and the highest education degree, Z_{jt} are district controls (including the 2009–2011 averages of district-level characteristics interacted with year fixed effects, and indicator for whether the district had a union recertification election in year t and whether the election was successful), and τ_t are year fixed effects. The solid line coefficients ("Selection + Effort") are estimated via OLS. The dashed line coefficients ("Effort (IV)") are estimated using IV. The sample is restricted to tenured teachers (with more than three years of experience) working in FP and SP districts. Standard errors are clustered at the district level.

however, is consistent with the above decomposition (point estimate equal to 0.052 standard deviations, Table 7, column 3). Estimates on the matched sample of FP and SP districts are larger and more precise (Table 7, columns 4-6).

The above estimates are all based on a measure of VA that is allowed to change after Act 10. The endogenous sorting of teachers after the reform could generate non-classical measurement error in VA, for example if teachers who move to FP districts are assigned to systematically different students (within schools) than movers to SP districts. Non-classical measurement error in the dependent variable would lead to biased estimates of β; the direction and extent of the bias depend on the sign and magnitude of the correlation between the error and FP_j. Appendix Table A15 tests for changes in the characteristics of the students of the teachers who move to FP and SP districts, in the year in which they move, before/after Act 10 and controlling for school fixed effects.[45] Movers to FP districts serve more low-SES students than movers to SP districts; no differences can be registered in other student characteristics. A possible explanation for this is that movers to SP districts after Act 10 are assigned slightly better students within each

[45]The table shows estimates of γ in the equation $Z_{ijt} = \gamma FP_t * post_t + \sigma_{s(jt)} + \tau_t + \varepsilon_{ijt}$, where Z_{ijt} are characteristics of the students of teacher i, who moves to district j in year t, and σ_s are school fixed effects. Z_{ijt} include the share of low-SES, Black, Hispanic, and English-language learners.

school for a compensating differentials motive (since SP districts do not use their flexibility to compensate teachers). This pattern of sorting could lead to overestimating the VA of movers to SP districts relative to movers to FP districts, which would bias the estimate of β towards zero. Lastly, no evidence exists of sorting of *students* across districts: the share of students who change district remains constant around 2011 (Appendix Figure A12).

Taken together, these results indicate that a change in teacher pay from one based on seniority to one that rewards quality affects both the composition of the teaching workforce and teachers' effort. Since a one standard deviation increase in VA leads to a 0.2 standard deviations increase in student test scores (Chetty et al., 2014a), my estimates imply a 0.02 standard deviations improvement in test scores in FP districts relative to SP districts after Act 10.[46]

The estimated increase in effort is in partial contrast to some existing works that show no effects of financial incentives on teachers' effort and productivity (Goodman and Turner, 2013; Fryer, 2013; de Ree et al., 2018) and that conclude that alternative hiring and firing practices are the only effective policies to improve teachers' quality (Staiger and Rockoff, 2010; Rothstein, 2014). My findings are, however, based on a substantially different policy change, which does not simply involve bonuses (such as Goodman and Turner, 2013; Fryer, 2013) or across-the-board salary increases (de Ree et al., 2018), but instead dramatically and permanently changes the entire structure of teacher pay.

7 A Model of the Teachers' Labor Market

The evidence presented so far shows that the introduction of flexible pay in a subset of Wisconsin districts led to an improvement in the composition of the teaching workforce in these districts compared with the rest of the state. Albeit small in the short run, this effect could become larger over time as more low-quality teachers leave FP districts and more high-quality teachers move from SP to FP districts. This, however, assumes that SP districts maintain a seniority-based pay scheme in the medium and long-run. What would happen if, instead, flexible pay were introduced in *all* districts?

The answer to this question is key to assessing the general-equilibrium effects of policies designed to attract and retain high-quality teachers. The selection patterns outlined above, however, are the combination of both demand and supply forces; it is therefore difficult to provide

[46] As a benchmark, a reduction in class size from 22–25 to 13–17 students (35-40 percent) leads to a 0.2 standard deviations increase in test scores (Krueger, 1999).

an answer extrapolating from these partial-equilibrium results.

To address the limitations of the reduced-form approach, I build and estimate a model of the teachers' labor market and I use it to simulate the effects of alternative salary schemes on the composition of the teaching workforce.[47] The model is an extended version of a simple Roy model with endogenous labor demand. Utility-maximizing teachers supply labor to districts; districts hire teachers to maximize a payoff function that depends on teachers' characteristics, subject to budget and capacity constraints. Salaries are exogenously determined and district-specific. In equilibrium, each side of the market maximizes its payoffs taking the choices of all other agents as given. Matches are formed as a result.[48]

For the sake of tractability, the model does not fully capture all the features of teachers' labor markets. First, teachers cannot choose effort (or decide where to teach based on how hard they want to work), and they cannot have a comparative advantage in teaching certain districts or students. The outside option is fixed across teachers, and the model does not allow for endogenous entry.[49] Lastly, I assume that salaries are exogenously determined, ruling out the possibility that districts set them strategically. Despite these limitations, the model is able to capture and replicate the sorting patterns observed in the data and can be used to study the effects of counterfactual pay schemes on the composition of the teaching workforce.

7.1 Model Setup

The framework is a two-sided static choice model in which job vacancies and salaries are exogenously determined. Matching between teachers and districts happens in two steps. First, each district decides whether to make an offer to each teacher. Each teacher then reviews her offers and either chooses the one that maximizes her utility or leaves the market. Job matches are realized as a result.

[47]Older studies of teachers' labor markets, such as Antos and Rosen (1975), estimate teacher labor supply using a hedonic-salaries approach based on the consideration that, if salaries are set to clear the market, then the salaries and the teacher-district matches observed in equilibrium are implied by (and can be used to derive) the preferences of teachers and districts. Teacher salaries, however, are typically rigid and unable to fully adjust for differences in either workers' characteristics or the non-pecuniary attributes of their jobs. Hedonic models are hence not appropriate for this setting.

[48]This model is similar to that of Boyd et al. (2013), who use many-to-one matching to estimate teachers' and schools' preferences. My paper builds on this approach in two ways. First, I model districts' choices as the outcome of a constrained maximization problem, explicitly incorporating a budget constraint and a capacity constraint. Second, I exploit the unique variation in salaries introduced by Act 10 (documented above) to estimate the parameters.

[49]Previous studies on this topic include, among others: Dolton and Van der Klaauw (1999), who affirm the importance of salaries and opportunity wages in teachers' turnover decisions and illustrate the insight gained from differentiating between multiple destinations or exit types; Boyd et al. (2005); and Goldhaber et al. (2011), who find heterogeneity in mobility behavior across the performance distribution and evidence that teacher mobility is affected by student demographics and achievement.

Districts' Problem. District j's payoff from hiring teacher i, u_{ij}, is a function of teacher i's characteristics such as experience, education, and VA.[50] The total district payoff is the sum of teacher-specific payoffs across all hired teachers. Each district decides which teachers are to be extended job offers. This choice is summarized by the vector $\mathbf{o}_j = [o_{1j}, o_{2j}, ..., o_{N_j}]$, where $o_{ij} = 1$ if district j extends a job offer to teacher i, and N is the number of teachers. Lastly, district j can spend up to B_j in salaries and can hire up to H_j teachers. District j's problem is as follows (I omit the subscript j for ease of notation):

$$\max_{\mathbf{o}} \sum_{i=1}^{N} h_i o_i u_i \tag{15}$$

$$\text{s.t.} \qquad \sum_{i=1}^{N} h_i o_i w_i \leq B, \sum_{i=1}^{N} h_i o_i \leq H, o_i \in \{0,1\} \; \forall i = 1, .., N \tag{16}$$

where w_i is the salary paid to teacher i, and h_i is the probability that teacher i accepts the district's offer, if one is made. In other words, each district maximizes the *expected* payoff from making a set $\mathbf{o}$ of offers, with respect to the probability of acceptance. Constraints in (16) are "soft," i.e. they must only hold in expectation. Intuitively, districts incorporate the fact that an offer made to a teacher i is only accepted with probability h_i. Since offers are made simultaneously, districts choose the offer set that maximizes their expected payoff and, in expectation, allows them to spend at most B and hire at most H teachers.

Salaries. In keeping with the reduced-form analysis I assume that salaries are not competitive, i.e. they do not adjust to equate demand and supply in equilibrium; they are instead exogenously determined and district-specific. The advantage of this assumption is that it makes the model more tractable and realistic. If salaries were competitive, in practice all districts should have switched to flexible pay after Act 10; yet this did not happen.[51] The drawback of this assumption is that it rules out the possibility that each district's salary structure is dependent on other endogenous variables of the model, for example the pre-Act 10 composition of the teaching workforce.

Teacher's Problem. Teachers have preferences over job characteristics. In each period they receive a set of offers O_i from school districts and choose the one that maximizes their utility. I define the utility of teacher i from working in district j as v_{ij}. Each teacher faces an outside

[50] This framework can be reconciled with one in which districts maximize a function of student achievement, which is in turn an additive function of teacher characteristics.

[51] If salaries were competitive, one could simply use the hedonic approach of Antos and Rosen (1975) to estimate teachers' preferences.

option, with an associated utility $v_{i0} = v_0$. The teacher's problem can be expressed as follows:

$$\max_{k \in O_i \cup \{0\}} v_{ik} \tag{17}$$

7.1.1 Equilibrium

The equilibrium of the model can be defined as a set of offers $\mathbf{o}^* = [\mathbf{o}_1^*, \mathbf{o}_2^*, ..., \mathbf{o}_J^*]$, where J is the number of districts, such that all agents in the market make the choice that is optimal for them given all other agents' optimal choices. The equilibrium can be formally defined as follows:

$$\begin{aligned}
\forall j, \mathbf{o}_j^* &\in \arg\max_{\mathbf{p}_j} \sum_{i=1}^{N} h_{ij} p_{ij} u_{ij} \text{ s.t. constraints} \\
h_{ij} &= \begin{cases} P(v_{ij} \geq \max_{k \in O_i \cup \{0\}} \{\{v_{ik}\}_{k \neq j}, v_{i0}\}) & \text{if } j \in O_i \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

7.2 Estimation

To estimate the parameters of teachers' utility and districts' payoffs I make the following assumptions.

Districts. They have identical and linear payoffs from hiring teacher i: $u_{ij} = \beta x_i + \varepsilon_{ij}$, where β is a vector of parameters, and the vector x_i includes teacher VA, years of experience, and an indicator for having a master's degree. The variable ε_{ij} is an idiosyncratic component, independent across teachers and districts and identically and normally distributed with mean 0 and variance σ^2. This formulation implicitly assumes that each district is able to perfectly observe all teachers' attributes, including VA.

Each district's problem can be solved using linear programming techniques. The problem is analogous to a two-constraint version of the 0-1 knapsack problem (Dantzig, 1957). I solve it using the algorithm of Martello and Toth (2003), based on the "continuization" of the discrete problem. The algorithm is detailed in Appendix D.

Teachers. They have identical and linear preferences from a job in j: $v_{ij} = \alpha z_{ij} + \xi_{ij}$, where α is a vector of utility parameters. The vector z_{ij} includes salary (in \$1,000), distance from the district where teacher i is an incumbent (in miles), an indicator for teacher i being an incumbent in district j (which captures the cost of moving across districts, assumed constant across teachers and districts), the share of disadvantaged students, and an indicator for urban districts. The variable ξ_{ij} is an idiosyncratic utility component, independent across districts and identically distributed with an Extreme Value Type 1 distribution. Teacher i's utility from the outside option is constant in expectation and equal to $v_{i0} = \alpha_0 + \xi_{i0}$, where ξ_{i0} is independent across teachers, orthogonal to ξ_{ij}, and identically distributed with an Extreme Value Type 1 distribution.

Salaries. Estimating teachers' preferences requires observing the characteristics of all the job alternatives available to a teacher, including salaries. In the data, however, I only observe salaries when a match is realized. To construct salary offers for unrealized matches, I back out each district's post-Act 10 salary structure by estimating a wage function separately for each district:

$$w_{ijt} = \gamma_{0j} + \gamma_j f(X_{it}) + \delta_j VA_{it} + \omega_{ijt} \tag{18}$$

where X_{it} is a full set of interactions between indicators for two-year seniority classes and highest education degree, and VA_{it} is time-varying VA of teacher i.

Budget and Capacity Constraints. I construct each district's budget limit by multiplying the previous year's total salary bill by the pre-Act 10 growth rate of total salaries. Similarly, I construct the capacity limit by multiplying the district's enrollment in the previous year by the average, district-specific number of teachers per student in the years until 2011.

7.2.1 Estimation Procedure

I first estimate the salary parameters γ_{0j}, γ_j, and δ_j in Equation (18) outside of the model and separately for each district, using OLS and data on post-reform teacher-district matches. I use these estimates to back out salaries for each teacher in each district. I then estimate the parameter vectors α (teachers' utility), α_0 (teachers' outside option), β (districts' payoff), and σ^2 (variance of the district's shock) using maximum likelihood. I divide Wisconsin into 12 separate geographic labor markets, corresponding to the 12 Cooperative Educational Service Agencies (CESAs), and I exclude the CESAs of Milwaukee and Madison. I assume that teachers can only move within CESAs and that districts can only make offers to teachers already working in their CESA.[52] I estimate the parameters using data from 2014. The final sample contains 12,573 tenured teachers working in 410 districts.[53] Table A16 shows summary statistics of the estimation sample. The estimation procedure is outlined in more detail in Appendix D.

[52] In 2014, about 60 percent of teacher movements happened within a CESA.
[53] In order to fully capture movements, I exclude teachers whose previous district is missing in 2014.

7.3 Identification

The model allows for a transparent identification of the parameters of teachers' utility. Identification relies on cross-district heterogeneity in district characteristics (such as location and student composition) and on the variation in salaries introduced by Act 10. Movements of teachers across districts and exits help identify the utility parameters α and α_0.[54]

Identification of the parameters of districts' payoff function is more subtle. The parameters β and σ are identified out of cross-district variation in optimal offer strategies. While I assume that districts have identical preferences, their optimal strategies might differ due to differences in their budget and capacity constraints. These differences, in turn, arise from the attrition of different types of teachers over time. To see this, consider the following example. Suppose that districts *A* and *B* are identical in terms of student and teacher composition, size, salary structure, and *ex ante* budget. At a certain point in time they both lose one teacher and thus have one vacancy to fill. If district *A*'s exiting teacher has 30 years of experience (and was therefore being paid a high salary) but district *B*'s teacher only has one (and was being paid a lower salary), then district A has more money "freed up" (and therefore a larger budget) than district *B*. To the extent that the characteristics of leavers are random (which the reduced-form results show to be true before Act 10), the hiring choices of district *A*, compared with *B*, reveal how teacher attributes are valued, and they identify β and σ.[55]

7.4 Parameter Estimates and Elasticities

Table 8 shows estimates and standard errors of the model's parameters. Teachers receive positive utility from salary and negative utility from distance. A positive and significant estimate for the incumbent dummy indicates that fixed moving costs are important. Lastly, teachers prefer suburban and rural districts to urban ones (Table 8, column 1).

To interpret the magnitudes of these coefficients, I compare the elasticities between the prob-

[54]For example, suppose teacher *x* is an incumbent in district *A* where she earns a wage w_A. She receives an offer from district *B*, located five miles from *A* and offering a wage w_B, with $w_B > w_A$, and an offer from district *C*, located seven miles from *A* and offering a wage w_C, with $w_C > w_B$. The choice of teacher *x* identifies the parameters her utility. For example, if she chooses C, this implies that the desire for higher salaries offsets the drawback of a longer commute and translates into a higher utility parameter on salaries and a lower parameter on distance. Similarly, teachers' exit will identify the value of the outside option.

[55]Another useful example is the following. Suppose that, given teachers' preferences and districts' budgets, for given values of β and σ districts' optimal strategies are such that one district ends up hiring too many teachers and the other ends up hiring too few with respect to their capacity. To bring the market into equilibrium, β and σ need to adjust in order for each district to maximize its payoff and satisfy both constraints. Similarly, if the optimal strategies given teachers' preferences, districts' capacities, and given parameter values are such that one or both districts violate the budget constraint, β and σ need to adjust to bring the market back into equilibrium.

Table 8: Model Parameters

Teacher				District		
parameter	interpretation	estimate (1)	elasticity (2)	parameter	interpretation	estimate (3)
α	salary ($1,000)	0.0037	0.2117	β	VA	0.8291
		(0.0010)				(0.0005)
	distance	-0.6521	-0.1202		seniority	0.1099
		(0.0005)				(0.0019)
	incumbent	5.4838	0.9660		master's degree	0.9052
		(0.0007)				(0.0020)
	% disadvantaged	-0.0659	-2.4291	σ	s.d. shock	0.1111
		(0.0016)				(0.0015)
	urban	-0.0630	-2.2979			
		(0.0010)				
α_0	outside option	0.0003				
		(0.0402)				

Notes: Estimates of the parameters of the structural model. Parameters are estimated by maximum likelihood. Defining p_{ij} as the probability that teacher i moves to district j, the elasticity of p_{ij} to a continuous job characteristic z_{ij} (implied by the logit assumption on the error term of teachers' utility) is $\alpha_z(1-p_{ij})z_{ij}$, where α_z is the parameter estimate on z_{ij}. The elasticity of urban and incumbent is defined as $(1-p_{ij})(1-\exp(-\alpha_z))$. Elasticities are evaluated at the median of each variable, equal to $59,000 for salary, 0.19 miles for distance, and 38 percent for the share of disadvantaged students. Standard errors in parentheses are calculated as the square root of the inverse of the information matrix using numerical derivatives.

ability of matching with a district and various district characteristics, shown in column 2 of Table 8.[56] A 1 percent higher salary (equivalent to $590 at the mean) is associated with a 0.21 percent increase in the match probability (Table 8, columns 1-2). A 10 percent increase in distance is associated with a 1.2 percent lower probability.[57] Moving costs (which correspond to the opposite of the estimate of the incumbent dummy) are equal to approximately 4.56 percent of salary (0.9660/0.2117), or $2,690.

These elasticities allow me to assess the importance of job characteristics for teachers' labor supply and to calculate compensating differentials, i.e. the required salary increases to attract and retain teachers to districts with certain characteristics. For example, consider two identical districts that want to hire a teacher employed in another district, one of which is 10 miles farther from where the teacher is currently working. The estimated elasticities imply that the farther-away district must offer a large 28 percent larger salary to attract the teacher (since a two mile longer distance requires a 5.68 percent higher salary, or 10*0.1202/0.2117).

Estimates of the parameters of districts' payoffs imply that districts prefer higher VA teach-

[56]Defining p_{ij} as the probability that teacher i matches with district j, the elasticity of p_{ij} to a job characteristic z_{ij} implied by the logit assumption on the error term of teachers' utility is $\beta_z(1-p_{ij})z_{ij}$. The elasticities shown in the table are calculated at the mean of p_{ij} and z_{ij}.

[57]As a comparison, Levy and Wadycki (1974) analyze the migration patterns of a sample of Venezuelan workers and estimate a distance elasticity of -0.43 and an income elasticity of 1.9.

ers, as well as those with more experience. Districts are indifferent between a teacher who has one extra year of seniority or 0.13 standard deviations higher VA (0.1099/0.8291, Table 8, column 3).

8 Alternative Pay Schemes and Workforce Composition

The structure of the model and the parameter estimates can be used to simulate the effects of alternative salary schemes on the composition of the teaching workforce. I focus on two types of counterfactuals. The first is a change in the salary component associated with VA (captured by the coefficient δ in Equation (18) *only in one district*, with salaries unchanged in all other districts. The second is a change in δ *in all districts at the same time*. In both cases I model the change as budget neutral by letting base salaries adjust to the change in δ.[58]

8.1 Increase in Quality Pay in One District

I start by simulating the effect of a change in δ only in one district. This change affects both teachers' labor supply and demand. First, it affects the budget and the salaries paid by the district. Second, it affects the preference ordering of *all* teachers, including those employed in other districts. This will, in turn, influence the probability that a teacher matches with *any* district, not only with the one affected by the policy.

I first solve the model for values of δ ranging from zero to 1.5 standard deviations of δ. I then plot the change in the probability that teachers in different quartiles of the distribution of VA move to, move out of, or exit from the district, as well as the change in the composition of the district's teaching workforce. For exposition, I perform the analysis on the school district of Ashland, an urban SP district.[59]

Figure 11 illustrates the changes in the simulated probability of moving to, moving out of, or exiting the district as δ increases from zero to 1.5 standard deviations, by quartile of VA. Teachers in the bottom quartile are 2 percent less likely to move to Ashland when δ increases by one standard deviation (compared to when it is equal to zero); teachers in the top quartile are instead 25 percent more likely to move there (Figure 11, panel A). In addition, teachers with VA in the bottom quartile are 6 percent more likely to move out, whereas those with VA in the

[58]To keep the budget neutral, I assume that base salaries, captured by γ_{0j} in Equation (18), adjust immediately depending on the new value of δ and the current composition of the district's teaching workforce.

[59]The school district of Ashland is located in the north-west part of the state. This urban district runs five schools, including four elementary and middle schools; it enrolled 2,101 students in 2014 year, 61 percent of whom are economically disadvantaged, and employs 35 teachers in my sample in 2014.

top quartile are 6 percent less likely (Figure 11, panel B). Lastly, teachers with VA in the bottom quartile are 6 percent more likely to exit and teachers in the top quartile are 6 percent less likely (Figure 11, panel C).

Figure 12 shows the change in the average VA of teachers moving in, moving out, and exiting the school district for different values of δ (panel A) and the change in the overall composition of the district's teaching workforce (panel B). The figure shows that the average VA of movers in the district increases slightly, the average VA of movers out of the district decreases slightly, and the average VA of exiters from the district declines more sharply. As a result, the overall composition of the district's workforce improves by 0.2 percent when δ increases by one standard deviation (Figure 12).

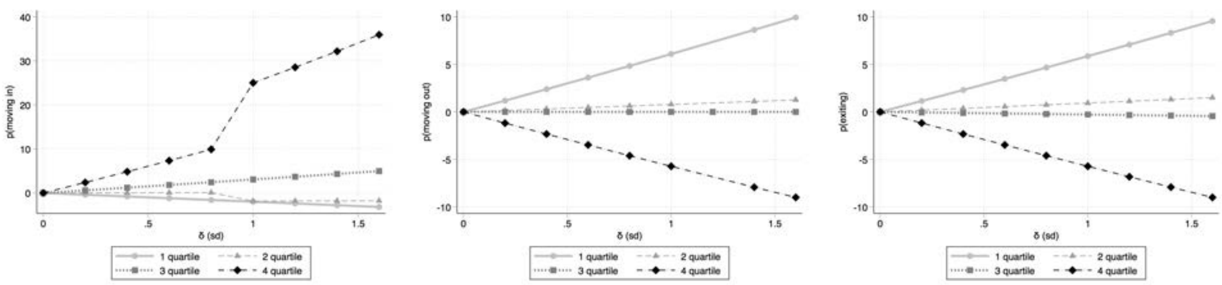
The results from this simulation exercise are in line with the reduced-form results: An increase in the share of salaries related to teacher quality is associated with an improvement in the composition of the district's teaching workforce, which in the case of the Ashland district is very small. This improvement is driven by higher VA teachers moving to the district from neighboring districts, attracted by higher salaries, and by lower VA teachers moving out to other districts or leaving teaching altogether.

8.2 Introduction of Quality Pay in All Districts

I now simulate the compositional effect of a change in δ in *all* districts. The results from this second counterfactual exercise do not trivially follow from the first. To see this, consider

Figure 11: Counterfactual 1 - Teacher Responses to an Increase in δ in One District

Panel A: Movements to the district | Panel B: Movements out of the district | Panel C: Exits from the district



Notes: Percentage change in the probability that a teacher moves to the Ashland school district (panel A), out of the district (panel B), or exits from the district (panel C), by quartile of VA, and for different values of δ (as defined in Equation (18)), relative to $\delta = 0$, under the first counterfactual.

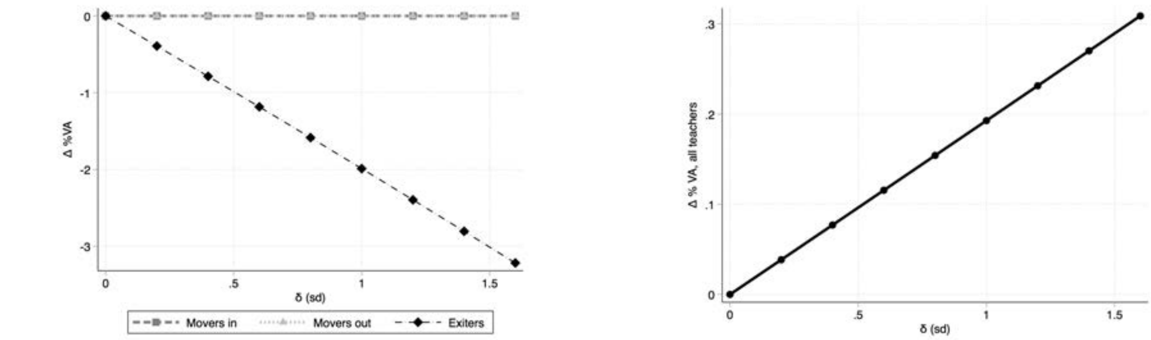
the decision of a teacher working in Ashland. She must decide whether to stay where she is, move to another district, or exit teaching. The first counterfactual directly affects the first option (staying). The second counterfactual directly affects two out of three options (because it affects salaries in all districts); it therefore also changes the value of leaving relative to remaining in public schools. As a result, the effect of a change in salaries in all districts on the exit behavior of teachers could in principle be very different from the one outlined in the previous subsection.

Results from this simulation indicate that teachers with VA in the three bottom quartiles are

Figure 12: Counterfactual 1 - Compositional Changes

Panel A: VA of movers in, movers out, and exiters from the districts

Panel B: VA of all teachers in the districts



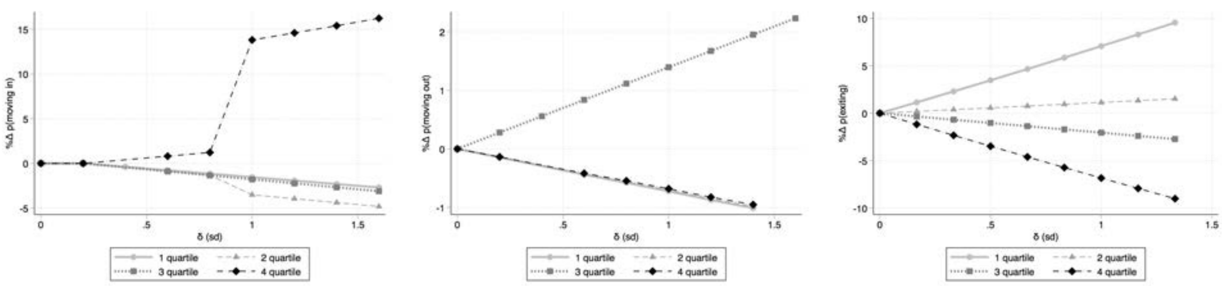
Notes: Percentage change in average VA of teachers moving to the Ashland school district, out of the district, and exiting public schools from the district (panel A), and average VA of teachers working in the district (panel B), for different values of δ (as defined in Equation (18)), relative to $\delta = 0$, under the first counterfactual.

Figure 13: Counterfactual 2 - Teacher Responses to an Increase in δ in All Districts

Panel A: Movements to the district

Panel B: Movements out of the district

Panel C: Exits from the district



Notes: Percentage change in the probability that a teacher moves to the Ashland school district (panel A), out of the district (panel B), or exits from the district (panel C), by quartile of VA, and for different values of δ (as defined in Equation (18)), under the second counterfactual.

less likely to move to Ashland when δ increases by one standard deviation *in all districts* (for example, teachers in the second quartile are 4 percent less likely), whereas teachers with VA in the top quartile are 14 percent more likely to move there (Figure 13, Panel A). Teachers with VA in the top and bottom quartile, however, are also equally less likely to move out of Ashland, whereas teachers in the third quartile are 2 percent more likely. Lastly, teachers with VA in the bottom quartile are 7 percent more likely to exit, and teachers with VA in the top quartile are 6 percent less likely (Figure 13, panel C).

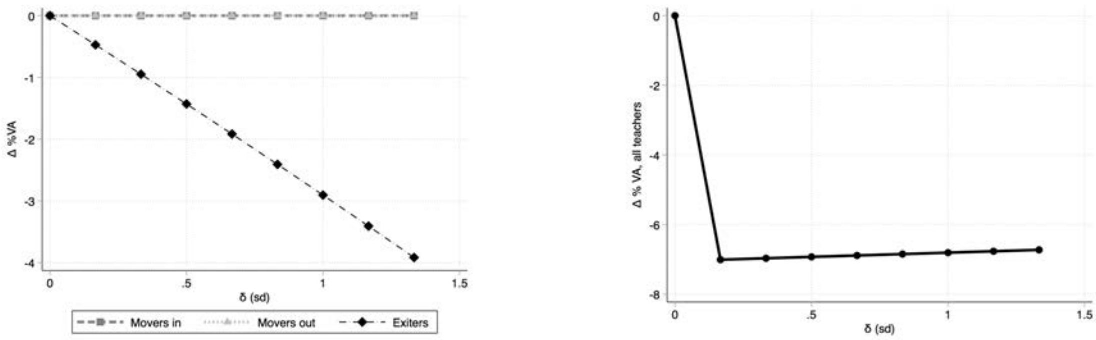
Figure 14 shows the average VA of teachers moving in, out, and exiting the school district for different values of δ and the overall composition of the district's teaching workforce. The figure shows that the composition of the teaching workforce in Ashland worsens by 7 percent of a standard deviation of VA when δ increases by one standard deviation (Figure 14, panel B).

The results from these simulations show that an increase in the quality component of salaries in all districts could lead to a much different change in the composition of a district's workforce compared with the case in which δ only increases only in one district. The reason is that, when quality pay increases only in one district, part of the resulting compositional improvement is driven by better teachers moving in and worse teachers moving out. When δ increases in all districts, this net inflow of high-quality teachers might be absent because quality is rewarded at the same rate everywhere.

Figure 14: Counterfactual 2 - Compositional Changes

Panel A: VA of movers in, movers out, and exiters from the districts

Panel B: VA of all teachers in the districts



Notes: Percentage change in average VA of teachers moving to the Ashland school district, out of the district, and exiting public schools from the district (panel A), and average VA of teachers working in the district (panel B), for different values of δ (as defined in Equation (18)), relative to $\delta = 0$, under the second counterfactual.

The results from the two counterfactuals suggest that the observed improvement in the composition of the teaching workforce in FP districts might be limited to the short run. If all districts eventually introduce merit pay in order to compete for the best teachers, the longer-term effects of Act 10 in each district, and in the whole state, might be more limited in size.

When interpreting the results of these simulations, a few important caveats apply. First, due to the impossibility of accurately measuring the quality of new teachers, the model does not incorporate entry and implicitly assumes that the quality of entrants is constant and equal to the quality of incumbents. The new pay scheme, however, could change workers' incentives to enter public school teaching, either by attracting more talented workers or by discouraging risk-averse workers. Second, the model does not capture the effects of a flexible pay scheme on teachers' effort. Even if the compositional changes become smaller as more districts switch to FP, overall teacher productivity could rise in response to this change in pay.

9 Discussion and Conclusion

The role of teachers' unions and the powers enjoyed by these associations have come under scrutiny in recent times, culminating with the Supreme Court decision in *Janus v. AFSCME*. Given the importance of individual teachers in shaping children's educational opportunities (Rockoff, 2004), policies affecting teachers' labor markets can have very large effects on students. This paper provides an initial assessment of these effects by exploiting a recent change in the scope of CB for teachers' unions that has only affected one US state so far, but that could be replicated in other states in the near future.

I exploit this policy change to assess its effects on the composition of the teaching workforce and to study teachers' labor supply and demand. A switch away from seniority pay toward flexible pay in a subset of Wisconsin districts, following the interruption of CB on teachers' salary schedules mandated by Act 10 of 2011, resulted in high-quality teachers moving to FP districts and low-quality teachers either moving to SP districts or leaving the public school system altogether. As a result, the composition of the teaching workforce improved in FP districts compared with SP districts. Effort exerted by all teachers also increased.

As cross-district movements and exits are rare events, the magnitudes of these compositional changes (and the associated increase in student test scores) are limited in size in the short run, but they could become larger over time as more teachers move and exit each year and as salaries become more strongly related to teachers' quality and effort. If, however, SP districts also switch to a FP scheme over time, the long-run effects of a policy change such as Act 10 could be very different. To understand what would happen under this scenario, I estimate teachers' labor demand and supply using a structural model of this labor market and I use the variation in post-Act 10 salary schemes across districts, as well as teachers' movements and exits, to identify the model's parameters. Simulations of this model on alternative pay schemes show that the introduction of flexible pay in all districts would lead to a much smaller (and possibly negative) compositional improvement than the one experienced by FP districts so far. This suggests that the observed gains in teacher composition and achievement in FP districts might be short lived and that the longer-term effects on each district (and on the whole state) might be more contained.

While this paper has focused on movements of teachers across districts and on exits from the profession, the effects of a policy such as Act 10 on the supply of new teachers could also be important and represent an interesting and important avenue for future research.

References

Aaronson, D., L. Barrow, and W. Sander (2007). Teachers and student achievement in the Chicago Public High Schools. *Journal of Labor Economics 25*(1), 95–135.

Abadie, A. and G. W. Imbens (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics 29*(1), 1–11.

Abramitzky, R. (2009). The effect of redistribution on migration: Evidence from the Israeli Kibbutz. *Journal of Public Economics 93*(3), 498–511.

Adnot, M., T. Dee, V. Katz, and J. Wyckoff (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis 39*(1), 54–76.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy 113*(1), 151–184.

Antos, J. R. and S. Rosen (1975). Discrimination in the market for public school teachers. *Journal of Econometrics 3*(2), 123–150.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater, and D. Wilson (2009). Evaluating the impact of performance-related pay for teachers in England. *Labour Economics 16*(3), 251–261.

Ballou, D. and M. J. Podgursky (1997). *Teacher pay and teacher quality*. WE Upjohn Institute.

Bandiera, O., I. Barankay, and I. Rasul (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics 120*(3), 917–962.

Biasi, B. (2017). Retirement age effects of pension and salary reforms: Evidence from Wisconsin teachers.

Boyd, D., P. Grossman, H. Lankford, S. Loeb, and J. Wyckoff (2008). Who leaves? teacher attrition and student achievement. Technical report, NBER Working paper n. 14022.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review 95*(2), 166–171.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2013). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers. *Journal of Labor Economics 31*(1), 83–117.

Brehm, M., S. A. Imberman, and M. F. Lovenheim (2017). Achievement effects of individual performance incentives in a teacher merit pay tournament. *Labour Economics 44*(C), 133–150.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? evidence from Project STAR. *The Quarterly Journal of Economics 126*(4), 1593–1660.

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

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9), 2633–2679.

Clotfelter, C., E. Glennie, H. Ladd, and J. Vigdor (2008). Would higher salaries keep teachers in high-poverty schools? evidence from a policy intervention in North Carolina. *Journal of Public Economics 92*(5), 1352–1370.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources 45*(3), 655–681.

D'Andrea, C. (2013). Limits on collective bargaining. *Education Next 13*(4).

Dantzig, G. B. (1957). Discrete-variable extremum problems. *Operations Research 5*(2), 266–288.

de Ree, J., K. Muralidharan, M. Pradhan, and H. Rogers (2018). Double for nothing? experimental evidence on an unconditional teacher salary increase in indonesia*. *The Quarterly Journal of Economics 133*(2), 993–1039.

Dee, T. S. and B. J. Keys (2004). Does merit pay reward good teachers? evidence from a randomized experiment. *Journal of Policy Analysis and Management 23*(3), 471–488.

Dee, T. S. and J. Wyckoff (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management 34*(2), 267–297.

Dobbie, W. (2011). Teacher characteristics and student achievement: Evidence from teach for america. *Unpublished manuscript, Harvard University*.

Dolton, P. and W. Van der Klaauw (1999). The turnover of teachers: A competing risks explanation. *Review of Economics and Statistics 81*(3), 543–550.

Dolton, P. J. (1990). The economics of UK teacher supply: The graduate's decision. *The Economic Journal 100*(400), 91–104.

Duflo, E., R. Hanna, and S. P. Rya (2012). Incentives work: Getting teachers to come to school. *American Economic Review 102*(4), 1241–1278.

Eberts, R. W. and J. A. Stone (1987). Teacher unions and the productivity of public schools. *Industrial and Labor Relations Review 40*.

Figlio, D. N. (2002). Can public schools buy better-qualified teachers? *Industrial & Labor Relations Review 55*(4), 686–699.

Figlio, D. N. and L. W. Kenny (2007). Individual teacher incentives and student performance. *Journal of Public Economics 91*(5), 901–914.

Fitzpatrick, M. D. and M. F. Lovenheim (2014). Early retirement incentives and student achievement. *American Economic Journal: Economic Policy 6*(3), 120–54.

Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from New York City Public Schools. *Journal of Labor Economics 31*(2), 373–407.

Glewwe, P., N. Ilias, and M. Kremer (2010). Teacher incentives. *American Economic Journal: Applied Economics 2*(3), 205–227.

Goldhaber, D., B. Gross, and D. 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.

Goodman, S. F. and L. 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.

Han, E. S. (2016). The myth of unions' overprotection of bad teachers: Evidence from the district-teacher matched panel data on teacher turnover.

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review 61*(2), 280–288.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2004). Why public schools lose teachers. *Journal of Human Resources 39*(2), 326–354.

Hensvik, L. (2012). Competition, wages and teacher sorting: Lessons learned from a voucher reform. *The Economic Journal 122*(561), 799–824.

Hoxby, C. M. (1996). How teachers' unions affect education production. *The Quarterly Journal of Economics 11*(3), 671–718.

Hoxby, C. M. and A. Leigh (2004). Pulled away or pushed out? explaining the decline of teacher aptitude in the united states. *American Economic Review 94*(2), 236–240.

Imberman, S. A. and M. 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.

Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics 27*(2), 213–256.

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.

Jackson, C. K., J. E. Rockoff, and D. O. Staiger (2014). Teacher effects and teacher-related policies. *Annual Review of Economics 6*(1), 801–825.

Kane, T. J. and D. O. Staiger (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, NBER working paper n. 14607.

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2015). Tax farming redux: Experimental evidence on performance pay for tax collectors. *The Quarterly Journal of Economics 131*(1), 219–271.

Kimball, S. M., H. Heneman, R. Worth, J. Arrigoni, and D. Marlin (2016). Teacher compensation: Standard practice and changes in Wisconsin. *Madison, WI: Wisconsin Center for Education Research*.

Kraft, M. A. (2017). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, 0916–8265R3.

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

Krueger, A. B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics 114*(2), 497–532.

Ladd, H. F. (1999). The Dallas School Accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review 18*(1), 1–16.

Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on Optimization 9*(1), 112–147.

Lavy, V. C. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy 110*(6), 1286–1317.

Lazear, E. P. (2000a). Performance pay and productivity. *American Economic Review*, 1346–1361.

Lazear, E. P. (2000b). The power of incentives. *American Economic Review 90*(2), 410–414.

Levy, M. B. and W. J. Wadycki (1974). What is the opportunity cost of moving? reconsideration of the effects of distance on migration. *Economic Development and Cultural Change 22*(2), 198–214.

Litten, A. (2016). The effects of public unions on compensation: Evidence from Wisconsin. Working paper.

Lovenheim, M. F. (2009). The effect of teachers' unions on education production: Evidence from union election certifications in three Midwestern states. *Journal of Labor Economics 27*(4), 525–587.

Martello, S. and P. Toth (2003). An exact algorithm for the two-constraint 0-1 knapsack problem. *Operations Research 51*(5), 826–835.

Moe, T. M. (2013). *A primer on America's schools*, Volume 486. Hoover Institution Press.

Muralidharan, K. and V. Sundararaman (2011). Teacher performance pay: Experimental evidence from india. *Journal of Political Economy 119*(1), 39–77.

Neal, D. et al. (2011). The design of performance pay in education. *Handbook of the Economics of Education 4*, 495–550.

Papay, J. P. and M. 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.

Podgursky, M. (2006). Teams versus bureaucracies: Personnel policy, wage-setting, and teacher quality in traditional public, charter, and private schools. *Education Working Paper Archive*.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review 94*(2), 247–252.

Ronfeldt, M., S. Loeb, and J. Wyckoff (2013). How teacher turnover harms student achievement. *American Educational Research Journal 50*(1), 4–36.

Roth, J. (2017). Union reform and teacher turnover: Evidence from Wisconsin's Act 10.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics 125*(1), 175–214.

Rothstein, J. (2014). Teacher quality policy when supply matters. *American Economic Review 105*(1), 100–130.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Sojourner, A. J., E. Mykerezi, and K. L. West (2014). Teacher pay reform and productivity panel data evidence from adoptions of q-comp in Minnesota. *Journal of Human Resources 49*(4), 945–981.

Springer, M. G., D. Ballou, L. Hamilton, V.-N. Le, J. Lockwood, D. F. McCaffrey, M. Pepper, and B. M. Stecher (2011). Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching (POINT). *Society for Research on Educational Effectiveness*.

Staiger, D. O. and J. E. Rockoff (2010). Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives 24*(3), 97–117.

Stinebrickner, T. R. (2001). A dynamic model of teacher labor supply. *Journal of Labor Economics 19*(1), 196–230.

Todd, P. E. and K. I. Wolpin (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review 96*(5), 1384–1417.

Willén, A. L. (2018). From a fixed national pay scale to individual wage bargaining: The labor market effects of wage decentralization.