Shortage Field Incentives: Impacts on Teacher Retention and Recruitment

Rachel B. Rosen

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

© 2012
Rachel B. Rosen
All Rights Reserved

Table of Contents

List of Table and Figures

Tables

Figures

Acknowledgements

I am especially grateful to Doug Ready, Judith Scott-Clayton, Jonah Rockoff, and Henry Levin for their invaluable advice and feedback. I also thank the Dissertation Working Group: Brooks Bowden, Mina Dadgar, Michelle Hodara, Aleksander Holod, Megan Silander, and Miya Warner for ongoing feedback and support. I thank Tom Bailey and the Community College Research Center for providing me access to the SASS data, and Matt Zeidenberg for data assistance. This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under Grant #DRL-0941014. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

On a more personal note, I thank my parents for always believing in me, and supporting me in every way possible. I also thank my faithful friend Miley, who reminds me daily to get out and smell the flowers.

Chapter 1

Introduction

Currently, most teachers in the United States are paid on a single-salary schedule that provides uniform increases for additional credentials and years of experience to all teachers in a given district. These kinds of contracts were mainly instituted during the 1940s as a way to equalize pay between elementary and high school teachers, with additional pay for credentials serving as an appeasement to secondary teachers, more of whom were male and held masters degrees (Murphy, 1990). Although they served for many years to equalize gender disparities in pay, in the last several decades researchers have begun to question whether an unintended consequence of the single-salary schedule has been a reduction in quality of teachers who both enter and remain in teaching. In addition, although individual teacher salaries are typically calculated by a combination of credentials and experience, another problematic feature of these compensation schemes is the fact that there is usually no pay differentiation between teachers' skills and knowledge. Because all teachers with similar experience and education within a district are treated equally in terms of pay, teacher salaries are unresponsive to labor market realities in which workers command different wages, depending on the demand for specific skills and knowledge. By contrast, single salary schedules fail to provide incentives for teachers with skills that are in-demand in non-teaching fields, such as in mathematics and science. Despite these issues, for the most part the salary schedule remains an intractable feature of the majority of teacher contracts.

Meanwhile, there is extensive evidence to suggest that a chronic shortage of teachers in fields like mathematics and science, as well as foreign languages, bilingual and special education, do exist and has persisted from at least the 1940s until present day (Ingersoll & Perda, 2010). Basic supply and demand theory suggests that the market response to shortage is for prices to increase. Thus, if teacher labor markets were unconstrained, the wage for teachers in shortage fields would rise. However, because of the salary schedule, instead there is market failure in which shortages persist, seemingly indefinitely. To correct these specialized teacher shortages, economists have recommended incentivizing teacher salaries to be more responsive to market conditions, by paying premiums for qualified teachers in shortage fields (Kershaw & McKean, 1962; Levin, 1985).

However, there has been strong resistance to changing the single salary schedule. Although this resistance has frequently been attributed to union resistance, these compensation schemes, which came into favor many decades before public school teachers gained collective bargaining rights, are also common across right-to-work states and remain an intractable feature of the way teachers are paid. Therefore, historically few districts have implemented pay plans that are designed to address shortages in specific teaching fields, and there is very little empirical evidence to support the efficacy of labor market incentives for either improving recruitment or increasing retention of shortage field teachers.

It is this gap in the literature that this dissertation seeks to address. I begin by addressing whether shortage field incentives work as theory predicts they should; namely, I examine whether the minority of public school districts that do offer shortage-field incentives report better recruitment conditions and experience higher rates of retention of shortage field teachers than comparable districts that do not offer incentives. I also explore whether incentives serve as an effective mechanism for raising the quality of shortage field teachers recruited into districts. Once I explore these two issues between districts, I will further explore whether these two outcomes respond to changes within district policy over time.

Review of the Literature

Are There Teaching Shortages?

Beginning in the 1980s, American public schools began to grapple with reports that predicted a "perfect storm": an aging teacher population heading toward retirement would meet historic increases in student enrollments. In the three decades since those warnings were first sounded, a teacher shortage has surfaced as predicted, but its causes have now been attributed to the protractedly short careers of many new teachers, rather than to a confluence of demographic trends. The problem is not one of supply, rather each year large numbers of teachers leave their public sector teaching jobs for reasons other than retirement (Ingersoll, 2001). This phenomenon means that teacher attrition patterns are u-shaped, such that the greatest turnover occurs in the early years, and toward retirement (Dolton & van der Klaauw, 1999).

Although teacher attrition around retirement is to be expected, much of the policy concern is focused on early career teachers who exit the profession at high rates. For example, Ingersoll estimates that within the first five years of teaching, approximately 46 percent of new teachers will leave the classroom (Ingersoll, 2001, 2003). Ingersoll derives his estimates from four waves of the Schools and Staffing Survey, collected between 1987 and 2000. Using these data, he estimates that after the 2nd year, almost 24 percent of new teachers have left teaching, but more recent findings from the 2009-2010 school year places the two-year new teacher attrition rate at 12.5 percent (Kaiser & Cross, 2011), or half the estimate from the previous decade. At present, however, it is unclear whether the more recent findings are due to long-term changes in teacher attrition trends, or if the lower rates are a temporary outgrowth of overall high unemployment rates in the national labor market.

Although the number reported by Ingersoll is frequently cited and discussed as being too high, whether rates of turnover amongst young teachers are actually higher than the turnover rates for other early career workers is debatable. For example, Loprest finds that the average voluntary job change rate for workers within their first four years is between 22 and 24 percent (Loprest, 1992), while Royalty estimates that after the first year of working, just 55 percent of all young workers remain in the same job (Royalty, 1998). However, she also finds that more educated workers under 30 have lower average turnover rates of between 18-20 percent per job tenure. Together, these findings suggest that there is variation in turnover of young workers that is likely intimately linked to both sector, and worker qualifications.

To that end, other studies seek to make direct comparisons between turnover in teaching and non-teaching careers. For example, one study found that young teachers are more likely to leave teaching than young nurses are to leave nursing, which is thought to be a comparable profession due to high participation of women, and the social service nature of the work involved (Ingersoll, 2001). However, Harris and Adams (2007) also compared teachers to nurses, as well as to social workers and accountants, and found that, at the aggregate, teacher turnover rates were similar to these other professions, and concluded that differences that did exist were largely driven by the earlier retirements of teachers at the upper end of the age spectrum, rather than by the turnover of younger teachers.

One factor that may make turnover in teaching difficult to compare to other professions is the fact that teaching is one of only a few heavily female dominated professions. This distinction is important because research on the wider labor market finds differences in turnover and its relationship to how men and women negotiate career decisions around issues of marriage and childrearing. Specifically, women are far more likely to leave jobs for non-employment family options than men (Royalty, 1998). Thus, exit rates in teaching may be somewhat exacerbated by the decisions of substantial portions of women teachers who leave their jobs for family reasons (Stinebrickner, 2002; Wayne, 2000). One study found that after the first four years in the profession, female teachers are considerably more likely than male teachers to exit teaching for non-employment reasons (Stinebrickner, 2001). Although these kinds of exits do contribute to the overall turnover rates in teaching, many of them are temporary, and evidence also suggests that 25-35 percent of the pool of teachers from which districts hire is comprised of teachers re-entering the profession (Ingersoll & Perda, 2010; Wayne, 2000).

One issue that is not clear in the literature is whether the continued female dominance in teaching is related to its perceived family-friendly structure. In other words, do more women continue to select into teaching because of its reputation as a career that can be more easily balanced with family responsibilities than other kinds of careers? If this is the case, policy prescriptions designed to stem attrition may need to account for the non-employment reasons many women leave teaching, as well as consider how to alter recruitment efforts to capitalize on changing social norms around gender and child-rearing. For example, one survey that asked teachers about non-monetary incentives that could induce them to remain in the same district found that more than half of teachers would stay in a district if employer-provided daycare were a benefit (Kelly, Tejada-Delgado, & Slate, 2010).

The Problem with Attrition

Whether or not teacher attrition is comparably higher than other professions, for men or women, there are nevertheless policy concerns about the distribution of new teacher attrition amongst different kinds of schools. In particular, schools that serve comparatively high proportions of poor and minority students are also more likely to experience higher turnover, and have greater difficulty filling vacancies than are other schools, even within the same school district (Goldhaber, Gross, & Player, 2010; Ingersoll & Perda, 2010; Lankford, Loeb, & Wyckoff, 2002). For example, in Chicago, although the overall annual teacher retention rate is 80 percent, every five years a majority of schools in that city loose over half of their teaching staff. After four years, schools with the highest attrition rates retain less than 30 percent of their teachers (Allensworth, Ponisciak, & Mazzeo, 2009). These kinds of findings illustrate the magnitude of the problem of new teacher turnover, concentrated in urban schools. At the same time, high-poverty schools also present greater professional challenges for teachers, and there is also significant evidence to suggest that the rapid turnover of new teachers in these schools is linked to teacher dissatisfaction with low status, poor working conditions and low salaries (Borman & Dowling, 2008; Darling-Hammond, 2004; Loeb & Darling-Hammond, 2005).

Teacher shortages are also not evenly distributed across geographic areas. For example, because of a combination of increases in population growth, combined with the added burden of efforts to maintain a sweeping statewide class-size reduction policy, shortages have been more problematic in California, which is home to about eight percent of the nation's school children, than in many other states (Podgursky, 2006). Other Western states and some in the South, Southeast, and Northeast have also experienced difficulties filling vacancies, while states in the Midwest have comparatively less difficulty (Murphy, DeArmond, & Guin, 2003).

Furthermore, because of evidence that new teachers, particularly those in their first three years in the classroom, do not perform as well as their more experienced peers (Kane, Rockoff, & Staiger, 2006; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), there is concern that schools that are staffed primary by novice teachers and face constant turnover also systematically provide a lower standard of educational opportunity than schools with more stable and experienced teaching staffs (Lankford et al., 2002). High levels of churn undermine a school's organizational capacity to implement either coherent curricular reforms, or develop collaborative relationships amongst staff (Guin, 2004). Chronic turnover also drains instructional expertise from schools, and deprives them of appropriate mentors for novice teachers (Darling-Hammond, 2004). Recent research findings also suggest that there is a significant and negative impact of constant teacher churn within schools on student achievement (Ronfeldt, Loeb, & Wyckoff, 2012).

Other concerns about chronic turnover in some schools and districts are based on the idea that transaction costs associated with the constant effort to replace new teachers strains the resources of districts that are already financially constrained. Although it is difficult to assess the exact costs of turnover within each district, and costs differ between districts, one study places the cost per novice teacher replacement at between a low of

$4,366 for a small, rural district, and a high of $17,872 for a large urban district (Barnes, Crowe, & Schaefer, 2007). Another study estimates an average new teacher replacement cost of $9,061 per teacher (Milanowski & Odden, 2007). Reducing turnover of new teachers, therefore, is also associated with potentially substantial cost savings for public schools.

Conversely, some have argued that teacher turnover can be beneficial, if the least effective teachers are the ones leaving the profession. To that end, there is mixed evidence about retention rates based on various measures of teacher quality. For example, some studies find that teachers with stronger academic backgrounds, as measured by quality of their undergraduate institution, are more likely to leave teaching (Boyd, Lankford, Loeb, & Wyckoff, 2005; Hoxby & Leigh, 2004), while other studies have found similar results for teachers who had high ACT or SAT scores (Podgursky, Monroe, & Watson, 2004; Stinebrickner, 2001).

College board scores and academic competitiveness are, however, arguably weak indicators of pedagogical talent, and findings are mixed regarding whether or not the most effective teachers, as measured by student achievement, are more likely to remain in teaching. While some studies have found that the most effective teachers are the most likely to stay in the classroom (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Hanushek, Kain, O'Brien, & Rivkin, 2005), others have found greater subtlety in variation in attrition across the effectiveness distribution. For example, several studies find that the greatest teacher mobility rates are found at both the lowest and highest ends of the effectiveness distribution, as measured by teacher value-added contributions to student achievement on standardized tests (Goldhaber et al., 2010; West & Chingos, 2009). Furthermore, the Goldhaber et al. study finds that while the least effective teachers tend to either leave teaching, or to transfer to other schools in the same district, the most effective teachers are more likely to leave the profession all together.

Shortages in Teaching Subfields

In addition to the chronic shortages in teaching attributed to turnover in more recent decades, shortages of qualified math and science teachers have been documented as far back as the 1940s and 1950s, when the percent of mathematics teachers without a major or minor in math ranged from 35-60 percent in different states. Shortages in the sciences were even worse, with fewer than 20 percent of chemistry teachers having majored in chemistry, and fewer than 10 percent of physics teachers having majored in physics, in some states (Levin, 1985). Furthermore, these numbers have held steady for at least 60 years, such that more recent figures suggest that nationally, 30 percent of secondary mathematics teachers did not major or minor in either mathematics or related fields such as engineering (Ingersoll, 1999). Moreover, although the supply of new math and science teachers in the past decade has kept pace with both teacher retirements and increases in demand created by increases in curricular requirements in mathematics and science, there is some evidence that the greatest source of shortages comes from the fact that more teachers of these subjects leave either their school or the profession pre-retirement than do teachers of other subjects (Ingersoll & Perda, 2010; Murnane & Olsen, 1990). For example, in one study, after three years of teaching, 30 percent of those with majors in mathematics, engineering or the natural sciences had left teaching, compared to 14 percent of those who had majored in education (Henke, Zahn, & Carroll, 2001).

In addition to pre-retirement exits in mathematics and science subject areas, other subfields face shortages based in the supply pipeline of qualified teachers. For example, the shortage of qualified Special Education and English as a Second Language (ESL) teachers are both subject areas in which there are simply not enough qualified teachers to meet demand (Billingsley & McLeskey, 2004; Diaz-Rico & Smith, 1994; Murphy et al., 2003). The shortages in these areas, however, are potentially worsened by existing policy. For example, federal NCLB legislation requires that, in order to be considered "highly qualified", secondary special education teachers must be certified in both special education and the content area of expertise, potentially creating excessive barriers to entry to the teaching of this sub-field (Billingsley & McLeskey, 2004). Additionally, many schools of education have failed to keep pace with rapidly changing student demographics, and have been slow to incorporate the development of bilingual/ESL teachers into their curricula (Diaz-Rico & Smith, 1994), which has meant that not enough certified teachers are available to teach rapidly growing populations of language minority students.

Theoretical Framework

Opportunity Costs and Declines in Quality

There are two kinds of teacher supply problems. There are supply shortages caused by lack of adequately prepared teachers, and shortages created by excess attrition. Both kinds of supply problems impact the ability of schools to staff classrooms with qualified, effective teachers. The first kind of supply problem can be caused either by policies that create barriers to entry, like those that impact the training and recruitment of special education and bilingual teachers, or by a failure of qualified people to choose to enter the profession. While the supply shortages created by policy barriers are relatively easier to both identify and correct, shortages caused by self-selection, as well as those created by higher rates of turnover amongst math and science teachers, can be understood through economic models of how both workers and employers create appropriate job matches across the labor market.

Specifically, economic theory suggests that workers self-select into occupations for which they are likely to receive the highest possible remuneration. The Roy Model (Roy, 1951) describes how, if the skills required to do two jobs are equal, but the potential wages of one are more compressed than the other, then the highest quality workers will select into the occupation with the greatest wage dispersion, even when average wages are the same. This is because the most productive workers will assume that their own superior ability will place them in the top end of the salary distribution. One implication of this model is that there will be an unequal distribution of quality between the two professions, which will lead to decreases in productivity in the profession with more wage compression. The implication of this model for teaching is that highly able graduates will have an incentive to choose non-teaching options, if they can earn substantially higher wages in those fields.

For mathematics and science teachers, who can command greater financial rewards in the private sector than other teachers (Bradely & Loadman, 2005), the empirical evidence supports the theoretical model. Specifically, the single salary schedule, which offers considerably more compressed wages than are found in the private sector market for professionals generally (Vigdor, 2008), provides significantly lower salaries than are frequently available to college graduates with degrees in math and science related subject areas. Thus, these graduates face greater opportunity costs to teaching that do teachers of other subjects. Research estimates indicate that beginning salary differentials for graduates with technical majors are almost $2,000 between teaching and non-teaching jobs (Goldhaber & Liu, 2005). Additional findings also suggest that secondary teachers face an almost $3,400 greater annual opportunity cost to teaching than do elementary school teachers (Goldhaber & Player, 2005). Although these numbers are not large, they may be consequential. One study found that each $1,000 predicted starting wage increase for teaching is associated with between three and four percent growth in interest in the field reported by college students majoring in technical subjects (Milanowski, 2003). Based on these figures, an increase in starting salary of $4,000 could increase the supply of potential shortage-field teachers by as much as 16 percent. With 485,800 bachelor's degrees granted in science and engineering fields in 2007[1], such a raise in starting salaries could potentially increase the national math and science teacher supply by 77,728. A starting salary raise increase of $5,000 would serve

[1] Data found at NSF: National Center for Science and Engineering Statistics http://www.nsf.gov/statistics/seind10/c2/c2h.htm

to nearly satisfy President Obama's call for 100,000 new math and science teachers (Obama, 2011).

Evidence of comparatively low starting salaries only demonstrates opportunity cost decisions faced by early career shortage-field teachers and non-teachers. However, as discussed above, teacher retention is also a pressing issue that disproportionately impacts shortage-field subjects, where the average stay of science teachers has been found to be between two to seven years shorter than the average elementary school teaching career (Murnane & Olsen, 1989, 1990). Moreover, differential exit rates imply that opportunity costs continue into mid-career, where the salary gap between math and science related teaching and non-teaching careers grows to as much as 25 percent, favoring non-teaching careers. This is compared to the differential for people with education degrees, which is roughly 17 percent (Goldhaber, DeArmond, Liu, & Player, 2008). In addition, after a career interruption, those teachers who chose to re-enter teaching are those who have fewer opportunities for wage increases outside of teaching (Beaudin, 1993). These studies support the notion that the lock-step salary schedule becomes less attractive as teachers gain experience (Vigdor, 2008), and may help explain the high turn-over of shortage field teachers, who start out their careers on closer footing with their non-teaching peers, but who notice themselves falling behind as they move into mid-career. Combined, these studies serve to illustrate that shortage-field teachers face additional opportunity costs to stay in teaching, compared to other teachers.

One caveat to these findings is that potential salary losses and attrition are not uniform across different types of labor markets. This is because some local labor markets provide greater comparative wages, depending on the demands of the local non-teaching economy. To that end, Rumberger (1987) found that districts located in areas with high concentrations of engineers, indicating high levels of demand for those kinds of credentials, reported greater shortages of mathematics and science teachers, as well as greater levels of teacher turnover within those markets.

Reductions in Teacher Quality

As predicted by the Roy Model, the opportunity costs created by wage compression in teaching have also led to documented declines in teacher quality. For example, aggregate levels of decline in the academic preparation and average skill levels between current teachers and those who taught in the past have been observed. In particular, historical shifts in the labor market for women have precipitated changes in the measureable characteristics of the teacher corps over time. Specifically, the impact of salary schedule constrained wages was less relevant when women had fewer occupational opportunities. In the 1940s, teacher salaries were relatively higher than those for other occupations open to college educated women. However, as those women moved into other careers, relative teacher salaries dropped considerably, making teaching steadily less financially attractive to the kinds of women who had filled those jobs in the past (Hanushek, Rivkin, Rothstein, & Podgursky, 2004). The effect of this shift of women into significantly more non-teaching occupations has been that for the highest performing female graduates, the opportunity cost to teaching has also increased. Although teaching is still dominated by women, the academic ability of those in the profession has shifted downward such that, even though the mean academic achievement level of teachers has fallen only slightly since mid-century, the number of high achieving women who enter the profession has dropped dramatically (Corcoran, Evans, & Schwab, 2002; Hoxby & Leigh, 2004).

In addition to changes in women's employment, the decline in teacher quality has also been linked to technological progress, which has served to increase the specialized skills of workers in non-teaching fields, but which has not affected the productivity of skilled teachers. This phenomenon occurs because the skills required for teaching have remained relatively unchanged since the inception of public schooling (Tyack & Cuban, 1995). The result of the stasis of skills needed within the teaching profession has actually been the decline of the skill and knowledge required of teachers, relative to other non-teaching occupations (Lakdawalla, 2001). Thus, although teaching was once one of the most intellectually challenging occupations open to smart young workers, the advances made in almost all other professional category jobs has dramatically shifted the dynamic.

Overall, the documented evidence of opportunity costs, combined with evidence of decline in the quality of teachers suggests that the Roy Model predictions about the consequences of compressed wages are applicable for understanding the dynamics of teacher retention difficulties, particularly as related to teachers of mathematics and science.

Imperfect Information and the Recruitment of Teachers

Although the Roy Model may be effective for illuminating issues of teacher self-selection into the teaching profession, it is limited in its usefulness for understanding how schools and districts select teachers from within the existing pool of applicants. Because the model does not address principal agent information asymmetries in the hiring process, it may be helpful to temper teacher labor market understanding with further theoretical models of the role of imperfect quality information on market dynamics. Specifically, in Akerlof's *Lemons* model, uncertainty about the quality of the good being offered serves to drive high quality goods from the market. This occurs because buyers will pay lower prices if they have a reasonably high enough probability of actually purchasing a lemon (Akerlof, 1970). This model provides a useful lens for understanding the market for teachers because the attributes of a quality teacher are hard to both define and measure, particularly before a teacher actually starts teaching.

The problem of hiring high-quality teachers is two-fold. The first problem is that the goal of schooling has been both vague and controversial. Historically, competing ideas regarding the purpose of education have made it difficult to either identify or evaluate good teachers. Social and political debate about the purpose of public schooling itself has created an environment in which there is a lack of clarity around determining what goals a good school and a good teacher should pursue.

For example, for many decades, beginning with Horace Mann's common schools and through the development of Carnegie Units and the comprehensive high school, the policy debate over public education has swung between the two opposing ideals of democracy and capitalism (Carnoy & Levin, 1985; Tyack & Cuban, 1995). Although many authors have addressed these societal tensions and their role in framing public education policy, Labaree (Labaree, 1997) goes a step further by outlining the history of education reform as an evolutionary process that has moved from a tension between the economy and society, into a single mission of individual consumerism, that is still in the process of emerging today. In this conception, Labaree describes deep tensions between the conviction that the primary purpose of schooling is to raise good citizens and enhance social cohesion through the transmission of common knowledge and values, and the belief that the primary purpose of schooling is to create good workers, prepared to slot into the various strata of the labor market. Historically, the socio-political environment swung between these competing ideals, and ultimately gave rise to the development of a massive public institution that nonetheless had ambiguous aims. This system lacked policy coherence, and for decades the role of teachers was caught within it, in such a way that it was necessarily impossible to define good teaching. Essentially, although is seems obvious that good teachers must be able to facilitate learning, the problem of determining quality in carrying out this end is highlighted if the question is reframed against the backdrop of conflict over what students should be learning: is the teacher who is best suited to molding good citizens, the same as the teacher best suited to shaping agile workers? Which end should the individual teacher pursue? Thus because historically the desired product of schooling was both debated and poorly articulated the role of teacher was also unclear. This lack of clarity of goals gave rise to a "loosely coupled" system that manifested as a lack of connection between teacher inputs and outputs (Weick, 1976).

Given this policy backdrop, hiring quality teachers has been complicated, both by the lack of clarity about the role of teachers, and the fact that the kind of teachers that different schools and districts wish to hire likely differed significantly as well. This complexity has meant that, because it is unclear what teachers should be able to do, the signals of quality have been weak. Therefore, the second problem in the hiring of teachers is that identifying a quality teacher at the point of hire is difficult if the ultimate role of teaching is up for debate.

As previously discussed, the traditional single salary schedule rewards teachers for the acquisition of credentials, advanced degrees, and experience. These are also the criteria that have generally been used to indicate qualified applicants in the teacher labor pool. Because these indicators are not tied to output, they may also be well suited to systems of identification and reward within an institution in which the productive aim is ambiguous.

However, as described by Labaree, the competing goals in education have begun more recently to give way to what he calls the single *social mobility* goal. This goal conceptualizes education as an individual commodity, rather than a public good. Proponents of this view believe that education is a private good, the purpose of which is to obtain credentials that will make individuals competitive in the labor market. They also believe that the allotment of education should be handled in a meritocratic manner. The purpose of education is to allow individuals to obtain comparative advantage, but that advantage should be allotted on merit, rather than on such inequitable criteria as social position, or family. Therefore, by combining individual labor market ends with a system of merit-based delivery, the social mobility goal can also be seen as an emerging hybrid of the two earlier policy goals.

Moreover, the increasing dominance of the social mobility goal can be seen as the impetus behind many modern educational developments and reforms, including judicial decrees around desegregation and the development of educational adequacy legislation, as well as the No Child Left Behind act, all of which can be characterized as efforts to make educational competition more fair. One important outgrowth of these developments has been the development of data systems designed to measure whether or not educational enterprises are meeting the aims with which they have been charged. Namely, as policy makers and reformers have coalesced around the social mobility conception of education, they have also demanded that the existing "loosely coupled" education system be made accountable for ensuring equitable educational access through the development of measureable outcomes (Hannaway & Mittleman, 2011). For the most part, the indicators that have been developed are standardized tests of individual student achievement. While these evaluations have been used to track student progress, their increasing prominence has also highlighted gross educational inequities that serve to undermine the mobility ideal.

Thus, it is through the very recent and growing availability and analysis of data, that the problem of identifying and hiring quality teachers has become apparent. On the one hand, individual teacher ability has been found to be profoundly related to student achievement (Rockoff, 2004), with differences in teacher quality across schools linked to as much as 7.5 percent of the variation in student achievement (Hanushek, Kain, & Rivkin, 1998). On the other hand, most of the teaching qualifications that have traditionally been used to indicate quality, and that can be assessed from a resume, such as undergraduate degree, certification, and advanced teaching degrees have been found to have either weak or no relationships to student performance (Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008).

At the same time, the teacher-level indicators that have shown the strongest positive relationship to student achievement, such as teaching experience (Hanushek et al., 2005) and National Board Certification (Clotfelter, Ladd, & Vigdor, 2011), can only be measured after a teacher has been teaching for some time (i.e. teachers are ineligible to apply for National Board Certification until they have a minimum of 3 years of experience[2]). Value-added models, which measure the effectiveness of individual teachers on student achievement scores, also require teachers to have some experience with students in classrooms to be utilized. These factors mean that schools and districts must hire teachers without indicators of quality that are reliable at the time hiring decisions are made.

Meanwhile, in the absence of observable quality at the time of hire, employers will nonetheless rely on signals of quality, such as undergraduate institution or major, to proxy for actual ability (MacLeod & Urquiola, 2009; Spence, 1973). The lack of reliable indicators of teacher effectiveness and the reliance on weak signals means that there is quality uncertainty in the market for new teachers, especially those that are inexperienced. The *Lemons* model predicts that in this kind of scenario, the salary offered to new teachers will be depressed, and not aligned with what highly skilled teachers are actually worth. Because human resources personnel are likely unable to accurately predict from the hiring process which candidates will either be great teachers, or which ones will leave the profession prematurely, low starting salaries may serve to save districts from investing resources in teachers who may have high probabilities of revealing themselves to be lemons. In this framework, the implication of this theoretical model is that even without being resource constrained, districts might still offer low starting salaries if teacher quality uncertainty is high.

[2] http://www.nbpts.org/become_a_candidate/eligibility_policies

However, the model also predicts that when a mid-level salary is actually offered, it will serve to further dilute the quality of teachers in the market by attracting low quality teachers, who can get a higher salary then their actual skills should warrant. All of this suggests a potential reason for the continuing wage gap between teachers and non-teachers with equal education, which is that private sector employers may offer higher salaries and greater wage dispersion if they are hiring for industries in which college majors and degrees provide greater certainty of a match between observable qualifications and productivity, than can be found in teaching.

Evidence on Quality Teacher Hiring

Because of the tension between the difficulty in identifying quality teachers from the point of hire and the need to simultaneously raise the quality of teachers, in more recent years, policy makers have made efforts to improve the pre-hire information about teachers that is available to schools and districts. For example, many states have adopted to use of competency tests to be used as screening mechanisms for prospective teachers. By setting an objective floor, these tests are designed to prevent the least qualified from entering the classroom. However, evidence indicates that the predictive value of these tests differs across racial subgroups of both teachers and students. Specifically, Goldhaber and Hansen (2010) found that Black students who were assigned a Black teacher who scored in the low end of the distribution of scores on teacher competency exams performed as well as Black students assigned a high-performing White teacher. These findings highlight that hiring quality teachers may also be difficult, in part, because measures designed to provide objective signals of quality are nevertheless subject to context, and that the most appropriate matches may need to be made using a variety of both objective and subjective indicators.

To that end, several other studies examine the predictive value of subjective pre-hire information about teachers, including teacher self-report surveys about such issues as ability, content knowledge, and personality, and rigorous screening interviews used by an alternative certification program (Rockoff, Jacob, Kane, & Staiger, 2008; Rockoff & Speroni, 2011). Both of these studies found some positive predictive value in terms of identifying teachers who had positive impacts on student achievement. However, both also concluded that combinations of traits were more indicative of quality than any single identifiable elements alone. In addition, while both of these studies indicate that it is possible to develop more rigorous systems for screening and hiring effective teachers, it is a process that is complex, and will likely require that school districts significantly increase the resources they dedicate to the hiring process. Although the process is still in its nascency, together, these studies represent a glimmer of how teacher hiring may be able to move forward to reduce information uncertainty, and provide more justification to allow districts to increase salaries and attract more qualified teachers.

Overall, the Labaree and *Lemons* frameworks compliment each other. One highlights the consequences for both prices and quality when observable indicators are unreliable, while the other illuminates the socio-historical policy context that has served to create uncertainty in the market for teachers. At the same time, the second also provides a theoretical lens for understanding the current evolution from a system dominated by uncertainty to one in which reforms are designed to increase what is known about teachers by tightening the relationship between their input and the student achievement output. Moreover, the empirical evidence serves to support the notion that established indicators of teacher quality are inadequate to support the goals of the emerging theoretical construct. Research findings also indicate that although new models of more tightly coupled teacher signaling mechanisms are possible, there is much that remains unknown about how to develop screening tools that efficiently identify those teachers who will be best situated to carry out the emergent schooling goals around equity and individual student achievement.

The Policy Prescription and Evidence on Economic Incentives

As discussed above, economic theory suggests that one way to attract more and better qualified secondary math and science teachers is to offer them higher salaries than their counterparts in the elementary grades, or even in other secondary subjects. The Roy Model implies that salary differentials would serve to reduce the greater opportunity costs to teaching faced by shortage-field teachers. These differentials are also hypothesized to increase teacher quality because equalizing wages between teaching and non-teaching would mean that income rewards were no longer an effective sorting mechanism. Salary differentials could also be effective for changing recruitment of shortage-field teachers if teaching provides not just lower salaries than other professions open to graduates with training in shortage-field subjects, but if it is also perceived as offering worse working conditions, such as less autonomy, fewer opportunities for advancement, inadequate supplies and materials, or greater potential exposure to violence or germs, for example. In this case, higher salaries could provide what is known as a compensating differential, which is the additional amount of incentive necessary to induce workers to take a job that provides worse working conditions to other jobs available requiring the same background or skills (Smith, 1979).

Despite the theory, there is little research evidence on whether or not financial incentives for shortage field teachers are actually effective, and the evidence that is available provides a mixed picture. For example, controlling for other factors such as school demographics and conditions, similar schools that pay more are able to attract both new and experienced teachers with better credentials (Figlio, 2002). In addition, a program in North Carolina that offered an annual $1,800 bonus to certified shortage-field teachers who worked in high-needs schools reduced mean turnover rates by 17 percent, with the strongest impacts on math teachers as well as those with relatively more experience (Clotfelter, Glennie, Ladd, & Vigdor, 2008).

By contrast, a signing bonus program in Massachusetts that paid new recruits $20,000 over three years to receive training and teach in high-needs schools was found to be ineffective for stemming attrition when large numbers of recruits nevertheless left teaching in the first few years of the program (Fowler, 2003; Liu, Johnson, & Peske, 2004). What these findings suggest is that although pay incentives may be ineffective for addressing initial principle agent information problems at the point of hire, they may be somewhat more effective for increasing motivation and stemming attrition from the field for those teachers who have already demonstrated commitment to the classroom beyond the entry level years.

The scarce evidence on differentiated pay in the public school arena is mostly due to lack of implementation. However, differentiated pay is used routinely in other educational sectors. Most notably, it is common in higher education for engineering, medical, and law faculty to earn considerably more than their counterparts in more traditional academic fields (Tuckman & Tuckman, 1976). However, although salary variation in higher education has been defended as a reflection of non-academic market demand for skills, there is also evidence that gender bias plays an independent role in faculty salary setting at the post-secondary level. Specifically, Bellas (1997) finds that over time, there is an inverse relationship between average disciplinary salaries and the proportion of women faculty in a given subject area, suggesting that salaries become depressed by the entrance of women into given disciplines. Additionally, the same study finds that greater wage dispersion undermines worker satisfaction, productivity, and collaboration amongst faculty. Although these findings are for faculty within the same department, rather than across departments, they provide some evidence of potential unintended consequences of differentiated pay.

Differentiated pay is also frequently used in private schools, which pay premiums for math, science and special education teachers (Ballou & Podgursky, 2001). However, despite these practices, there is little empirical evidence to indicate whether or not the policies either reduce retention, or improve the quality of the shortage-field educators in these settings. Moreover, the most relevant findings suggest that, despite the greater flexibility in compensation available to administrators in private schools, teacher attrition is nevertheless higher amongst private school teachers, as compared to their public school counterparts (Guarino, Santibanez, & Daley, 2006; Ingersoll, 2001).

Overall, the argument for differentiated pay in public schools is that implementing it will both increase quality and stem attrition of shortage field teachers. However, despite the fact that differentiated pay is utilized in other educational settings, these outcomes have not actually been studied in those settings, so there is scarce comparative evidence of their efficacy. Moreover, the impact on outcomes that have been studied, particularly in higher education, suggest potential unforeseen consequences of differentiated pay policies, such as increased gender bias in pay, and decreases in staff cohesion.

That said, in the absence of flexibility in teacher compensation schemes, there is also evidence that public schools will provide non-monetary inducements in order to retain high quality teachers. In particular, one study found evidence that elementary school principals provided more favorable classroom conditions, including those with more high achieving math students, lower percentages of IEP and ELL students, and ratios favoring female students, to teachers with higher competency exam scores, and those who had achieved higher value-added with prior classes (Player, 2010). These findings are suggestive because they indicate that not only could differentiated pay arguably increase equitable distribution of teachers by removing the school incentive to reward the best teachers with more favorable classes, increases in salary in areas where there are shortages could possibly help lower-performing students obtain greater access to the most qualified teachers.

Finally, there is some evidence that salary enhancements may need to be considered as both corrections to opportunity cost mechanisms, and as compensating differentials, at least in terms of keeping teachers in the lowest performing schools, where teachers have been found to leave as much for reasons of salary as for poor working conditions (Darling-Hammond, 2004). That said, in a further investigation of turn-over in low performing schools, women teachers were found to be unresponsive to salary bonuses when the decision was to remain in schools where working conditions were poor (Hanushek, Kain, & Rivkin, 2004).

Other Financial Incentives in Teaching

Currently, many policy and research discussions center around the idea of changing teacher compensation to make teaching more closely tied to student achievement. The push to incentivize teachers to raise student test scores is also part of the *social mobility* shift in educational goals that seeks to move away from a loosely coupled school system to one in which inputs and productive output are more closely tied together. While shortage-field incentives are designed to create shifts in the labor market that will ultimately lead to more high quality teachers entering and remaining in the classroom, another class of policy incentives around teachers' labor are also currently in vogue. Namely, a variety of policy experiments in recent years have offered teachers financial incentives to increase student achievement on standardized tests. The theory behind these kinds of incentives is that teachers will work harder to improve student performance if they know that doing so will earn them a personal reward. The assumption behind these kinds of incentives is that student performance has been weak in the existing institutional structure because teacher effort was loosely coupled from student output. Teacher effort is seen as one of the main determinants of student performance, thus bringing these elements closer together will necessitate both more teacher effort, and higher student performance. Although both shortage-field incentives and performance incentives are rooted in the same over-arching theoretical model, they aim to address different aspects of teachers' work, and have also been found to have different outcomes.

Namely, although financial incentives have shown some success in helping to keep shortage-field teachers in the classroom, the evidence that financial rewards are successful at raising quality, or increasing motivation for teachers once they are in the classroom is mixed. For example, Lavy (2002) finds that schools that competed for and won financial rewards did experience some improved outcomes for students, such as reduced drop-out rates, and increases in student acquired credit hours, compared to similar schools that were not eligible to compete. On the other hand, a similar randomized group incentive plan administered to schools in New York City did not result in any significant differences in either student achievement, or most teacher outcomes (Fryer, 2011; Goodman & Turner, 2009). In addition, even when teachers were paid large performance-based bonuses, measurable differences in student achievement gains on standardized tests were not found (Springer et al., 2010).

Overall, the research that has been done on pay enhancements for teachers provides a muddy picture of how teachers respond to pay-based incentive policies. However, those that provide positive impacts seem to do so on issues related to supply and retention, while those that are related to improving quality seem to be more likely to have null findings. Nevertheless, the dearth of evidence and the mixed findings from what does exist, particularly about differentiated pay, suggests a need for more evidence about how pay enhancement policies impact teacher labor markets.

Chapter 2

Research Focus

My research focuses on three aspects related to the role of shortage-field incentives in recruitment and retention of teachers with specialized skills.

1. I begin by examining whether the presence of incentives in a district impacts teacher attrition. Are shortage-field teachers in districts that offer shortage-field incentives more likely to be retained after one year than similar teachers in districts that did not offer shortage-field incentive?
2. Second, I examine whether the presence of shortage-field incentives impacts recruitment of shortage-field teachers in those districts. Specifically, I examine whether districts that offer incentives are able to recruit shortage-field teachers with stronger academic backgrounds than districts without incentives.
3. My next research question asks whether changes to shortage-field incentive policies over time lead to changes in shortage-field teacher recruitment or retention outcomes. Districts that change their incentive policy by either adopting a new incentive policy or abandoning an old one are expected to experience accompanying changes in teacher outcomes. In this set of analyses I examine whether or not these changes actually occur.

Data and Methods

Data

The Schools and Staffing Survey (SASS) for 1999-2000, 2003-2004, and 2007-2008 provide nationally representative survey data about schools and teachers. These three combined waves of the SASS contain data from 106,930 public school teachers in 6,540 public school districts, from all 50 states. Included in these surveys is information about districts that do and do not offer shortage field incentives, as well as information on whether surveyed teachers were retained one year after the initial survey was collected. The data also include information about teachers' backgrounds. For a full description of all variables used in these analyses, please see Appendix Table 1.

Methods

A comparison of means for shortage-field teachers in public school districts pooled from the three survey waves show significant differences between teachers in districts that did and did not offer shortage-field incentives (Table 1)[3]. These results suggest that districts that offer shortage-field incentives are more often located in rural and urban areas, as well as in the South and West. These districts also have higher percentages of Black and Hispanic students, as well as higher percentages of students eligible for free or reduced price lunch. Although more rural districts offer incentives, they impact more urban teachers. Specifically, the mean number of teachers sampled in each rural district is 9, compared to 27 in each urban district. Thus in these surveys, there

---

[3] Although there are only a total of 6,540 unique districts in the 3 waves of SASS, these data are for more than 12, 000 districts. This discrepancy reflects that fact that many districts were surveyed more than once, and applied different shortage-field incentive policies in the different years they were surveyed. Thus, districts were counted based on their incentive and survey year status, so some districts were counted more than once. Demographics reflect the mean for the district year.

are almost 11,000 teachers teaching in urban districts that offer shortage-field incentives, compared to 7,300 teaching in rural districts that provide the benefit. Nonetheless, these descriptive findings provide *prima facie* evidence of non-random distribution of shortage-field incentives among districts.

Table 1:
Characteristics of Districts that Do and Do Not Offer Shortage-field Incentives

Variable Name (at district level)	Shortage-field Incentives (n=2,030)	No Shortage-field Incentives (n=10,460)
District Mean % Black	0.16	0.10
District Mean % Hispanic	0.18	0.08
District Mean % in lunch program	0.47	0.38
Urban	0.20	0.10
Suburban	0.40	0.41
Rural	0.40	0.48
South	0.44	0.29
Midwest	0.21	0.33
Northeast	0.09	0.18
West	0.25	0.20

**T-statistic mean differences significant at the p<.001 level. Suburban not significant.*

Causal Framework: Research Questions 1 & 2

The fact that certain kinds of districts are more likely to provide shortage field incentives introduces selection bias into the data, jeopardizing the ability to make causal inferences about any potential incentive effects. A straight comparison of outcomes between districts that do and do not provide shortage-field incentives is likely to conflate the effect of incentives themselves with other district characteristics that have known relationships to teacher recruitment and retention outcomes. For example, high-poverty schools are known to be associated with higher teacher turnover, therefore, given that districts with more high-poverty students are also more likely to offer shortage-field incentives, it is difficult to parse the direction of causality for the impact of those incentives on teacher retention. A straight comparison of teacher retention between districts that do and do not offer shortage-field incentives would likely capture only the difference in outcomes between high and low poverty districts. In order to identify the impact of incentives themselves, rather than the effect of other district characteristics that are associated with the outcome, I utilize an analysis strategy that creates a comparison group that is similar enough to the treatment group that received the incentives to indicate what would have happened, had the shortage-field incentives not been implemented in the districts that utilized them.

In order to address the selection bias between districts that choose to offer shortage-field incentives and those that do not, I utilize a two-stage least squares instrumental variable (IV) approach. In this method, I utilize a variable that is related to the policy of interest—the offering of shortage-field incentives—but which related to the outcomes only through the policy of interest, in order to create a natural experiment in which assignment to treatment or control conditions is driven by a source of variation that is external to the model (Angrist, Imbens, & Rubin, 1993; Angrist & Pischke, 2009; Gelman & Hill, 2007). Because assignment to treatment is exogenous this method allows me to create a plausible counterfactual control group, permitting valid causal inferences about whether shortage-field incentives increase retention and increase the quality of teachers recruited.

In this case, I exploit the exogenous variation provided by teacher union representation. Specifically the two major national teachers unions, the American Federation of Teachers (AFT) and the National Education Association (NEA), have differing official platforms on the use of differentiated pay for shortage-field teachers. While both organizations state that they favor the maintenance of a strong single-salary schedule as a measure of base pay, the NEA explicitly does not support additional pay for teachers in hard-to-staff subjects. Their justification for opposing subject-specific pay enhancement is that doing so "sends the signal that we value certain subjects over others and destroys internal equity" (Roekel, 2008). The AFT, on the other hand, supports placing teachers in shortage-fields, such as mathematics, science, and special education, higher on the salary schedule when they begin teaching (Goldhaber, 2009; *Where We Stand: Teacher Quality*, 2003).

Given the divergent views on shortage-field incentives of the nation's two largest teachers' unions, I hypothesize that AFT or NEA union affiliation is likely significantly related to whether or not a local district's teaching contract provides shortage-field incentives to teachers in shortage-fields. At the same time, I hypothesize that the affiliation of a district's teachers' union is unrelated to the recruitment or retention of teachers in shortage-fields within districts because teachers are unlikely to choose to remain in teaching because of the national affiliation of their labor union. Union affiliation itself is also unlikely to be related to the quality of teachers recruited into districts, because teachers are unlikely to choose the work in a given district because of its union affiliation. Moreover, union affiliation is also unlikely to cause districts to systematically hire more qualified candidates than other districts, represented by other unions.

In addition, whether a given districts' teachers are represented by the NEA or the AFT is, in most cases, a historical artifact related to the different development, growth and competition of the two unions for members during and leading up to the period when collective bargaining laws were first enacted, rather than to specific district characteristics (Murphy, 1990). The divergent development of the two unions makes union affiliation a potentially robust variable that provides a strong source of exogenous variation between districts. Furthermore, while specific union representation will likely strongly influence the type of pay contracts offered to shortage field teachers within given districts, union affiliation alone is unlikely to directly effect the retention of those teachers in teaching, or the recruitment of quality teachers into districts.

Analysis Sample: Research Questions 1 & 2

The Schools and Staffing Survey (SASS) does not contain information on local union affiliation. Therefore, in order to create the union instrument variable, I utilized publically available data on teachers unions and manually assigned NEA or AFT to the districts in SASS, which could only be identified by linking them to the Common Core of Data. I began this process by utilizing lists of state and local union affiliates available on the state level websites of the NEA and the AFT. However, because these lists were incomplete, I followed this process by conducting searches of local news reports about teachers' unions, as well as searches of legal documents involving local teachers' unions. Other sources of evidence included documents such as copies of negotiated teachers' contracts posted by policy research organizations, and state labor relations boards. I was able to find positive information on union affiliation for 67 percent of the districts in the analysis sample. For the most part, districts where local union affiliation was not found are smaller suburban and rural districts. In addition, a small minority of districts are represented by local unions that are unaffiliated with either of the two national teachers' unions.

In order to focus the analysis on the kind of wage agreements that are negotiated by the different unions, I limited the sample to states that require full collective bargaining for public sector unions. I did not include five states that have no collective bargaining rights (GA, NC, SC, TX, VA). I also did not include states with weaker "meet-and-confer" negotiation rights (AL, AR, AZ, CO, KY, LA, MO, MS, UT, WV, WY)[4]. In addition, in order to ensure that within-state comparisons could be made, I further limited the sample to states in which both the AFT and the NEA have local affiliates. The AFT is a much smaller union than the NEA, and does not have any K-12 local teacher affiliates in some rural states. In total, I eliminated 12 states for lack of union variation (AK, DE, HI, IA, ID, ME, ND, NE, NV, SD, VT). Because the District of Columbia has only one school district, represented by one union, it was also eliminated. Moreover, the NEA and the AFT have merged in four states and these were also eliminated from the sample (FL, MN, MT, NY), as were a handful of districts where local district affiliation is with both unions. Finally, I eliminated New Mexico from the sample because collective bargaining laws in that state were allowed to sunset in 2000 (Lindy, 2011). Although they were re-authorized in 2003, this was in the middle of the second wave of SASS, and it is unclear whether teachers in the SASS sample would have been subject to collectively bargained contracts. After all of these restrictions on the data, as well as the exclusion of teachers in districts with missing national union affiliation, the final analysis sample includes 26,230 teachers in 18 states. Because of the sample

---

[4] Public employees in Missouri gained the right to collectively bargain in May, 2007, which is just before the final year of the SASS, making it unlikely that teachers in the dataset from that state were governed by bargained contracts.

limitations, any results are likely generalizable only to the population of teachers who teach in states with similar union contract negotiating rights.

In addition, because the sample contains data from teachers in less than half of the nation, the design weights that accompany the SASS were not utilized. These weights were created for the purpose of making findings from SASS nationally representative and generalizable to the entire population of teachers in the United States. Findings from these analyses do not make a claim of being generalizable to the entire population of teachers. Rather, any findings reported here are generalizable only to specific subgroups of public school teachers, working under certain policy conditions.

Methods: Research Questions 1 & 2

Retention Outcomes: My first outcome measure is a dichotomous variable indicating whether a teacher was retained in their school for one year following the initial survey. Teachers' self-reported follow-up information is only available from the subset of teachers who participated in the Teacher Follow-Up Survey (TFS). Therefore, to construct this variable for all the teachers in SASS in my analysis sample, I utilized the teacher status survey item, which was collected from schools as a preliminary indicator of teacher whereabouts one year after the survey. Participating schools were sent a questionnaire asking whether each teacher who had been surveyed in the previous year was still teaching at the same school one year later. Survey item choices included if the teacher was at the same school, if the teacher was at another school, no longer teaching, or deceased. Because the item does not indicate if teachers who moved schools also remained in the same district, I constructed a dichotomous outcome of retention indicating whether or not a teacher remained at the same school, after one year. I also constructed a dichotomous outcome indicating if the teacher remained in the teaching profession.

Analytic Model: Research Question 1

The key assumption of this model is that shortage-field incentives specifically, and not a bundle of work conditions or contract provisions associated with AFT affiliation, have a causal relationship to retention of shortage field teachers. To test this premise, I utilize an IV difference-in-differences model in which the excluded instrument is the interaction of the dichotomous union affiliation variable with the dichotomous indicator of whether a teacher is a shortage-field teacher. Teachers in AFT affiliated districts who do not teach in shortage-fields should not be impacted by the availability of shortage-field incentives within their district. Therefore, by controlling for other potential fixed difference between AFT and NEA teachers, a difference in outcomes between shortage and non-shortage field teachers represented by the AFT increases the likelihood that results are driven by the incentives themselves, and not by union affiliation as a whole. This model allows for an outcome that makes specific distinctions between shortage and non-shortage field teachers in AFT and NEA districts. In addition, it serves to increase the plausibility that the exclusion restriction is satisfied because it increases the likelihood that the identified relationship is through shortage field incentives, rather than through some other unmeasured pathway between union affiliation and either recruitment or retention outcomes.

Thus, in the first stage of this model, I predict the likelihood that a district will offer shortage-field incentives to shortage-field teachers if the teachers' union is affiliated with the AFT. I then use that predicted variable to further predict whether a given teacher was retained in teaching, subject to affiliation of union. The first stage of the model is:

$$D_k = Z_k(AFT * shortfield) + X_k(AFT) + X_i(shortfield) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where D is modeled as the likelihood that district k offers shortage field incentives, as a function of whether the teachers' union in district k is affiliated with the AFT, controlling for a vector of other covariates, as well as state and survey year fixed effects for teacher i in school j. The second stage takes the form:

$$Y_{ik} = a_{ijk} + \hat{D}_{ik} + X_k(AFT) + X_i(shortfield) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where Y is the probability that teacher i in district k is retained for one year after the initial survey. This outcome is a function of $D\text{-}hat_{ik}$ which is the predicted likelihood that shortage field teacher i in AFT district k taught in a district that offered incentives; X_{ijk} is a vector of teacher, school, and district covariates; state and survey year fixed effects are included, and in both stages, e_{ijk} is the uncorrelated error term.

Covariates in both stages are designed to control for known associations between school environment and a given teacher's experience at work. For example, there is evidence of strong relationships between teacher turnover, the level of poverty in a school, and the ethnic composition of students (Darling-Hammond, 2004; Hanushek, Kain, et al., 2004). Therefore, control variables include district urbanicity, school percent

Black and Hispanic students, school size, school percent of students classified as English learners (ELL), school percent of students eligible to receive free or reduced lunch, school percent of students with individual evaluation plans (IEPs), whether a school is a charter school, teacher age, gender and years of experience, school level taught, whether teacher is fully certified, and whether the teacher is a new teacher (less than 3 years experience). In order to control for variation in state teacher labor market conditions, I also include state fixed effects in the model. In addition, to control for potential confounding factors associated with time specific changes in external labor market conditions between the first and last data collection periods, I include survey year fixed effects in the model.

Methods: Research Question 2

Recruitment Outcomes*:* In order to investigate the impact of shortage-field incentives on recruitment, I will utilize two different recruitment outcomes.

Given that the presence of economic incentives are hypothesized to increase the quality of teachers who will be attracted to teaching, I will also utilize an outcome that measures quality of teachers' academic background. However, because teacher quality is both difficult to define and measure, the data do not contain any direct measures of teacher quality. Thus, I will use a proxy measure of teacher academic ability. Following several studies that have utilized data on competitiveness of teacher undergraduate institution as measured by Barron's college rankings (Angrist & Guryan, 2004; Boyd et al., 2008; Clotfelter et al., 2011; Hoxby & Leigh, 2004), to indicate teacher quality I will construct an outcome measure of teacher quality by matching the SASS teachers' undergraduate institution to the ranking the school received from Barron's in the decade during which the teacher graduated. Specifically, the file of Barron's rankings available from NCES provides the ranking assigned to a given school for the years 1972, 1982, 1992, and 2004.

The Barron's rankings are constructed from several measures of the academic quality of undergraduate students enrolled in a particular institution, including High School GPA and class rank, as well as median SAT and ACT scores, and percentage of students accepted of those that applied. All of the categories changed in terms of admission averages over the years, particularly in relation to the SAT and ACT scores of admitted students (full rankings criteria by decade can be found in Appendix Table 2). Therefore, I assigned the teachers who graduated 1967-1976 the 1972 ranking; teachers who graduated 1977-1986 were assigned the 1982 ranking, teachers who graduated 1987-1996 were assigned the 1992 ranking, and all teachers who graduated 1997-2006, which is the latest year a teacher could have graduated and been surveyed for the last year of the SASS, were assigned the 2004 ranking.

To construct a ranking index variable, I created a 9-point scale in which the most competitive colleges were ranked 9, and those rated as non-competitive were coded as 1. Some of the colleges ranked in three of the categories- highly competitive, very competitive, and competitive- also include "plus" ratings, meaning that specific colleges are ranked at the higher-end of their ranking category. To account for these nuances in ranking, I created additional categories, which I called "Highly Competitive Plus", "Very Competitive Plus", and "Competitive Plus". This allowed me to expand the scale from including 6 ranking categories, to including 9.

I also conducted a robustness check of this quality measure by analyzing whether shortage-field teachers in districts with shortage field incentives are significantly more or less likely to be teaching out-of-field, as defined by the field of their undergraduate major. To construct this measure I utilized data from SASS on teacher assignment and major field of Bachelor's degree. Teachers with bachelor's degrees in math, science, or other related fields such as statistics or engineering, were coded as having a STEM degree. If they were also assigned to teach math or science, they were counted as being an in-field teacher. Teachers with degrees in special education, who were also teaching special education classes were coded as being in-field teachers as well. Finally, teachers who had degrees in teaching English as a second language, bilingual education, or Spanish and who were assigned to teach bilingual classes were also coded as infield teachers.

Analytic Model: Research Question 2

In the first stage of this model, I again predicted the likelihood that a district offered shortage-field incentives to shortage-field teachers if the district teachers' union is affiliated with the AFT. In the second stage, I use that predicted variable to further predict whether districts with shortage-field incentives hired teachers who graduated from more competitive undergraduate institutions, or were more likely to employ teachers teaching in-field. The first stage of the model for the recruitment outcomes is identical to the first stage model above:

$$D_{ik} = Z_k(AFT * shortfield) + X_k(AFT) + X_i(shortfield) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where D is modeled as the likelihood that teacher i in district k teaches in a district that offers shortage field incentives, as a function of whether the teachers' union in district k is affiliated with the AFT, controlling for a vector of other covariates, as well as state and survey year fixed effects for teacher i in school j.

The second stage takes the form:

$$Y_{ik} = a_{ijk} + \hat{D}_{ik} + X_k(AFT) + X_i(shortfield) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where, in the first recruitment outcome, Y is the competitiveness of the undergraduate institution of teacher i in district k. For the measure of in-field teaching, Y is a dichotomous outcome in which Y is the likelihood that teacher i in district k has a bachelor's degree in the subject in which he/she is assigned to teach.

Each of these outcomes is a function of $D\text{-}hat_{ik}$ which is the predicted likelihood that shortage field teacher i in AFT district k taught in a district that offered incentives; X_{ijk} is a vector of teacher, school, and district covariates; state and survey year fixed effects are included for the district recruitment outcome. In both stages, for all outcomes, e_{ijk} is the uncorrelated error term.

Covariates in both stages again include district urbanicity, school percent Black and Hispanic, school size, school percent of students classified as English learners (ELL), school percent of students eligible to receive free or reduced lunch, school percent of students with individual evaluation plans (IEPs), whether a school is a charter school, teacher age, gender and years of experience, school level taught, whether teacher is fully certified, and whether the teacher is a new teacher (less than 3 years experience). In order to control for variation in state teacher labor markets, I also include state fixed effects in the model. In addition, to control for potential confounding factors associated with time specific changes in external labor market conditions between the first and last data collection periods, I include survey year fixed effects in the model.

Causal Framework: Research Question 3

Within three waves of the SASS, there are 106,930 teachers in a total of 6,540 school districts. Although many of these districts were only surveyed once in three survey administrations, 3,340, or slightly more than half, were surveyed more than once in different years. Of these repeatedly surveyed districts, approximately one third provided shortage field incentives during at least one of the years they were surveyed. Some of these districts offered incentives in only one of the two or three years they were surveyed, while others offered them two of three survey years, and still others offered them every year they were surveyed. Henceforth, I will refer to these districts as multi-year incentive districts. In the majority of multi-year incentive districts the shortage-field incentive policies were inconsistent across time. Only 164 districts provided incentives during all the years they were surveyed, while 900 multi-year incentive districts offered shortage field incentives at least once, but less than every year they were surveyed.

Exploiting the changes in shortage-field incentive policy over time, I propose a fixed effects strategy to investigate the impact of incentives on recruitment and retention outcomes. By comparing changes in district outcomes in one time period to district outcomes in the same district, in another time period, this model allows me to make causal inferences about the viability of shortage-field incentives for improving the quality of teachers recruited into districts as well as the retention of teachers. Because each district essentially serves as its own counterfactual, the key assumption of this model is that the outcomes in a district over a given time period would be the same if a new incentive had not otherwise been added to district policy, or if a previously existing incentive had not been removed from district policy. By comparing the level of change within districts over time, I can more precisely isolate the effect of the incentive policies on teacher outcomes.

At the same time, because changes in policy are likely endogenous to the outcome, in the sense that districts that change incentive policies may do so because they are the districts most in need of creating change in either recruitment or retention conditions, I will also need to further isolate the effect to be able to examine the impact of incentive policy change within districts. Therefore, in order to increase the robustness of my model, I will also examine the change in outcomes between shortage and non-shortage field teachers in districts that report policy changes. Outcomes for non-shortage field teachers should remain unaffected by changes in policy in shortage-field incentives and so this dimension improves the causal inference of the results.

Analysis Sample: Research Question 3

In order to conduct these analyses, I utilize the full sample of teachers and districts that were surveyed more than once, in order to be able to make comparisons between districts with and without incentive changes over time. This will give me a sample of 82,580 teachers in 3,340 districts, which represents more than half of all teachers surveyed by SASS in the three survey administrations. Of these teachers, 36,110 were in districts that offered surveys at least once during the multi-year period, with

20,300 of them surveyed the same year their district offered the incentive. Of this sample, 254 districts, with 7,100 teachers, eliminated an incentive policy between the first and the second, or the second and third survey administration. However, a total of 540 districts, with 17,900 teachers added an incentive policy after not having one. Finally, 104 districts with 4,910 teachers either had an incentive, dropped it, and added it again, or added an incentive after not having one, and then dropped it again. It is possible, however, that these districts represent measurement error, and will be dropped from the analysis sample.

In addition, of the 1,064 multi-year incentive districts, a total of 210, with 1,630 teachers, did not have any shortage field teachers who responded to the survey. These were removed from analysis, making the final sample 80,950 teachers. In this final subset, there are a total of 18,590 shortage field teachers.

Although this model examines the same teacher outcomes as the models proposed by research questions 1 and 2, by examining the issue over time, rather than cross-sectionally, the analysis allows me to asses the way that recruitment and attrition dynamics respond to changes in policy within the same districts. In addition, by utilizing data from across all 50 states and the District of Columbia, these findings are more generalizable. Furthermore, there is only minimal overlap between the analysis sample used in these models and the analysis sample used in the IV models, meaning that results of this analyses are not be derived from the same teachers.

Once again, the SASS design weights were not utilized in this analysis. In this case, although the sample is much larger, because of the way the weights were constructed, teachers who were surveyed from the same school and district, had different probabilities for selection in different years of the survey. At the same time, teachers sampled from the same school and district in the same year, have the same weights. Therefore, utilizing the weights as constructed by NCES would mean that comparable teachers within the same school and district would be given unequal weighting in the analysis. Therefore, the inclusion of district, school, and teacher level controls, as well as state and survey year fixed effects provide balance for any potential residual over-sampling effects of various characteristics created by the initial sampling design.

Analytic Models: Research Question 3

For my fixed effects analyses, I estimated the following model for both the retention and recruitment outcomes:

$$Y_{ik} = X_1(incentive) + X_2(sf) + X_3(incent * sf) + X_{ij} + D_{district} + S_{survey} + \varepsilon_{ik}$$

where Y_{ik} represents the recruitment or retention of teacher *i* in district *k*; X_1 is an indicator variable for whether or not a given district offers shortage-field incentives. The X_2 parameter represents and indicator for whether a teacher is a shortage-field teacher, and X_3 is an interaction term, as well as the variable of interest, of whether or not being a shortage-field teacher in a district that provide incentives has an impact on the outcome *Y*. X_{ij} is a vector of school and teacher level covariates, which include means of student minority, ELL, IEP, and free and reduced price lunch enrollment, as well as school size, grade level, and if the school is a charter. Teacher covariates include teacher gender, if the teacher is a new teacher, fully certified, or holds BA in a STEM field. The parameter *D* represents individual district fixed effects. Survey year fixed effects are indicated by *S*.

I estimate versions of this model on a variety of different sub-samples. For example, the first estimation only includes districts that actually reported a change in incentive policy between survey years. Those that never offered an incentive policy, or those that reported offering incentives every year they were surveyed are not utilized in the analyses, which allows me to better estimate the impact of policy change on outcomes.

Chapter 3

Policy Context: Setting the Stage

Teaching Contract Document Analysis

In order to understand the empirical findings presented in the next chapter, I wanted to first set the stage for understanding the policy context in which the empirical data are located. Because my identification strategy is predicated on the assumption of variation in teaching contracts, I begin by conducting a document analysis of a subsample of negotiated teacher contracts from districts that were identified in the SASS as offering shortage-field incentives. This is followed by an analytic estimation of the magnitude of incentives offered to teachers in shortage fields.

The document analysis helps to contextualize the policy environment within which shortage field incentives have been both negotiated and applied. Unlike the theoretical assumption that all shortage-field teachers be paid salary differentials, my findings indicate that the actual implementation of incentive policy is much less clear, and potentially more arbitrary than the theoretical assumption of uniform differentials would suggest.

In total, the analysis sample for the full empirical analysis includes 350 districts in 18 states that offer a shortage-field incentive of some kind. In order to understand exactly how these incentives were applied, I searched for publicly available negotiated teacher contracts for these districts. In 3 states, (MI, NH, NJ) I have near complete teacher contract information for all districts in the state. In an additional 6 states (IL, OH, IN, KS, PA, WI) I have found some, but not all, contracts available for shortage-field incentive districts. I was unable to locate any teacher contracts available in the remaining 9 states in my sample (CA, CT, MA, MD, OK, OR, RI, TN, WA)[5].

In total, I reviewed the language on shortage-field incentives in 129 teacher labor contracts from districts that indicated they provide shortage-field incentives in the SASS. I then classified the type of language I found about shortage-field incentives in these contracts in the following manner:

Strong: Any contract that listed a specific dollar amount to be paid to teachers in shortage-fields. Some examples of language classified as strong included:

> "Teachers with an ELL certification will be paid an additional $1,000 per year."
>
> "Special ed teachers hired prior to August, 1977 get a $220 differential".
>
> "The District shall have the option of paying a one-time contract signing bonus to newly hired Speech Pathologists, Occupational Therapists, Physical Therapists, and other employees in areas defined by the Board in agreement with the Union President as chronic-certified and/or licensed shortages.
> Hard to Fill Bonus: 2009-2010=$2,454.17
> 2010-2011=$2,454.17"

The main criterion for the classification of a "strong" contract was that it discussed a specific dollar amount to be paid to teachers in specific fields. Out of the 164 contracts I reviewed, 22 were classified as containing strong, specific language about incentive magnitude.

Vague: I classified a further 32 contracts as containing vague language about incentives or placement of teachers higher on the salary schedule as a recruiting tool. This category was more varied. Some examples of contracts classified as vague include:

> “Newly hired teachers who are in critical shortage areas and who are properly certificated may be placed up to four steps above the beginning salary step”
>
> "The board of education reserves the right to exceed the salary schedule to obtain the services of a teacher if it is for the good of the district".
>
> “For teachers recommended to fill positions of ‘critical shortages’, placement may be advanced up to two (2) steps beyond applicable teaching experience.”
>
> “Upon initial employment teachers for positions that are critical in need as designated by the [Name] County Superintendent of Schools may be placed up to step 11 in 2000-2001, step 12 in 2001-2002, step 13 in 2002-2003, and step 14 in 2003-2004.”

The main criteria for the “vague” classification was that there was enough language to explain the positive incentive reported in SASS, but which was otherwise unclear about whether or if specific teachers of specific subjects were actually paid differentials of any specific amount, at any point in time.

No Language: Despite the positive reporting for district offered incentives in the SASS, I could find no language to indicate that incentives might be offered in 75 of the 164 (45 percent) contracts I reviewed.

No Contract Available: I was unable to find contracts for 35 of the districts in the 9 states where I was able to locate some or all teacher contracts.

The final column of Table 25 may potentially offer some explanation as to why so many of the teacher contracts in districts that reported offering incentives did not contain any language indicating any detail that would further illuminate the type of incentives offered. The SASS sample contains data from surveys beginning in the 1999-2000 school year, and collected again in 2003-2004, and 2007-2008. However, publicly available teacher contracts tend to be provided for that which was most recently negotiated. Many of the contracts I found were negotiated in 2008 or later. The final column of Table 11 provides an accounting of the number of districts that reported positive for incentive offerings, where I also have the corresponding negotiated contract for the same year. In total, I have just 9 districts where the available contract matches the survey year in which the incentive was reported.

Table 2

Contract Language Classification by State

State	Strong/Specific contract language	Vague contract language	No contract Language	Contract not found	Number of districts where contract and survey year overlap
NH (n=14)	3	7	4	0	1
NJ (n=16)	5	5	4	2	3
MI (n=24)	1	5	17	1	3
IL (n=13)	2	0	10	1	1
OH (n=18)	1	0	5	12	0
IN (n=6)	0	0	5	1	0
KS (n=26)	7	7	9	3	0
PA (n=24)	3	2	12	7	1
WI (n=23)	0	6	9	8	0
Total (n=164)	22	32	75	35	9

Summary

This document analysis provides greater context for which to understand the results of the empirical models, presented below. For example, even for those contracts that contained language about incentives, the application of those incentives remains questionable. This is because although the language gives district personnel freedom to offer incentives, it does so in a manner vague enough to make it plausible that many shortage-field teachers hired into those districts were not actually offered incentives upon hire. Furthermore, the conditions under which Human Resources staff may actually be authorized to make incentive offers to prospective hires is so vague, it seems unlikely that in many of these districts the availability of incentives are used as a recruitment tool, let alone as part of outreach efforts to populations of prospective teachers who may not otherwise consider a career in teaching.

In essence, although theory predicts that a change in salary should serve to change the selection mechanism of those who consider teaching, in practice if salary increases are to function as a recruitment tool, then reasonably those who are currently deterred from teaching because of salary need to be given positive information that salary conditions have changed. Thus, if a higher salary is used only as a tool for recruitment in a time of emergency, it is unlikely that it is applied as part of a concerted effort to change the pool of applicants, but rather is used only as a means to secure a specific candidate from an existing pool of job seekers.

Incentive Magnitude Estimation

In addition to the somewhat uncertain policy environment illuminated by the labor contracts described above, it is also important to motivate any discussion of financial incentives by framing it in terms of magnitude. For example, when theorists discuss shortage field incentives as a tool for reducing opportunity costs, does this mean that incentives for shortage-field teachers should match the average differential between teaching and non-teaching options for teachers with STEM degrees? Or is there some additional amount that will provide enough of a difference to reduce attrition to a more tolerable level? Does the amount needed to make a difference in teacher outcomes differ by labor market context?

Because the SASS does not contain information specifically about how much extra money was provided to shortage-field teachers in districts that offered shortage-field incentives, I am unable to answer all of these questions. However, utilizing the teacher annual salary data information provided in the survey, I am able to employ empirical models to estimate some trends, including how much more money shortage-field teachers in districts that provide incentives earn, relative to other teachers, both across the full sample and in different regions of the country. In addition, I can estimate how much more shortage-field teachers who taught in AFT affiliated districts earned the year they were surveyed.

Thus, in order to try to understand the salary context in which shortage field incentives are offered, I first calculated simple means of teacher salaries across the full sample, and by region of the country. Across all regions, I find an average of $53,331. Salaries are slightly higher in the Northeast, where the average is $58,548. The lowest average is for teachers in the South, at $45,346. Mean salaries are $51,066 and $54,630 in the Midwest and West, respectively.

I then conducted simple linear regression models (Table 3) in which teacher salary is the outcome variable[6]. These models reveal substantial variation in teacher salaries in different regions of the country. In these OLS models, the variable of interest is the coefficient on the variable for shortage field teachers teaching in districts that offer shortage-field incentives. This coefficient provides an estimate of the size of the salary difference between shortage field teachers in districts that offer incentives and all other teachers in the sample, controlling for other characteristics.

In these models, the effect of being a shortage-field teacher in a district that provides shortage-field incentives is negative, but non-significant in both the full analysis sample, and across regions of the county. These findings suggest that, on average, shortage-field teachers teaching in districts that offer incentives receive lower salaries than other teachers, even when controlling for teacher experience, or being a new teacher. That said, none of the estimates are statistically significant, and the confidence intervals suggest that there is considerable variation in the amount of money teachers in these districts were paid. For example, in the full sample, the range for the differential is between earning $2,299 less than other teachers to earning $852 more than other teachers, annually. In the Northeast, the earnings range for shortage-field incentives is estimated between $3,330 less to earning $3,031 more. In the Midwest, the confidence interval ranges from $3,958 less to $1,577 more, and in the South the range is between

[6] All salary variables are CPI adjusted to 2007 dollars, which is the last and most recent year of the survey.

$5,031 less to $4,228 more. Finally, the confidence interval in the West ranges between earning $2,327 less to earning $1,653 more.

One possible explanation for this finding may be that districts that offer incentives are more likely to employ teachers with fewer credentials and years of experience generally, which are the main drivers of teacher salaries. If this is the case, than districts that offer additional incentives may be paying shortage-field teachers more than they pay other teachers in their own districts, but still less than more experienced or better credentialed teachers receive elsewhere. This is plausible because in both the Midwest and the Northeast, shortage-field teachers across all districts earn approximately $1,500 more than other teachers, and the results are statistically significant ($p<.01$). Since these regions combined comprise more than half the analysis sample, the findings indicate that districts without formal shortage-field incentive policies are paying their shortage-field teachers more than other teachers, even when controlling for factors such as experience or being a new teacher.

Moreover, in the South, districts that offer incentives pay all teacher in their districts an average of almost $1,900 ($p<.05$) less a year than other districts in that region, which suggests that because these models place all teachers together and do not adequately distinguish the selection bias between districts that do and do not offer incentives, the models are likely capturing the effects of the non-random characteristics of teachers and districts that do and do not offer incentives.

Table 3

OLS Salary Models, Full Sample and by Region

	Full Sample	Northeast	Midwest	South	West
Shortage Field Teacher*Incentive	-731	-174	-1,236	-401	-290
	(804)	(1,618)	(1,413)	(2,338)	(1024)
District offers incentive	-285	-413	-30	-1,882*	-346
	(450)	(813)	(862)	(946)	(730)
Shortage field teacher	918**	1,574**	1,570**	2,199	737
	(364)	(582)	(467)	(1,568)	(688)
Constant	40,018**	41,270**	25,328**	51,603**	49,229**
	(1,941)	(3,020)	(2,898)	(4,918)	(4,225)
Observations	23,110	7,060	7,040	3,420	5,590
R-squared	0.49	0.47	0.46	0.47	0.44

* significant at 5%; ** significant at 1%; Robust standard errors in parentheses

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clustered at the district level. All Salaries adjusted to 2007 levels.

Thus, in order to attempt to gain a better understanding of the salary differentials paid to shortage field teachers in AFT affiliated districts, I also conducted an analysis in which I utilized the available salary information as the outcome in my first stage equation from my IV models, as follows:

$$I_i = Z_k(AFT * shortfield) + X_k(AFT) + X_i(shortfield) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where *I* is the yearly income from teaching received by teacher *i.* In this model, the coefficient on the interaction term is an estimate of how much more shortage-field teachers in AFT districts are earning than either other teachers in their district, or more than teachers in NEA districts. This is a more targeted analysis, and is likely to more specifically illuminate a salary dynamic that is specific to the model estimates discussed in the next chapter.

When I run this model on the salaries of all teachers, I find that, on average, shortage-field teachers in AFT districts earn only about $260 more a year than non-shortage-field teachers (Table 4). However, the result is not statistically significant, and the confidence interval for the estimate places the true value in a range between $860 less to $1,370 more. This confidence interval may be indicative of a variety of implementation scenarios. It is possible that incentives are generally very small, that they are not large enough to overcome gaps between districts, or that many of the shortage-field teachers in the sample who work in districts that report offering incentives likely did not actually receive them.

However, I also run this model again, by regions. I find that shortage field teachers teaching in AFT affiliated districts in the Northeast earn and average of $3,800 more than non-shortage AFT field teachers, and that the result is statistically significant at the 0.01 level. By contrast, shortage field teachers in the Midwest, earn an average of $2,222 less than other teachers. The result is also statistically significant at the 0.01 percent level.

I also find negative estimates in both the South and the West. However, these results are not statistically significant, and the confidence intervals are very large, particularly in the South, where the standard error is more than 3 times the size of the coefficient estimate. In addition, while the sample sizes in the Northeast and Midwest regions are similar at approximately 7,000 teachers teaching in each of these regions, they are smaller in the West, and particularly the South, which likely contributes to the uncertainty of these estimates.

Based on these findings, I also run a model in which I exclude only the Midwestern states, and I find that the average estimated salary for shortage-field teachers in AFT districts is $1,610 higher than for other teachers. This result is statistically significant at the 5 percent level. This finding suggests that outside of the Midwest, shortage field teachers teaching AFT districts are earning somewhat more than their counterparts in the NEA, as well as more than non-shortage field teachers. In addition the average amount more they are earning is approximately the size of a single salary step in may districts, which is in line with the AFT policy of placing shortage field teachers higher on the salary scale.

In addition, the findings for these analyses suggest that the teacher labor markets and compensation in the Midwest is not comparable to labor markets in other regions of the country, and suggest that analyses on other teacher outcomes may be impacted by dynamics that differ across regions.

Table 4
First Stage Salary Models, Full Sample and by Region

	Full Sample	Northeast	Midwest	South	West	All regions except Midwest
AFT*Shortage Field	256	3,828**	-2,222**	-719	-1,760	1,607*
	(569)	(1094)	(805)	(2,234)	(1,107)	(780)
Constant	37,583**	34,540**	16,938**	35,789**	45,918**	41,977**
	(1,038)	(1,959)	(1,658)	(2,667)	(2,157)	(1291)
Observations	23,110	7,060	7,040	3,420	5,590	16,070
R-squared	0.49	0.48	0.46	0.47	0.44	0.50

* significant at 5%; ** significant at 1%; Robust standard errors in parentheses
Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clustered at the district level. All Salaries adjusted to 2007 levels.

Summary

In this chapter, I set the stage for understanding the policy context in which shortage field incentives are implemented in two ways. First, I conducted a document review of a sub-sample of union negotiated contracts from districts that indicated they offered incentives. These analyses suggest that the implementation of incentives may be limited to securing teachers who are already in the teacher labor market pool, for specific positions. There is little evidence, however, that these incentives are utilized as marketing tools, or to expand the appeal of teaching to non-traditional candidates, which may undermine their ability to influence the selection mechanisms of potential teachers into the profession.

Interestingly, in line with the type of policy endorsed by the AFT, the language of many of the contracts that do discuss incentives suggests placing shortage-field teachers on higher steps of the salary scale rather than providing specific dollar amounts, or otherwise deviating from the teacher compensation scheme. In addition to language, the contracts also detail the size of the steps on the individual district salary schedules. In many districts, a single step is typically worth somewhere between $1,000 and $1,500 dollars, which is approximately the size of the average incentive magnitude estimated for the non-Midwestern regions, in my second analysis.

In this step, in the absence of data about the size of shortage-field incentives, I utilized teacher salary data to estimate an average incentive magnitude. I found evidence to suggest that teacher compensation is lower in the Midwest than in other regions of the country. In addition, I found a potential difficulty for the viability of my identification strategy in the data from the Midwest. Namely, shortage-field teachers in AFT affiliated districts in this region actually earn significantly less than other teachers, which suggests that I might also find differences in recruitment and retention outcomes in that region, using the models I proposed in Chapter 2. Moreover, once I remove the Midwest from the pool for estimating the salary differential, I find that shortage field teachers teaching in AFT affiliated districts earn an average of $1,600 more than their non-shortage field counterparts. This estimate supports the descriptive evidence found in the contract document review because this amount of money is also approximately the size of one salary step in many teacher contracts.

Chapter 4

Findings

Descriptive Results: Research Questions 1 and 2

Retention Outcomes

Descriptive analyses comparing districts that do and do not offer incentives from the analytic sample show similar differences as were found between districts from the complete sample (Table 5). Specifically, districts that offer incentives have higher proportions of Black and Hispanic students, more students eligible for the federal lunch program, and more ELL and IEP students. They are also larger, and more likely to be urban. Interestingly, the mean salaries teachers receive are about $600 a year higher in districts that offer incentives. Although this is a statistically significant difference, it is questionable whether such a difference, which amounts to about 50 pre-tax dollars a month, is a substantively meaningful difference in the lives of teachers.

Table 5
Characteristics of Districts that Do and Do Not Offer Shortage-field Incentives (Analytic Sample)

Variable Name (at district level)	Shortage-field Incentives (n=310)	No Shortage-field Incentives (n=1,690)
District Mean % Black	0.23	0.11
District Mean % Hispanic	0.19	0.11
District Mean % in lunch program	0.45	0.34
District Mean Enrollment	35,245	22,525
District Mean Teacher Salary	$50,046	$49,450
District Mean % ELL Students	0.23	0.18
District Mean % IEP Students	0.12	0.08
Urban	0.46	0.26
Suburban	0.41	0.55
Rural	0.13	0.18
South	0.17	0.14
Midwest	0.27	0.32
Northeast	0.28	0.30
West	0.26	0.22

**T-test statistics. Mean differences all significant at the p<.001 level*

Descriptive comparisons of teachers in districts represented by the AFT and the NEA show that 24 percent of AFT teachers teach in districts with incentives, compared to 13 percent of NEA teachers (Table 6). The statistics show that there are significant differences between the kinds of districts represents by the two unions. Those represented by the AFT have higher percentages of Black and Hispanic students, are larger, have more students eligible for free or reduced lunch, and are more likely to be urban. However, there are no significant differences between the numbers of ELL or IEP students in the districts represented by the two unions. Mean salaries in AFT districts are also higher.

Table 6
Characteristics of Districts Represented by the AFT and the NEA

Variable Name (at district level)	AFT (n=210)	NEA (n=1,790)
Has Shortage-field incentives	0.24	0.13
District Mean % Black	0.29	0.10
District Mean % Hispanic	0.21	0.11
District Mean % in lunch program	0.49	0.33
District Mean Enrollment	73,299	14,791
District Mean Teacher Salary	$51,041	$49,243
% ELL Students	0.19	0.19
% IEP Students	0.09	0.08
Urban	0.65	0.22
Suburban	0.30	0.58
Rural	0.05	0.20

**T-test statistics. Mean differences all significant at the p<.001 level, except differences in district enrollment in ELL students, and students with IEPs which are not significant.*

Descriptive statistics also show differences between teachers who left their school after one year, and those who remained (Table 7). Overall, 14 percent of the teachers in the analysis sample were not retained in their school after one year. Those who left were more likely have taught in urban schools, and Southern schools, whereas in the Midwest fewer teachers left teaching. Leavers were also more likely to have taught in AFT affiliated districts. Teachers more frequently left schools with higher proportions of Black students, but there was no difference in the average number of Hispanic students in the schools of teachers who were and were not retained. There was also little difference in the average rates of ELL students, or percentages of students with IEPs. However, the schools teachers left had an average of 4 percent more of students eligible for free/reduced price lunch. Shortage-field teachers were also somewhat more likely to leave their schools. However, 46 percent of leavers were high school teachers, compared to 52 percent of those who remained, compared to elementary and middle school teachers who, combined, comprised 48 percent of those who left and 42 percent of those who stayed, indicating that high school teachers as a whole have higher retention rates than other teachers. Fully certified teachers were also more likely to remain, as were teachers with a BA in a STEM field. Male teachers comprised 32 percent of those who left, and 34 percent of those who stayed, which is interesting in light of the teacher gender literature presented above, suggesting that males who select into teaching, are more likely to remain. Finally, consistent with the literature, 24 percent of teachers who left their schools were new teachers, compared to just 14 percent of those who remained.

Table 7
Characteristics of Teachers Who Were and Were Not Retained in their School for 1 Year

Variable Name (at teacher level)	Not Retained (n=3,670)	Retained (n=22,560)
Total	0.14	0.86
Shortage-field incentives	0.17	0.15
City	0.31	0.29
Rural	0.17	0.18
South	0.17	0.14
Northeast	0.30	0.31
Midwest	0.31	0.33
West	0.23	0.23
AFT	0.19	0.16
NEA	0.81	0.84
Charter	0.02	0.01
School % Black	0.17	0.13
School % Hispanic	0.12	0.12
School % IEP	0.22	0.21
School % LEP	0.11	0.09
School % free/reduced lunch	0.37	0.33
New Teacher	0.24	0.14
Shortage field	0.25	0.23
Elementary	0.32	0.28
Middle	0.16	0.14
High School	0.46	0.52
Combined grades	0.07	0.06
Fully certified	0.82	0.89
Female teacher	0.68	0.66
Male teacher	0.32	0.34
STEM BA	0.12	0.14

T-test statistics. Mean differences significant at the $p<.01$ level. Mean difference not significant for rural, Northeast, Midwest, West, combined grades, or teacher gender.

A little more than half of all the teachers who left their schools after one year, did not move to a new school, but rather chose to leave teaching. Specifically, although 6 percent of teachers moved to another school, 8 percent left the profession. Descriptive statistics of the characteristics of the schools and teachers for this subgroup of leavers (Table 8), showed similar results to those found for the school retention outcome, suggesting that teachers in districts with shortage-field incentives are retained in teaching less often, as are teachers in AFT districts and charter schools.

Table 8
Characteristics of Teachers who Were and Were not Retained in the Teaching Profession for 1 Year

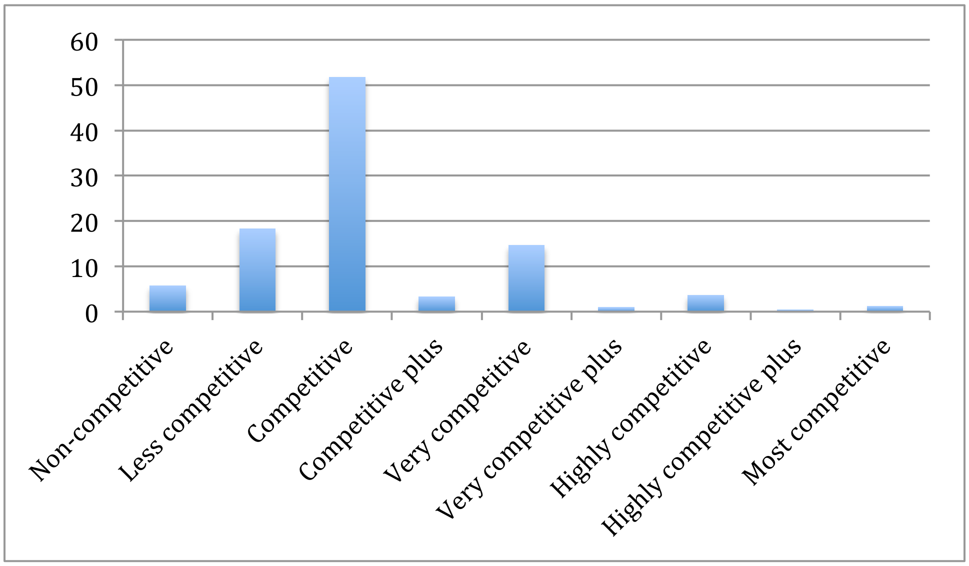
Variable Name (at teacher level)	Not Retained in Teaching (n=2,000)	Retained in Teaching (n=24,230)
Total	0.08	0.92
Shortage-field incentives	0.17	0.15
City	0.31	0.29
Rural	0.16	0.18
South	0.16	0.14
Northeast	0.30	0.31
Midwest	0.30	0.32
West	0.23	0.23
AFT	0.20	0.16
NEA	0.80	0.84
Charter	0.02	0.01
School % Black	0.18	0.13
School % Hispanic	0.12	0.12
School percent IEP	0.22	0.21
School % LEP	0.11	0.09
School % free/reduced lunch	0.38	0.34
New Teacher	0.18	0.15
Shortage field	0.22	0.24
Elementary	0.28	0.28
Middle	0.16	0.14
High School	0.50	0.52
Combined grades	0.06	0.06
Fully certified	0.83	0.88
Female teacher	0.66	0.67
Male teacher	0.34	0.33
STEM BA	0.12	0.14

T-test statistics. Mean difference significant at $p<.01$; Mean difference not significant for Northeast, Midwest, West, shortage field, elementary, high school, combined grades, gender, Stem BA.

Recruitment Outcomes

Unlike the descriptive evidence for the retention outcomes, the descriptive evidence for teacher academic background indicates that there is very little variation between teachers, at least as far as the competitiveness rankings of their undergraduate institutions. Specifically, over 50 percent of teachers graduated from colleges clustered in the Competitive and Competitive Plus ranking categories (Figure 1).

Figure 1
Distribution of Teachers by Barron's Rank of Undergraduate Institution



As such, there are almost no significant differences in the academic backgrounds of different categories of teachers (Table 9). For example, the mean academic rank of teachers in districts that both do and do not offer shortage-field incentives is 3.3. Notable exceptions to this trend are between new teachers, who have an average college rank of 3.7, compared to more experienced teachers, who have an average rank of 3.2. It is unclear from the data, however, if new teachers actually graduated from better colleges than their more experienced counterparts, or if these teachers benefit from coming of age in an era when colleges themselves have become more strategic about climbing the rankings ladder. This phenomenon could impact this difference because two teachers who graduated from the same college 20 years apart, could have different rankings that are due either to actual academic improvement, or to a given college increasing its rankings by engaging in competitive ranking game play (Farrell & Van Der Werf, 2007).

A 3.7 to 3.2 difference is also the difference between the college rank of teachers with a BA in a STEM field and other teachers. Interestingly, teachers who are fully certified graduated from colleges with slightly lower academic rankings than teachers who are not fully certified, although this could reflect the same difference as the new teacher difference since new teachers are also more likely to be uncertified. Likewise, charter school teachers also graduated from colleges with an average rank that is 0.02 points higher than teachers in traditional public school, but charter school teachers also tend to be younger than other teachers. Teachers in the Northeast graduated from higher ranked colleges than teachers in other regions. Specifically, the Northeast average is 3.7, compared to 3.1, with Southern teachers graduating from colleges with the lowest academic rankings. It is unclear from these data if the stronger academic backgrounds of Northeastern teachers is due to the clustering of competitive colleges in this region, or to other factors that attract graduates from these colleges to teach in this region.

Table 9

Characteristic of Teachers by Barron's Rank of Undergraduate Institution

Variable Name (Teacher Level)	Variable=1	Variable=0
District offers shortage-field incentives	3.3	3.3
City	3.3	3.3
Rural	3.0	3.4
South	3.0	3.3
Northeast	3.7	3.1
Midwest	3.4	3.0
West	3.3	3.3
AFT	3.4	3.3
Charter school teacher	3.5	3.3
New teacher	3.7	3.2
Shortage-field teacher	3.3	3.2
Elementary	3.2	3.3
Middle	3.3	3.3
High School	3.4	3.2
Fully certified	3.3	3.6
Female teacher	3.3	3.3
Has STEM BA	3.7	3.2
Infield	3.4	3.3

Looking at descriptive statistics for the infield outcome, I compared only shortage-field teachers who were either teaching infield or out-of-field (Table 10), as opposed to comparing them to non-shortage field teachers. For these analyses, infield was defined as a teacher who had a BA in any STEM field, in any special education field, or in bilingual or English language learner education, or in Spanish. Teachers with BA degrees in other foreign languages, which included German, Latin, French, Russian or Other Foreign Language, were not counted as infield because only a very small fraction of them were teaching bilingual education, and the vast majority were assigned to teach foreign languages, or other subjects, like social studies. Teachers with the aforementioned primary undergraduate majors were counted as infield teachers if their main teaching assignment included classes is math or science, special education, or bilingual education, respectively. Teachers were considered out-of-field if their major field of assignment was other than the subject area that best corresponded to their major. Surprisingly, considerably more shortage-field teachers were teaching out-of-field than infield. However, it appears that most of the out-of-field teaching was done by special education and bilingual teachers, given that 73 percent of shortage-field teachers who were teaching infield, also held a STEM BA, compared to just 1 percent of those teaching out-of-field. This descriptive finding suggests that when schools can hire a teacher with the appropriate subject matter background, they are likely to assign them to the classes that best fit their skills.

In addition, infield teachers taught in schools with slightly higher mean percentages of Black students, but also taught in schools with far fewer students eligible for free or reduced lunch, and lower mean percentages of Hispanic, IEP, and LEP students. These numbers however, may reflect the fact that infield teachers tend to teach in more wealthy schools, but they also may be less likely to be found in elementary schools, where students tend to have more IEPs, and where younger students with less exposure to English may be more likely to be classified as language learners than their secondary school counterparts. Lower percentages of infield teachers are found in the Midwest and the Northeast, as compared to the percentages of out-of-field teachers teaching in these regions, and the opposite is true for infield and out-of-field teachers located in both the West and South.

Table 10
Characteristics of Shortage-field Teachers who are Teaching Infield and Out-of-field

Variable Name (Teacher Level)	Infield (n=1,860)	Out-of-Field (n=4340)
Total	0.30	0.70
Shortage-field incentives	0.15	0.16
City	0.39	0.30
Rural	0.15	0.17
South	0.18	0.13
Northeast	0.28	0.32
Midwest	0.29	0.33
West	0.25	0.22
AFT	0.20	0.18
NEA	0.80	0.82
Charter	~	0.01
School % Black	0.16	0.14
School % Hispanic	0.10	0.12
School percent IEP	0.14	0.23
School % LEP	0.05	0.10
School % free/reduced lunch	0.34	0.36
New Teacher	0.15	0.15
Elementary	0.11	0.24
Middle	0.12	0.20
High School	0.71	0.47
Combined grades	0.06	0.09
Fully certified	0.87	0.85
Female teacher	0.58	0.69
Male teacher	0.42	0.31
STEM BA	0.73	0.01

T-test statistics. Mean difference significant at p<.01; Mean difference not significant for shortage field incentives, rural, new teacher.

Overall the descriptive picture presented here serves to illustrate much of the literature that was presented in the last chapter. Specifically, that teacher attrition disproportionately impacts urban schools that serve higher percentages of Black and Hispanic students, as well as those that serve higher proportions of English learners, and those students who live in poverty. They also affirm the notion that new teachers in their first three years leave both their schools and the teaching profession at much higher rates than their more experienced peers. In addition, teachers in schools with higher concentrations of low income and minority students are less likely to be the schools staffed by infield shortage-field subject teachers. They are also more likely to be affiliated with the AFT, and are more likely to offer shortage field incentives.

The quality outcome indicator examining teacher's undergraduate academic institution, was found to be less variable than all the other outcomes and variables of interest. These findings suggest a uniformity in the quality of teachers. However, although there is less variation, teachers are clustered at the lower end of scale, indicating an overall mediocre supply of teachers, at least as indicated by the types of institutions they attended as college students.

Empirical Results: RQ1 & RQ2

Analytic Results: Retention in School

The analytic models used to address the first research question utilize a total sample of just over 23,000 teachers, which represents a drop of 11 percentage points from the full analysis sample utilized for the descriptive findings, and is due to the automatic listwise deletion of observations with missing data on any variables. Data were complete for almost all control variables, except the school reported percentages of ELL students, and students eligible for the federal lunch program. Although it is possible that these missing values could bias results, utilizing data available through the Common Core of Data (CCD), I spot checked a random sample of schools that were missing the lunch variable and found that schools that were missing this information were fairly evenly split between those with high percentages of lunch eligible students and those with zero lunch eligible students. A small minority of schools from this random sample were not listed in the CCD, indicating that the school may no longer be in existence. I was unable to check data on the percentage of ELL students at schools missing that information.

Results from the first set of IV regressions for the school retention outcome suggest that shortage-field teachers in AFT districts are significantly more likely to teach in districts that provide incentives (Table 11). Specifically, in the first stage, shortage-field teachers in AFT teachers are 17 percent more likely ($p<.01$) to teach in districts with incentives than are shortage-field teachers in NEA districts, suggesting that the interaction of AFT and shortage-field is a strong instrument for use in these models. In addition, the first stage F-statistic of joint significance for this model is 29.4, which surpasses the weak instrument test rule of thumb threshold of 10 (Angrist & Pischke, 2009).

In the second stage, however, I find that shortage-field incentives are not statistically significantly related to school retention. In this model, the coefficient of 0.122 is estimated between -0.09 and 0.33 indicating that the true value for retention is likely positive, but that these data are too noisy to obtain a statistically significant result.

Table 11

Retention-in-school Outcome, Full Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.122
			(0.105)
AFT*Shortfield	0.021	**0.171****	
	(0.018)	(0.014)	
AFT	-0.012	**-0.010****	-0.011
	(0.012)	(0.002)	(0.011)
Shortage-field	-0.022**	**0.122****	**-0.037***
	(0.007)	(0.005)	(0.017)
Constant	0.721**	-0.053**	0.7601**
	(0.045)	(0.014)	(0.043)
Observations	23110	23110	23110

* significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

One of the concerns about this model is that the SASS does not contain very much information about the external factors that may influence teacher retention decisions. Teacher labor markets tend to be local, but there is no direct evidence in the survey data about external job market factors, such as prevalence of other kinds of local industries, or average local wages. Although the use of state level fixed effects does control for differences between states, it does so by holding constant the state level variation in things like local labor markets, and industries. However, these fixed effects may not be enough to account for larger demographic trends in given areas because they do not account for region level phenomena that cross state borders and impact large segments of the labor force.

For example, an earlier descriptive study that analyzed SASS 1999-2000 data found that administrators across the Midwest reported fewer difficulties filling vacancies than did administrators in other parts of the country (Murphy et al., 2003). Specifically, that study looked at late fill vacancy rates[7] for teaching positions and found that the average in many states in the Midwest, including 4 of the 5 from that region that remain in my analysis sample, was less than 1 percent, which was the lowest possible rating. In contrast, the national average late fill vacancy rate was 1.5 percent, while the rate in many Western states was more than 2 percent, and in some cases as high as 5 percent. Although these do not seem like large differences, the result is extremely consequential in terms of volume of teachers that must be hired and retained in the different regions. By way of comparison, according to the Murphy paper, California employed a total of 300,000 teachers and had a late-fill rate of 2.3 percent, which meant that that state hired 6,900 teachers after the start of the school year. By contrast, collectively Illinois, Michigan, and Ohio, employed 350,000 teachers, and had late-fill rates of 1.5 percent, 0.8 percent, and 0.5 percent respectively. Together, these states hired a total of 3,500 teachers after the start of the school year, which is just half the number that had to be hired in California, despite a three state teacher population that is 17 percent larger.

In addition, descriptive results from Table 7 above echo this trend and show that slightly more teachers from the Midwest are retained in their schools, than are teachers from other regions of the country. In addition, data from the American Community

---

[7] Late fill rates are the percent of teaching positions that remained unfilled at the start of the school year.

Survey[8] suggest that workers in the public education sector in the Midwest comprise a smaller average proportion of the overall workforce than in other states. Specifically, calculations conducted by the author using ACS data found that the five Midwestern states in the analysis sample had an average public education sector workforce that represented 6.7 percent of all workers, while the rest of the states in the analysis sample had an average public education workforce of 7.5 percent of the total, suggesting possible overall lower demand for teachers in these states. Additionally, descriptive analysis provided above indicated that Midwestern teachers were less likely to either teach in districts that offered shortage-field incentives or leave their schools after one year. Although these findings are all descriptive in nature, they nevertheless suggest that teacher shortages are a localized phenomena, and specifically that teacher recruitment is less of a problem in the Midwest than in other parts of the country. They also illustrate why state fixed effects may be inadequate for controlling for the regional variation in teacher hiring trends. Finally, the salary analysis in the previous chapter indicates that both general teacher salaries, and salaries for shortage-field teachers in AFT affiliated districts in the Midwest are lower than elsewhere. Therefore, I re-ran the analysis excluding teachers and districts in states located in the Midwest.

This sub-analysis reduces my sample to 16,000. At the same time, removing this region (Table 12) increases the predicted likelihood in the first stage that shortage-field teachers in AFT districts will teach in districts that offer incentives to 25 percent ($p<.01$). Additionally, the instrument in this model remains strong, despite the reduction in sample size ($F=26.4$). The estimated impact on retention also increases from 10.5 to almost 18

[8] Data from the 2010 American Community Survey, U.S. Census Bureau. Table title "Industry by Class of Worker for the Civilian Employed Population 16 Years and Over", by state.

percent (p<.05) and is also now statistically significant, indicating that shortage-field incentives have a positive impact on school retention in regions that face greater recruitment challenges.

Table 12

Retention-in-school Outcome, Non-Midwest Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			**0.178***
			(0.074)
AFT*Shortfield	**0.045***	**0.251****	
	(0.021)	(0.019)	
AFT	-0.022	-0.021**	-0.018
	(0.015)	(0.002)	(0.015)
Shortage-field	**-0.019***	**0.116****	**-0.040***
	(0.009)	(0.006)	-0.015
Constant	0.768**	-0.078**	0.782**
	(0.054)	(0.019)	(0.054)
Observations	16,070	16,070	16,070

* significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Thus far, all models have grouped all shortage-field teachers together, such that special education and bilingual teachers have been defined in the same group as math and science teachers. However, it is reasonable to assume that teachers in these different fields face different labor market conditions. Therefore, in order to determine whether the effect holds for different kinds of shortage-field teachers, I also conduct subgroup analyses for just math and science teachers, as well as for special education teachers and bilingual teachers. In order to operationalize these models, and not loose further sample size, I replaced the shortage-field variable in the models with indicators specifically for

STEM teachers, and special education and bilingual teachers, such that the excluded instrument becomes an interaction between AFT, and these new indicator variables. The control variables were also replaced with these new indicators, such that the first stages of each of these models is specified as follows:

$$D_{ik} = Z_k(AFT * STEM) + X_k(AFT) + X_i(STEM) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

$$D_{ik} = Z_k(AFT * SEB) + X_k(AFT) + X_i(SEB) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

where STEM indicates a teacher of math or science, and SEB indicates a teachers of special education, or bilingual education.

In the model that included math and science teachers alone the instrument remained strong (F=12.3), but the likelihood that these teachers would teach in districts offering incentives dropped to 19 percent (p<.01) in the non-Midwestern sample (Table 13). The estimated coefficient of on the incentive variable remained at about 18 percent, however in this specification it is only statistically significant at the 10 percent level. One likely explanation for this drop in significance is sample size. In particular, out of more than 16,000 teachers in non-Midwestern regions only 360 teachers actually taught math or science in a district affiliated with the AFT.

Table 13

Retention-in-school outcome, STEM Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			**0.182~**
			(0.113)
AFT*STEM	0.035	**0.190****	
	(0.022)	(0.026)	
AFT	-0.015	**-0.009****	-0.014
	(0.016)	(0.002)	(0.016)
Stem teacher	-0.009	**0.112****	-0.030
	(0.011)	(0.007)	(0.019)
Constant	0.763**	-0.039**	0.770**
	(0.055)	(0.014)	(0.055)
Observations	16,070	16,070	16,070

~ significant at 10%; *significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

By contrast, the results for the special education and bilingual sample alone (Table 14), suggest that these teachers are 30 percent ($p<.001$) more likely to teach in districts offering incentives, and the instrument also remains strong is this model ($F=14.5$). However, despite the increase in likelihood of a incentive, the estimate for retention suggest that these teachers are only 15 percent ($p<.10$) more likely to remain in their school after 1 year than comparable teachers who did not teach in districts providing incentives. These findings suggest that despite the fact that shortage-field incentives are more likely to be offered to teachers of special and bilingual education, they may be slightly more effective for actually impacting retention of teachers of math and science. This may be related to the distinction discussed above between shortages driven by high turn-over and opportunity costs, and shortages driven by lack of supply. It may be that more districts offer incentives for special education and bilingual teachers because more districts have difficultly finding teachers with these credentials in the recruitment phase.

Table 14

Retention in School, Special Education and Bilingual Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			**0.150~**
			(0.091)
AFT*Special ed and Bilingual	0.046	**0.306****	
	(0.031)	(0.027)	
AFT	-0.016	**-0.011****	-0.014
	(0.015)	(0.002)	(0.014)
Special ed and Bilingual	**-0.030***	**0.113****	**-0.047***
	(0.015)	(0.009)	(0.023)
Constant	0.763**	-0.036**	0.748**
	(0.054)	(0.027)	(0.054)
Observations	16,070	16,070	16,070

~significant at 10%; *significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Finally, because of substantial evidence suggesting that new teachers in particular leave the profession at disproportionately high rates (Darling-Hammond, 2004; Ingersoll, 1999), I also conducted the analysis specifically for new shortage-field teachers (those with 3 years or less of experience) in the non-Midwestern sample. In this model, also to maintain sample size while restricting the analysis to the subsample of interest, I again replaced the original shortage-field teaching indicator with an indicator specifically for new shortage-field teachers, such that the first stage of the model is now specified as follows:

$$D_{ik} = Z_k(AFT * New * SF) + X_k(AFT) + X_i(New) + X_i(SF)$$
$$+X_i(New * SF) + X_{ik}(New * AFT) + X_{ik}(SF * AFT) + X_{ijk}B + S_{state} + S_{survey}$$
$$+ e_{ijk}$$

Where the instrument is a triple interaction between AFT, New Teacher, and Shortage Filed. However, the first stage F-statistic in this model drops to 4.34, and fails the weak instrument test. This problem likely occurs because in the non-Midwestern sample, there were only 160 new shortage-field teachers who taught in AFT affiliated districts. Because of these findings, the findings for the model, presented in Table 15 should be considered merely illustrative, and be interpreted with caution.

That said, in this model I found that new shortage-field teachers in AFT districts have a predicted likelihood of teaching in a district offering a shortage-field incentive of 25 percent ($p<.01$), but that these incentives do not have a statistically significant impact on the retention of new teachers. Although the coefficient is estimated at 0.23, the confidence interval is very large, ranging from an impact on retention estimated at between negative -0.45, to positive +0.92. This finding suggests wide variation in the predicted impact of incentives on new shortage field teachers, as well as an inability to distinguish the estimate from zero.

Table 15

Retention in School Outcome, New Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.233
			(0.341)
AFT*New Shortfield	0.059	**0.254****	
	(0.087)	(0.051)	
AFT	-0.018	-0.001	-0.018
	(0.015)	(0.001)	(0.015)
AFT*Shortfield	0.035	0.001	0.035
	(0.021)	(0.004)	(0.021)
AFT*New	-0.025	0.0002	-0.025
	(0.047)	(0.001)	(0.047)
New Shortage-field	-0.043	**0.127****	-0.076
	(0.034)	(0.015)	(0.064)
Constant	0.767**	-0.028*	0.771**
	(0.054)	(0.007)	(0.055)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

To verify that the effect is driven by more experienced teachers, I also analyzed the sample of teachers who have more than three years of experience in the non-Midwestern region (Table 16). I again changed the interaction term in the model to reflect the different sample of interest:

$$D_{ik} = Z_k(AFT * Exp * SF) + X_k(AFT) + X_i(Exp) + X_i(SF)$$

$$+X_i(Exp * SF) + X_{ik}(Exp * AFT) + X_{ik}(SF * AFT) + X_{ijk}B + S_{state} + S_{survey} + e_{ijk}$$

In contrast to the results found for the new teachers, in the first stage the F-statistic is strong (21.3), and these teachers are also predicted to have a 25 percent greater chance of teaching in a district that offers incentives (p<.01). However, in the second stage, I do not have a significant result. Although the coefficient estimate is negative the confidence interval is very large, ranging from -0.90 to 0.43. This range suggests that the data in the model are simply too noisy to detect an effect. Moreover, the lack of significant findings for both the new and experienced teacher estimates suggest that the positive results above are driven by a combination of experience levels that cannot be parsed from the available sample.

Table 16

Retention in School Outcome, Experienced Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			-0.238
			(0.333)
AFT*Exp Shortfield	-0.059	**0.248****	
	(0.087)	(0.021)	
AFT	-0.043	**-0.023****	-0.048
	(0.045)	(0.003)	(0.051)
AFT*Shortfield	0.095	0.001	0.095
	(0.085)	(0.003)	(0.084)
AFT*Experienced	0.025	**0.004***	0.026
	(0.047)	(0.002)	(0.048)
Experienced Shortage-field	0.043	**0.112****	0.069
	(0.034)	(0.006)	(0.059)
Constant	0.767**	-0.049**	0.755**
	(0.054)	(0.017)	(0.057)
Observations	16,070	16,070	16,070

~ significant at 10%; *significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Analytic Results: Retention in Teaching

Results for the retained in teaching outcome were not statistically significant for any model specification. Although these models were specified using a larger sample of retained teachers, the retention outcome allowed for more variation than the previous models. The range of potential outcomes for teachers was expanded to include teachers who moved to another school in the same district, to a different school in another district, to a charter, or to a private school. At the same time, the alternative category- the leavers- was compressed to include only those who left teaching for other careers, for family obligations, for retirement, or who died.

As expected, in the first set of analyses for the retained in teaching outcome, utilizing the full sample (Table 17), the first stage results are identical to the model above, in which a shortage-field teacher teaching in an AFT affiliated district has an almost 17 percent ($p<.01$) greater likelihood of being offered an incentive than comparable teachers. However, the results in the second stage are not statistically significant. In addition, the coefficient on the incentive variable is negative, close to zero, and has a wide confidence interval. These findings suggests that, in this sample, shortage-field incentives may be less effective for teachers who are contemplating non-teaching options, in comparison to teachers who may want to stay in teaching, but who without and incentive would otherwise would look for a different teaching position, either in the same districts, or in another district.

Table 17

Retention in Teaching Outcome, Full Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			-0.006
			(0.100)
AFT*Shortfield	-0.001	**0.172****	
	(0.017)	(0.014)	
AFT	-0.012	**-0.010****	-0.012
	(0.009)	(0.002)	(0.009)
Shortage-field	0.006	**0.122****	0.006
	(0.005)	(0.005)	(0.015)
Constant	0.900**	-0.053**	0.899**
	(0.037)	(0.014)	(0.037)
Observations	23,110	23,110	23,110

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clustered at the district level.

The result for the non-teaching outcome in the sample that does not include the Midwest region is also not statistically significant (Table 18). In this sample, again the first stage results are identical to those presented above. In this case, however, the coefficient on the variable of interest is positive, and the confidence interval is smaller than the estimate for the full sample, indicating that the elimination of the Midwest region does produce a better specified outcome. However, the result nevertheless suggests that shortage-field incentives, at least in the magnitudes the have been applied in to the teachers in this sample, are still unlikely to impact the decisions of those considering non-teaching alternatives.

Table 18

Retention in Teaching Outcome, Non-Midwest Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.037
			(0.061)
AFT*Shortfield	0.009	**0.251****	
	(0.015)	(0.019)	
AFT	-0.013	**-0.021****	-0.013
	(0.011)	(0.002)	(0.011)
Shortage-field	0.008	**0.116****	0.004
	(0.006)	(0.006)	(0.011)
Constant	0.879**	-0.078**	0.894**
	(0.048)	(0.021)	(0.048)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

The impact of shortage-field incentives on the retained-in-teaching outcome is also statistically indistinguishable from zero in the models that are specified separately for both math and science teachers, and teachers of special education and bilingual teachers. However, in the model for math and science teachers the coefficient on the incentive variable in the second stage is similar in size to the result estimated for the retained in school outcome, but is not significant. By contract, the coefficient of interest in the model for special education and bilingual teachers (Table 20) is both much smaller than the previous model, and negative. The finding for STEM teachers helps confirm that these are the teachers whose retention behavior may be most impacted by the offering of shortage-field incentives, but the sample used in the retained in teaching outcome may be too small or otherwise imprecise to identify a statistically significant finding.

Table 19

Retention in Teaching Outcome, STEM Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.162
			(0.112)
AFT*Stem	0.031	**0.190****	
	(0.016)	(0.026)	
AFT	-0.015	**-0.009****	-0.013
	(0.011)	(0.002)	(0.011)
Stem teacher	0.007	**0.116****	-0.012
	(0.007)	(0.008)	(0.016)
Constant	0.891**	-0.039**	0.897**
	(0.048)	(0.014)	(0.048)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Table 20

Retention in Teaching, Special Education and Bilingual Teachers in Non-Midwestern Regions

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			-0.047
			(0.077)
AFT*Special ed. Bilingual	0.014	**0.306****	
	(0.021)	(0.028)	
AFT	-0.009	**-0.012****	-0.009
	(0.010)	(0.002)	(0.010)
Special ed. Bilingual	0.008	**0.113****	0.013
	(0.009)	(0.009)	(0.015)
Constant	0.891**	-0.037**	0.890**
	(0.048)	(0.012)	(0.048)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

The results for the retained-in-teaching outcome are also not statistically significant for models that are specified separately for new teacher and experienced teachers (Tables 21 and 22). The model for new teachers here also suffers from the same weak instrument problem discussed in the model that utilized the retained-in-school outcome, and should also be interpreted with caution. The coefficient for the incentive treatment for experienced teachers is positive, but also statistically indistinguishable from zero.

Table 21

Retention in Teaching Outcome, Non-Midwest Sample, New Teachers

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			-0.100
			(0.276)
AFT*New Shortfield	-0.025	**0.254****	
	(0.067)	(0.052)	
AFT	-0.008	**-0.001****	-0.008
	(0.011)	(0.0009)	(0.011)
AFT*Shortfield	0.012	0.0007	0.012
	(0.014)	(0.0004)	(0.014)
AFT*New	--0.030	0.0002	-0.030
	(0.036)	(0.0007)	(0.036)
New Shortage-field	0.006	**0.127****	0.018
	(0.019)	(0.015)	(0.044)
Constant	0.891**	-0.028**	0.887**
	(0.048)	(0.007)	(0.049)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Table 22

Retention in Teaching Outcome, Non-Midwest Sample, Experienced Teachers

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.102
			(0.285)
AFT*Experienced Shortfield	0.025	**0.248****	
	(0.067)	(0.021)	
AFT	-0.038	**-0.023****	-0.035
	(0.035)	(0.003)	(0.039)
AFT*Shortfield	-0.013	0.001	-0.013
	(0.066)	(0.004)	(0.066)
AFT*Experienced Teacher	0.030	**0.004***	0.029
	(0.036)	(0.002)	(0.037)
Experienced Shortage-field	-0.006	**0.112****	-0.017
	(0.019)	(0.006)	(0.041)
Constant	0.891**	-0.049**	0.895**
	(0.048)	(0.017)	(0.049)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Analytic Results: Recruitment

My next set of analyses examines two teacher quality outcomes, including the Barron's college plus ranking index, where each 1-unit increase represents a move up the Barron's ranking scale from non-competitive to most competitive, as well as a dichotomous outcome indicating whether or not a teacher has an infield BA. Although the first stage of this model should be identical to the results above, the sample size for these models is reduced by about 2,000 teachers for whom there is missing data on undergraduate institution. Therefore, in the first stage of this model, using the full sample of data (Table 23), the shortage-field teachers in AFT districts have a 15 percent greater likelihood (p<.05) of teaching in a district that provides shortage-field incentives. In addition, I do not find any statistically significant impacts of incentives on the academic quality of teachers recruited.

Table 23:

Barron's Index Outcome, Full Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.121
			(0.607)
AFT*Shortfield	0.019	**0.155***	
	(0.093)	(0.015)	
AFT	0.036	**-0.011****	0.037
	(0.059)	(0.009)	(0.056)
Shortage-field	-0.128**	0.124	-0.143
	(0.031)	(0.005)	(0.092)
Constant	4.69**	-0.065*	4.69**
	(0.183)	(0.016)	(0.185)
Observations	20,300	20,300	20,300

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Once I focus the sample to include only the non-Midwestern states (Table 24) that face significant recruitment challenges, the likelihood of shortage-field teachers teaching in AFT districts that provide incentives increases to 22 percent ($p<.01$). However, the impact of incentives on quality of teachers recruited remains insignificant in this reduced sample. Furthermore, in both models, the confidence interval on the coefficient of interest ranges from -0.96 to +1.2 in the full model, with a similar spread in the non-Midwestern sample. These results may be due to the fact that in both the full and non-Midwestern samples, there is a lack of variation in teacher undergraduate rankings. In both samples, 52 percent of all teachers attended a college with a “competitive” ranking, which is the 7th ranking out of nine, and coded as 3 in the variable rank. Also in both samples, 18 percent of teachers attended colleges with the second from lowest “less competitive” ranking. Sixteen percent of teachers attended colleges with “very competitive” rankings, which is the 5th ranking. Finally, 6 percent of teachers attended the lowest ranked colleges, while 7 percent attended colleges with the top 4 rankings combined. These figures hold for the samples of both shortage and non-shortage-field teachers. This lack of variation in the college ranking of teachers throughout the sample, makes it much less likely that shortage-field incentives can be identified as increasing the quality of teachers recruited into teaching. It also suggests that the issue of teacher quality and the teacher preparation pipeline may be a larger systemic issue that may need to be addressed through multiple means, of which incentives may only represent one part.

Table 24

Barron's Rank Outcome, Non-Midwestern Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.229
			(0.592)
AFT*Shortfield	0.051	**0.219****	
	(0.125)	(0.019)	
AFT	0.104	**-0.021****	0.108
	(0.077)	(0.003)	(0.073)
Shortage-field	**-0.134****	**0.118****	-0.161
	(0.039)	(0.006)	(0.092)
Constant	4.69**	-0.097**	4.72**
	(0.229)	(0.021)	(0.235)
Observations	14,100	14,100	14,100

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Since quality of undergraduate institution is an arguably distal measure of teacher quality, I also test whether districts that offer shortage-field incentives are more likely to recruit teachers with a BA in the subject they are teaching. Although I again find no statistically significant impacts of incentives on this measure of teacher quality (Table 23), it is interesting to note that the confidence interval on the coefficient of interest is much larger in the model that uses the non-Midwestern sample (Table 24) than in the full sample. In the non-Midwest sample, the range is from -0.23 to + 0.22, but in the full sample to range is from -0.092 to +0.604. Since the interval decreases when the sample size decreases, these differences suggest that when the Midwest is included, the overall likelihood that a teacher is infield is greater than in other parts of the country, lending further evidence to the idea that teacher labor markets in that part of the country are different than in other regions. Moreover, the regional disparities may help explain why the results for the retention in school outcomes in my first research question do not hold when those states are included in the model. Specifically, it is possible that teachers with undergraduate degrees in STEM fields may have fewer other employment options in the Midwest and are therefore more likely to be found in teaching there than elsewhere, even without the extra financial inducements provided by shortage-field incentives.

Table 25:

In-field Teacher Outcome, Full Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			0.257
			(0.177)
AFT*Shortfield	0.044	**0.172****	
	(0.026)	(0.014)	
AFT	**-0.015***	**-0.010***	-0.012
	(0.006)	(0.002)	(0.007)
Shortage-field	**0.242****	**0.122****	0.211
	(0.007)	(0.005)	(0.026)
Constant	0.095**	-0.053	0.108**
	(0.027)	(0.015)	(0.030)
Observations	23,100	23,100	23,100

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Table 26

Infield Outcome, Non-Midwestern Sample

	Reduced Form	1st stage	2nd stage
Shortage-field Incentive			-0.007
			(0.115)
AFT*Shortfield	0.002	**0.251****	
	(0.029)	(0.019)	
AFT	-0.007	**-0.021****	-0.007
	(0.006)	(0.003)	(0.006)
Shortage-field	**0.252****	**0.116****	**0.252****
	(0.009)	(0.006)	(0.019)
Constant	0.077*	-0.079	0.077*
	(0.035)	(0.019)	(0.037)
Observations	16,070	16,070	16,070

*significant at 5%; ** significant at 1%

Control variables included in the model: city and rural indicators, school percent Black and Hispanic, school percent receiving free/reduced price lunch, school percent ELL and IEP, school enrollment (logged), charter school indicator, total teaching experience, middle and high school indicators, new teacher indicators, teacher age, teacher gender, teacher has a STEM BA, fully certified teacher indicator, survey year and state fixed effects. Standard errors clusters at the district level.

Summary

The lack of significant findings for the academic background outcome, combined with the lack of significant findings about in-field status, particularly in the non-Midwestern states, suggests that although shortage-field incentives may be effective for stemming attrition of teachers who have already selected into the field, they may be ineffective, at least as they are currently administered, for enticing those who are not already predisposed to teaching, or those who have alternative options, to consider entering the field.

In this chapter, I conducted analyses examining the impacts of shortage-field incentives on the retention and recruitment of teachers. I utilized an instrumental variables, difference-in-differences approach and found that shortage field incentives increase the probability that a teacher will remain in their school after 1 years by approximately 17 percent, in all states that require collective bargaining, except those located in the Midwest. I also found that incentives were more effective for increasing the retention of more experienced teachers, but had little impact on attrition of those with less than 3 years of teaching experience. The findings for new teachers however, may have been impacted by sample size issues.

I also failed to find an impact of incentives on changes in the quality of teachers teaching in districts that offered shortage-field incentives. Specifically, shortage field teachers in incentive districts had similar academic backgrounds to their counterparts in districts that did not offer incentives, and were also as likely to be teaching infield. Lack of variation in academic backgrounds of all teachers likely undermined the ability to detect differences between teachers on this measure.

Chapter 5

Descriptive Results-RQ3

Because the analysis sample changes from just over 26,000 teachers in research questions one and two, to almost 81,000 teachers in research question three, I also present descriptive results for the expanded sample utilized in the fixed effects analysis.

Table 26 provides characteristics of teachers and schools in districts, by the frequency of the incentive availability. Largely, these findings indicate that the frequency with which a district provides incentives is likely endogenous to other district characteristics that are known to be associated with teacher shortages. In particular, the more frequently a district offers shortage-field incentives, the more likely it is to be urban, and the less likely to be rural. Districts in the South and West are also more likely to provide incentives on a more frequent basis, while the Midwest is less likely to provide incentives with each wave of the survey. Districts that more frequently offer incentives also have higher mean percentages of Black and Hispanic students, as well as more language minority students. Districts that provide incentives also more frequently have lower average rates of certified teachers. However, mean rates of teachers with bachelor's degrees in math and science subjects remains relatively static, regardless of how frequently a district offers shortage-field incentives. Finally, although teachers' unions have often been cited as an impediment to shortage-field incentive implementation, the results for collective bargaining are mixed. For example, the percentage of districts with collectively bargained contracts that never offer incentives, and those that offered it at least once are virtually the same. However, more than half the districts that offered incentives twice also had collectively bargained contracts. At the same time, only 30 percent of districts that provided the incentives every year had collective bargaining. These results suggest that there may be more variation in the flexibility of local union contracts than has been previously suggested.

Table 27

Descriptive Statistics of District Characteristics, by Frequency of Incentive Availability

Variable Name (District level)	Incentive never offered (n=2,140)	Incentive offered once (n=710)	Incentive offered twice (n=230)	Incentive offered three times (n=60)
Total	0.56	0.26	0.13	0.05
City	0.21	0.30	0.43	0.46
Rural	0.37	0.26	0.14	0.13
South	0.37	0.42	0.40	0.43
Northeast	0.14	0.07	0.10	NA
Midwest	0.27	0.18	0.07	0.06
West	0.22	0.32	0.42	0.52
Collective bargaining	0.47	0.46	0.52	0.31
Charter	0.01	0.02	0.02	0.00
% Black	0.13	0.18	0.21	0.20
% Hispanic	0.07	0.12	0.23	0.34
% IEP	0.14	0.14	0.15	0.12
% LEP	0.04	0.06	0.10	0.16
% Lunch eligible	0.38	0.40	0.50	0.53
Elementary	0.27	0.29	0.32	0.32
Middle	0.14	0.14	0.13	0.14
High school	0.50	0.50	0.46	0.49
Combined grade	0.09	0.07	0.09	0.05
New teacher	0.15	0.17	0.19	0.19
Shortage teacher	0.23	0.23	0.24	0.22
Fully certified	0.90	0.88	0.84	0.84
Female	0.68	0.69	0.69	0.69
Male	0.32	0.31	0.31	0.31
Stem BA	0.13	0.12	0.13	0.12

I also examined comparisons between shortage and non-shortage field teachers in the fixed effects analysis sample (Table 27). These descriptive findings suggest that these teachers teach in similar school settings. There is little difference in the socio-economic characteristics between the students taught by shortage and non-shortage field teachers, although shortage-field teachers as a whole are slight more likely to teach in urban areas, and slightly less likely to teach in the West. Shortage-field teachers are also less likely to teach in elementary school, and more likely to teach in high school. In addition, these teachers are more likely to be represented by collective bargaining units, and slightly more likely to be male. Shortage-field teachers are also less likely to be fully certified, which is unsurprising since districts are presumably more likely to hire an uncertified teacher in a field where there is a shortage.

Table 28

Characteristics of Shortage and Non-shortage Field Teachers, Fixed Effects sample

Variable Name (teacher level)	Shortage Field (n=18,590)	Non Shortage Field (n=62,360)
Total	0.23	0.77
City	0.30	0.27
Rural	0.29	0.30
South	0.39	0.39
Northeast	0.12	0.11
Midwest	0.22	0.21
West	0.27	0.29
Collective bargaining	0.50	0.46
Charter	0.01	0.01
% Black	0.16	0.16
% Hispanic	0.12	0.11
% IEP	0.15	0.14
% LEP	0.06	0.06
% Lunch eligible	0.41	0.41
New teacher	0.16	0.16
Elementary	0.21	0.30
Middle	0.17	0.13
High school	0.52	0.49
Combined grade	0.10	0.08
Fully certified	0.86	0.89
Female	0.67	0.69
Male	0.33	0.31
Stem BA	0.18	0.11

Mean differences significant at p<.01. Mean differences not significant for rural, lunch eligibility, and new teacher.

These descriptive findings about districts and teachers suggest that the frequency with which incentives are offered appears to be endogenously related to other district characteristics that are known to negatively influence teacher retention. Therefore, to mitigate the potential problem caused by the endogeneity of incentive frequency, and to create an analysis sample with plausible counterfactual comparison groups, I created a sub-sample for analysis that includes only teachers who taught in the same districts, either in the survey year prior to the offering of an incentive, or in the survey year during the time when incentives were offered. By only including teachers who taught in the same districts, but who responded to surveys at different times, I am able to approximate treatment and control groups in which pre-treatment differences in district characteristics are minimized because the districts are the same, except for the differences in time.

In order to understand if teachers who taught in districts before and after incentives are actually comparable, I also analyzed descriptive statistics of this subgroup (Table 26). In this analysis, my sample size is reduced by more than half, to just under 32,000 teachers, but the groups have similar distributions across urban and rural areas, as well as across the South and Midwest. However, more teachers in districts in the Northeast were exposed to incentive policies after they were added. Additionally, the schools where teachers taught after incentives were added had higher percentages of Hispanic and ELL students, although percentages of Black students were similar. In addition, there were more students eligible for the lunch program after incentives were added. The findings for Hispanic students likely reflect rapidly changing demographics and population growth of Hispanics between 1999 and 2007.

In addition, fewer teachers who taught in districts after incentives were offered were either fully certified or taught under collective bargaining agreements than those who taught in the same districts previously. The differences in union status may be due to state level changes in collective bargaining laws during this time period, such as those previously described in New Mexico and Missouri.

Table 29

Characteristics of Teachers Who Taught in Districts Before and After Incentives Were Introduced

	After (n=16,330)	Before (n=15,590)
Total	0.55	0.45
City	0.36	0.32
Rural	0.21	0.24
South	0.41	0.41
Northeast	0.09	0.07
Midwest	0.13	0.16
West	0.36	0.35
Collective bargaining	0.42	0.53
Charter	0.01	0.02
% Black	0.19	0.20
% Hispanic	0.18	0.13
% IEP	0.14	0.15
% LEP	0.08	0.06
% Lunch eligible	0.44	0.42
New teacher	0.19	0.17
Shortage-field teacher	0.23	0.24
Elementary	0.30	0.30
Middle	0.13	0.14
High school	0.49	0.49
Combined grade	0.08	0.07
Fully certified	0.85	0.87
Female	0.31	0.31
Male	0.69	0.69
Stem BA	0.13	0.12

T-test statistics. Mean differences significant at the 0.01 percent level. Mean differences not significant for South, West, Charter, % Black, %IEP, Shortage-field, Elementary, Middle, High School, Gender, Stem BA.

I also conducted descriptive analysis of the outcomes, utilizing the full fixed effects analysis sample. Table 29 provides results for teachers who were and were not retained in their school for one year. In this sample, a total of 16 percent of teachers were not retained in their school; this is compared to 14 percent in the analysis sample utilized in the IV analysis in research questions one and two. In this sample, teachers who are and are not retained are equally likely to be teaching in urban areas. However, those who are retained are more likely to be found in the Midwest, and to teach in schools that have collective bargaining for teachers. Those who were not retained were more likely to leave schools with higher average percentages of Black students, and students eligible for the subsidized lunch program. However, the schools of those who did and did not remain had similar populations of Hispanic, LEP and IEP students. Elementary and middle school teachers were more likely to leave their schools, compared to high school teachers. Finally, teachers who were not retained were much more likely to be new teachers, as well as uncertified.

Table 30

Characteristics of Teachers Who Did and Did Not Remain in their School After 1 Year

Variable Name (teacher level)	Retained in school (n=67,900)	Not retained in school (n=13,050)
Total	0.84	0.16
City	0.28	0.28
Rural	0.30	0.29
South	0.38	0.41
Northeast	0.11	0.10
Midwest	0.22	0.19
West	0.29	0.30
Collective bargaining	0.47	0.44
Charter	0.01	0.01
% Black	0.15	0.18
% Hispanic	0.12	0.13
% IEP	0.14	0.14
% LEP	0.06	0.07
% Lunch eligible	0.40	0.43
New teacher	0.15	0.25
Shortage-field teacher	0.23	0.25
Elementary	0.28	0.30
Middle	0.13	0.15
High school	0.50	0.46
Combined grade	0.08	0.09
Fully certified	0.89	0.81
Female	0.69	0.68
Male	0.31	0.31
Stem BA	0.13	0.13

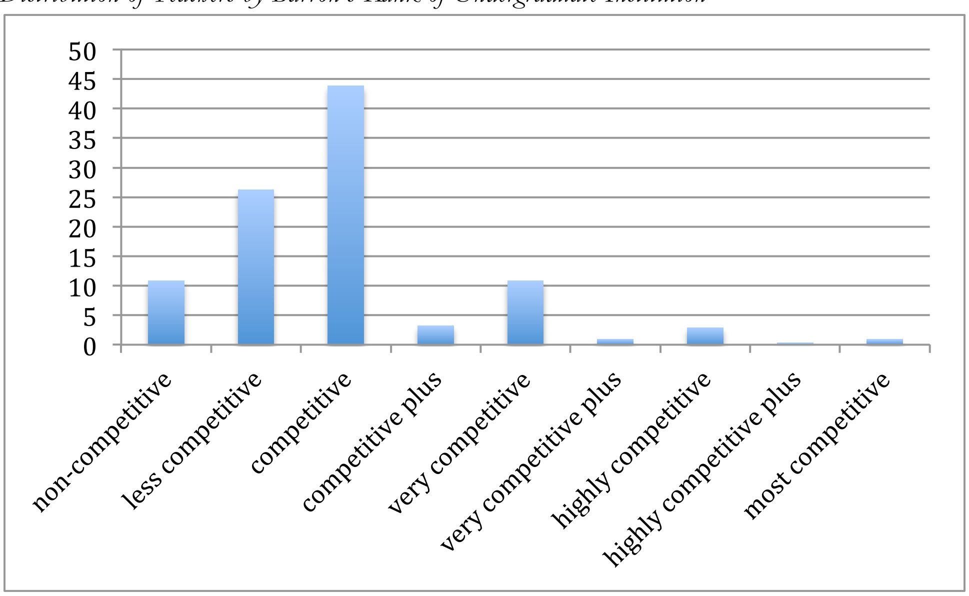
Mean differences significant at $p<.01$. Mean differences not significant for city, rural, male, female, STEM BA

Interestingly, descriptive findings for the Barron's undergraduate competitiveness rankings outcome show greater variation in the distribution of teachers in the larger fixed effects analysis sample than was found in the sample for the first two research questions

(Figure 2). In the previous sample, over half the teachers attended colleges ranked either competitive or competitive plus. In this sample, that combined number drops to less than 50 percent. In the previous sample, teachers with backgrounds from colleges ranked in the non-competitive and less competitive categories comprised less than 25 percent of the sample. In this case, teachers with those backgrounds make up approximately 37 percent of the sample. By contrast, in the previous analysis sample, teachers from colleges ranked very competitive or very competitive plus, made up over 15 percent of the sample, where in the fixed effects sample, teachers with these backgrounds are about 12 percent. All of these findings indicate that the previous sample, which was smaller and arguably less nationally representative, consisted of teachers who had stronger academic backgrounds than those found in the larger population of teachers, and that the results of those analyses are likely only generalizable to states with populations of teachers more similar to the teachers represented by that sample.

Figure 2

Distribution of Teachers by Barron's Rank of Undergraduate Institution



Despite the differences in the distribution of teacher academic backgrounds between the fixed effects analysis sample and the previous analysis sample, differences in the characteristics of teachers by undergraduate academic rank show similar patterns to those presented previously (Table 31). For example, as in the earlier analysis, new and charter school teachers have academic backgrounds that were rated more highly than other teachers. Teachers in the Northeast graduated from colleges that have an average Barron's rank that is almost a full point higher than teachers in the other regions of the country, and teachers who are fully certified graduated from colleges with lower academic rankings. Finally, teachers in districts with collective bargaining rights graduated from colleges with slightly higher college rankings than other teachers. Given that the previous sample only contained teachers with full bargaining rights, the differences between this sample and the last one is likely a reflection of this difference.

Table 31

Characteristic of Teachers by Barron's Rank of Undergraduate Institution, Fixed Effects Sample

Variable Name (Teacher Level)	Variable=1	Variable=0
District offers shortage-field incentives	2.9	3.1
City	3.0	3.0
Rural	2.7	3.0
South	2.8	3.1
Northeast	3.8	2.9
Midwest	2.8	3.0
West	3.0	3.0
Collective bargaining contract	3.1	2.9
Charter school teacher	3.6	3.0
New teacher	3.5	2.8
Shortage-field teacher	2.9	2.9
Elementary	2.9	3.0
Middle	3.0	3.0
High School	3.0	2.9
Fully certified	2.9	3.5
Female teacher	3.0	3.0
Has STEM BA	3.2	3.0
Infield	3.0	2.9

Finally, I conducted analysis of the descriptive differences between shortage field teachers who were teaching both in and out-of-field in this sample (Table 32). Here, only 25 percent of shortage field teachers are teaching infield. However, given that over 70 percent of teachers with undergraduate degrees in STEM fields were teaching infield, and only 1 percent were teaching out-of-field, it appears that most of the out-of-field teaching was done by special and bilingual education teachers. Infield teachers taught more frequently at schools with lower mean populations of Hispanic, ELL, IEP and lunch eligible students, although they are also found at schools with slightly more Black students. Infield teachers were more likely to teach at the high school level, and were also more likely to be both male, and fully certified. In addition, while 63 percent of infield teachers teach in districts with collective bargaining agreements, only 46 percent of out-of-field teachers worked under such contracts. This finding also further illuminates how the sample utilized for the previous analysis differs from a more nationally representative sample.

Table 32

Characteristics of Infield and Out-of-field Shortage-field Teachers

Name of Variable (teacher level)	Infield (n=4,560)	Out-of-field (n=14,030)
Total	0.25	0.75
City	0.30	0.30
Rural	0.31	0.28
South	0.39	0.39
Northeast	0.12	0.12
Midwest	0.22	0.22
West	0.27	0.28
Collective bargaining	0.63	0.46
Charter	0.00	0.01
% Black	0.17	0.16
% Hispanic	0.10	0.13
% IEP	0.14	0.16
% LEP	0.04	0.06
% Lunch eligible	0.37	0.42
New teacher	0.17	0.16
Elementary	0.11	0.24
Middle	0.12	0.18
High school	0.68	0.47
Combined grade	0.08	0.11
Fully certified	0.88	0.86
Female	0.63	0.69
Male	0.37	0.31
Stem BA	0.71	0.01

Mean differences significant at $p<.01$. Mean differences not significant for city, all regions, and new teacher

Empirical Results: Research Question 3

For the first fixed effects analysis with the retained in school outcome, I utilized the full analysis sample, and included control variables for districts that never offered an incentive, or offered one every year. The coefficient of interest is for the interaction indicating shortage field teachers who taught in districts the year the incentive was offered. In the naive model (Table 33), I do not include any covariates, but have 3,130 district fixed effects. In the second model, I include teacher and school control variables. In this model, due to missing data on some covariates, I loose about 4,000 teachers, but still maintain a sample of 77,000 teachers. Based on the regional analysis presented in the last chapter, I also conduct a sub analysis that eliminates districts located in the Mid-West. The coefficient of interest in all models is estimated at about 0.01, and is not significant in any specification. In addition, the confidence interval in all models is estimated between approximately -0.004 and 0.02, indicating that the effect is very small, at best.

Table 33

Retained in School Outcome, Full Sample Fixed-effects Model

	(1) No Controls	(2) With Controls	(3) Non-Midwest Sample
Shortage Field*Incentive	0.009	0.010	0.008
	(0.007)	(0.007)	(0.008)
Shortage Field Teacher	-0.021**	-0.022**	-0.019**
	(0.004)	(0.004)	(0.004)
Constant	0.853**	0.800**	0.788**
	(0.003)	(0.008)	(0.009)
Observations	80,950	76,930	60,530
Number of district fixed effects	3,130	3,110	2,250
R-squared	0.06	0.07	0.07

*significant at 5%; ** significant at 1%; Standard errors in parentheses

Models 2 and 3 contain controls for teacher gender, teacher STEM BA, teacher fully certified, New teacher, school percent Black, Hispanic, lunch eligible, IEP, LEP. Other controls include school level, charter school flag, and collective bargaining contract. Survey year fixed effects are also included.

I also conducted analysis on the reduced sample in which I included only teachers in those districts that reported a change in shortage-field incentive policy between survey years (Table 33). In these models, the variable of interest is the interaction between being a shortage-field teacher, and teaching in the district the year the incentive was offered. Again, in these models the effect is estimate near zero, and is not statistically significant. Moreover, the estimate, nor the confidence interval changes much with either the inclusion of control variables, or with the elimination of districts in the mid-west.

Table 34

Retained in School Outcome, Districts Before and After Incentive, Fixed-effects Model

	(1) No Controls	(2) With Controls	(3) Non-Midwest Sample
Incentive	0.010	0.006	0.008
	(0.010)	(0.010)	(0.011)
Shortage Field	-0.022**	-0.018*	-0.016*
	(0.007)	(0.007)	(0.008)
Incentives*SF	-0.001	0.002	-0.001
	(0.005)	(0.006)	(0.006)
Constant	0.844**	0.770**	0.765**
	(0.005)	(0.012)	(0.013)
Observations	31,910	30,520	26,009
Number of district fixed effects	940	930	730
R-squared	0.05	0.06	0.06

* significant at 5%; ** significant at 1%; Standard errors in parentheses

Models 2 and 3 contain controls for teacher gender, teacher STEM BA, teacher fully certified, New teacher, school percent Black, Hispanic, lunch eligible, IEP, LEP. Other controls include school level, charter school flag, and collective bargaining contract. Survey year fixed-effects are included.

In the next set of analysis, I examined the impact of adding an incentive on the Barron's rank of teacher's undergraduate institution, for teachers surveyed in districts before and after incentive policies were reported (Table 34). In these models, the Barron's outcome is standardized so that coefficients can be interpreted as standard deviation units. Although these models utilize the same sample as the second set of models estimated above, the sample sizes are reduced due to missing undergraduate institutions for some teacher. Note however that the number of district fixed effects remains constant.

For these analyses, I estimated two models, one with and one without control variables. The coefficient of interest is not statistically significant in either model, however. In the naïve model, the confidence interval ranges from -0.07 to 0.03, and the full model is spans from -0.04 to 0.07 which indicates that in both models the data are too noisy to approximate whether the effect is positive, negative, or actually zero. In these models it is interesting to note that model 1, which includes only the shortage field and incentive indicators as well as survey year, and district fixed effects, 19 percent of the variation in the outcome is explained by these factors, but that adding in all the teacher and school level covariates increases the R-squared by only 4 percentage points. This finding indicates that the explained variation in teacher quality, as measured by academic background, is greatest between districts, as opposed to within districts.

Table 35

Barron's Rank Outcome, Districts Before and After Incentive, Fixed-effects Model

	(1) No Controls	(2) With Controls
Incentives*SF	-0.020	0.012
	(0.027)	(0.027)
Shortage Field	**-0.040***	**-0.076****
	(0.019)	(0.019)
Incentives	**0.034***	0.020
	(0.015)	(0.015)
Constant	-0.009	-0.067*
	(0.012)	(0.034)
Observations	27,420	26,230
Number of district fixed effects	940	930
R-squared	0.19	0.23

* significant at 5%; ** significant at 1%; Standard errors in parentheses

Model 2 contain controls for teacher gender, teacher STEM BA, teacher fully certified, New teacher, school percent Black, Hispanic, lunch eligible, IEP, LEP. Other controls include school level, charter school flag, and collective bargaining contract. Survey year fixed effects are included.

In the final analyses utilizing the sample of teachers who taught in districts before and after incentive policies changed, I also examined the impact of incentives on whether shortage-field teachers were teaching infield or out-of-field. Because this further restriction on the data limited the analysis to only shortage-field teachers, these analyses are conducted on a limited sample of just 7,500 teachers. In addition, while they contain district fixed effects, the variable of interest is the single indicator for whether or not teachers taught in districts that offered shortage-field incentives. In both the naïve model, and the model containing all school and teacher covariates, the estimate is 0.013, and is not statistically significant. In addition, the confidence interval in both models ranges from -0.007 to 0.03, indicating that the true value may be positive, but is likely very small. In addition, here the R-squared in the first model is 0.41 meaning that 40 percent of the variation in whether or not shortage field teachers are teaching infield or not is explained by cross district variation, as opposed to within district variation. This finding is supported by the fact that the inclusion of school and teacher level variable increases the R-squared only slightly.

Table 36

Infield Outcome, Districts Before and After Incentive, Fixed-effects Model

	(1) No Controls	(2) With Controls
Incentives	0.013	0.013
	(0.010)	(0.011)
Constant	0.504**	0.378**
	(0.008)	(0.024)
Observations	7,520	7,189
Number of district fixed effects	940	930
R-squared	0.41	0.43

* significant at 5%; ** significant at 1%; Standard errors in parentheses
Model 2 contains controls for teacher gender, teacher STEM BA, teacher fully certified, New teacher, school percent Black, Hispanic, lunch eligible, IEP, LEP. Other controls include school level, charter school flag, and collective bargaining contract. Survey year fixed effects are included.

Summary

In this sub-analysis, I conducted fixed effects analysis examining changes in teacher outcomes over time, in districts that experienced changes in incentive policies from between rounds of the SASS. Although I was unable to detect statistically significant results, the study was limited to finding outcomes over a relatively short period of time. Specifically, all teachers in these analysis were only tracked for one year, and only a small minority of districts in this analysis were surveyed more than twice, meaning that most districts only accounted for one year prior, and one year after a change. Because there were only a relatively small number of teachers surveyed within districts each year, detecting a statistically significant result within a single district over one year may have been too difficult. In addition teachers included in these analyses were different from those included in the analysis sample in the last chapter. Specifically, this sample included teachers who did have collective bargaining rights, which likely impact both recruitment and retention outcomes in significant ways.

Chapter 6

Conclusion

In this dissertation, I utilized three waves of the national Schools and Staffing Survey to investigate the efficacy of financial incentives for increasing retention and improving recruitment of shortage-field teachers. In the first two of three research questions, I was able to identify a causal pathway between the provision of incentives and teacher outcomes by utilizing an instrumental variable difference-in-differences model. By finding an external instrument that is related to the policy or treatment of interest, but that is related to the outcome only via the treatment, this method allows for the approximation of comparable treatment and control groups, and is recognized as one of the stronger quasi-experimental methods available for obtaining causal findings from observational data. In this case, I exploited the exogenous variation in differing national teacher union policies on the provision of financial incentives for teachers in shortage-fields. Specifically, I was able to create comparable groups between teachers who worked in AFT affiliated districts and those who worked in districts affiliated with the NEA. In this case, the counterfactual was plausible because there was a strong link between teaching in an AFT district and the likelihood of receiving an incentive, but no readily identifiable link between union affiliation and actual teacher outcomes. Moreover, by combining the IV with the difference between shortage and non-shortage field teachers in the same districts, I was able to further improve the plausibility of the findings because non-shortage field teachers in AFT districts were likely not impacted by other potential effects of AFT affiliation.

Because the identification strategy relied upon district affiliation with one of the two major national teachers unions, the analysis sample utilized for research questions 1 and 2 was strictly limited to districts that were governed by collectively bargained contracts, as well as those that could be positively identified as being represented by either the AFT or the NEA. Therefore, the results of the analysis are only generalizable to the population of teachers and districts that are found in states that have collective bargaining policies for public sector employees.

Results of these analyses indicate that shortage-field incentives have a positive causal relationship to increasing the retention of shortage-field teachers in schools, for at least an additional year. In addition, because the effects were found only for in-school retention, and not for retention in teaching, it is likely that these kinds of incentives work best for keeping teachers from leaving some schools for better positions in other schools, rather than for convincing those who are not committed to teaching to remain in the profession.

Another aspect of the finding for the first research question was that the impacts were strongest in regions that are experiencing elevated levels of shortages. This trend is likely because in these areas, greater levels of shortages means there are more opportunities for experienced teachers to move schools, as positions in more desirable teaching environments become available more frequently. In addition, in areas that have experienced overall worse employment conditions and lower salaries, like the Midwest, there are potentially fewer options for teachers to move positions, which is likely why there was no incentive impact in that region.

Overall, however, the findings for the analysis for the first research question suggest that providing financial incentives to teachers in shortage-fields is an effective tool for keeping existing teachers in their schools.

I also utilized the same IV difference-in-differences framework to examine the impact of shortage-field incentives on two recruitment outcomes. Specifically, I analyzed whether districts that offered incentives employed shortage-field teachers with stronger academic backgrounds than other districts, as well as whether teachers in these districts were more likely to be teaching in the subject area for which they were trained. However, in these analyses I did not find any statistically significant relationship between the provision of incentives and the quality of teachers teaching in these districts.

These null results may have been related to the fact that there was very little variation in the backgrounds of teachers in the sample. Specifically, the measure of academic background I used was the Barron's rank of teachers' undergraduate institution. Descriptively, what I found was that despite 9 potential ranking categories, more than half of all teachers in the sample had graduated from schools ranked "competitive", making it difficult to find differences across the population of teachers.

I also undertook an additional fixed effects analysis in which I isolated districts that did and did not offer incentives in different years of the survey. By comparing teachers within the same districts, before and after incentive policies were adopted, I was able to examine the impact of policy changes over time. By utilizing district level fixed effects, I was able to hold constant all district characteristics so that the variation in the model was only between teachers and schools in the same districts. This strategy was designed to create a counterfactual group in which teachers in the same district, and arguably teaching within similar district environments at different times, were compared to one another. In this sub-analysis, I did not find any statistically significant relationships between the addition of a policy with a district, and changes in either retention or recruitment outcomes.

In addition, in order to provide some context to the results found in the empirical analysis, I also utilized the existing models and salary data to try to estimate the magnitude of the financial incentive. Although the results of these analyses do not claim to make a causal connection between specific dollar amounts and teacher outcomes, they do provide illustrative evidence of the size of incentives that teachers are likely being offered in districts that do provide them. These models expose regional variation in salaries received by shortage-field teachers generally, and between districts that do and do not offer incentives. These models have implications for the empirical models because they show that salaries in the Midwest particularly are lower than in other parts of the country. In addition, the incentive estimate in this region was actually negative which indicates that the empirical models are ineffective for estimating impacts of salary enhancements for shortage-field teachers in this region, using these data. Overall, however, having a range of dollar values with which to contextualize the empirical findings is useful from a policy standpoint because it helps frame the discussion in terms of potential costs for districts, and benefits for teachers.

Finally, I also conducted a document review of a subsample of union negotiated contracts from districts that were positively identified as offering incentives in the SASS. The findings from these analyses were particularly helpful in providing some evidence about how incentives are actually applied. They were particularly helpful for understanding possible reasons for null findings on the recruitment outcomes because they presented a picture of districts that utilize incentives mainly to secure specific candidates from the existing labor pool, rather than using incentives as marketing or recruitment tools to entice potential teachers who may not otherwise consider a career in teaching.

The vagueness of some of the language in these contracts may also help illuminate why I find differential effects in the Midwest than in other regions of the country. Namely, leaving in enough language to allow incentives in times of need without tying districts to extra financial commitments when they are unnecessary may provide flexibility to respond to changing local labor market conditions. Thus, districts in the Midwest with vague language contracts may still claim to offer shortage-field incentives because the option is available to them, without actually providing a financial increase for shortage-field teachers when there is a ready supply.

Limitations

One potential criticism of this study is that districts represented by the AFT and the NEA are not really comparable, either because the AFT tends to represent more urban districts, or because the AFT is "known" to be a more progressive union. To address the first point, although it is true that the AFT tends to represent more urban districts than the NEA, the relationship is not absolute. Both unions represent a mix of urban, suburban, and rural affiliates of all sizes. But more importantly, the fact that the AFT has a tendency to represent more urban districts potentially only serves to strengthen the understanding of the results of this study. Specifically, factors associated with urban districts such as high levels of poverty and high minority student populations tend to be associated with greater turnover of teachers, not less. Thus, the finding that AFT districts have a higher likelihood of offering incentives and in turn are more likely to retain shortage field teachers serves to increase rather than decrease the plausibility that the finding is a result of the incentives, and not a result of the AFT affiliation.

As to the idea that the AFT has a reputation for being a more progressive union than the NEA, this is also true. Under the direction of Albert Shanker for more than three decades, the AFT historically has supported many educational reforms, such as charter schools and the development of Peer Assistance and Review teacher evaluation systems that are considered progressive, but which have been resisted by the NEA (Kahlenberg, 2007). However, far from undermining the results of this study, the fact that the AFT offers different policy stances than the NEA was crucial to the development of this research. I needed variation between the unions to create an effective instrument. However, the only policy of which I am aware that the AFT supports that distinguishes between types of teachers within districts is the use of differentiated pay for shortage-field teachers. Otherwise, support for reform policies is generally designed to bolster the professional lives of all the teachers represented by either union. However, the reform-minded reputation of the AFT does mean that the difference-in-differences element in the IV analysis is crucial for providing additional plausibility for the findings. With this element of the study design, I am not just comparing AFT teachers to NEA teachers; rather I am comparing outcomes for teachers of different subjects, teaching under the same contract conditions to each other, as well as to the outcomes of teachers in NEA districts.

In addition to these substantive concerns, the data provided by SASS about the availability of incentives is an extremely noisy measure. The survey only asks a district official whether or not incentives are provided, but does not offer any further illumination on the issue. The 1999 version of the survey did ask districts to specify for which subjects incentives were offered, but the level of detail provided by these sub-questions was cut from the two subsequent survey administrations. Although I sacrificed some level of detail, I chose to use the more blunt question that was asked across all survey waves, rather than limiting the sample size by using only the wave of the survey that contained the greater level of detail.

Moreover, because the survey does not ask districts to answer any questions about the size of the incentives, I was left to estimate their magnitude using salary data and empirical models which provide only a very gross estimate of any differential actually received by shortage-field teachers. That said, considering the variability and vagueness of actual incentive implementation highlighted by the teacher contract document analysis, it is possible that these empirical estimates are at least as accurate as a direct survey question would have been. Moreover, the SASS did contain one teacher level question about how much teachers are paid for a variety of extra activities, including the receipt of shortage-field incentives. I did not employ this question because it asked teachers for a single bonus amount for any activity ranging from coaching activities to student achievement incentives, and parsing out shortage-field incentives would have been impossible. However, considered in light of the document analysis, teachers who are simply placed higher on the salary schedule may very well be unaware that they are even receiving a bonus, so the question, even if asked, may nonetheless garner noisy data.

Overall, the lack of accurate data about the size of incentives offered to specific teachers in specific districts limits these analyses in two ways. First, I was unable to fully explore the impact of shortage-field incentives in relation to the labor market in which they are located. Although the region-specific analyses were supportive of the idea that differing labor market conditions strongly impact the efficacy of shortage-field incentives, being able to match districts to specific local labor market conditions, including salaries offered in comparable non-teaching fields and local unemployment rates would have been more illuminating, and suggests a potential avenue for further research that utilizes richer local data sources.

In addition, the data available for incentive estimates were not robust enough to allow me to investigate whether or not incentive magnitude is impacted by local labor market conditions. For example, are larger incentives enough to overcome the larger opportunity costs created by greater teaching and non-teaching wage differentials in different locations? Or, what is the impact of larger incentives serving as compensating differentials in more challenging school settings? Essentially, I am unable to answer questions about the equilibrium point at which shortage-field incentives can be used to shift teacher supply curves, and whether those shifts differ by location. These too are questions that may be investigated in future research with more fine-grained data.

In addition to the issue of the data related specifically to shortage-field incentives, there are several limitations with the outcomes utilized in this study. The two retention outcomes measured in-school and in-profession retention for just one year. Because the SASS only contains follow-up data for one school year, it was not possible to determine if the incentives might have a cumulative effect over time. In addition, given the findings utilized in the fixed effects analysis that indicate that many districts change their policies over time, it is unclear from the available data how long teachers who were offered incentives actually maintained them. However, given that the AFT policy is to actually place teachers high on the salary scale when they are hired, and the fact that many of the contracts reviewed for the document analysis reflected that policy, it is possible that once teachers have been placed on a specific step, they are not moved from that position on the scale, even if the policy changes. If this is the case, however, it is likely that the implementation of incentives occurs only when new teachers are hired, and are not applied retroactively to existing shortage field teachers in the district. The lack of clarity in how the specific districts in this sample actually applied incentives means that there is also a lack of clarity about whether specific teachers in this sample actually received an incentive. Given this uncertainty, the fact that an effect on retention was nevertheless detected suggests that results may be understated.

One potential problem with the Barron's academic ranking index is that it may or may not accurately reflect teacher quality. Although there is some literature to suggest that teacher academic background is linked to teacher pedagogic ability (Loewenberg Ball, 2008), as with so many other measureable proxies of teacher quality, undergraduate ranking nonetheless provide a noisy signal of teacher ability. Thus, the fact that there is a lack of variation in teacher academic backgrounds, and that most teachers graduated from schools from the lower end of the competitiveness rankings is not necessarily a clear indicator of the quality of the overall teacher labor pool. However, the lack of variation does suggest that changes to the available teacher labor pool will likely need to address more systemic issues surrounding the preparation and development of teachers from a wider range of institutions than can likely be addressed either by shortage-field incentives, or this dissertation.

Although I originally designed the fixed-effects sub-analysis to provide support for the findings in research questions 1 and 2, I was unable to identify any statistically significant impacts within districts over time. At the same time, although I had null findings using these models, they do not undermine the main findings for the original research questions because the analysis sample used here represented mostly different teachers than those used in the first analysis sample. In particular, the sample used to address the first and second research questions was limited to teachers in districts with full collective bargaining rights. By contrast, almost half the teachers in the fixed effects sub-analysis were not covered by collective bargaining. This distinction is important because collective bargaining and the provision of union negotiated contracts likely influences the teacher work environment in ways that are significantly related to both recruitment and retention.

Another possible limitation to the sub-analysis may be that although different teachers were sampled in the same districts in different years, the teachers who responded after the incentive policy change may have actually worked in the district under the policy conditions reported the first time the survey was administered, but not been selected to respond until the second or third survey round. This sampling issue is important because a shortage-field teacher already hired into a district before an incentive policy was put into place may not be impacted despite the official policy change, especially if, as indicated by the document review, most incentives are only offered at the time of hire, and not retroactively. Ideally, these analyses would have only focused on new teachers hired into the districts with policy changes, during the year of the survey. However, the size of the new teacher sample in the SASS was too small to parse to the level of new teachers teaching in districts with policy changes.

Finally, although the before and after models sought to compare equivalent groups across time, because those models specifically focused on districts that were new incentive implementers, a span of a year of implementation may not have been enough for detectable impacts of incentives on within district change. For all these reason, the results of the IV difference-in-differences models should be considered more robust than those provided by the fixed effects models.

Links to Theory and Extant Literature

The finding that districts with incentives are 17 percent more likely to retain teachers than comparable districts without incentives is consistent with results found in previous work examining the efficacy of shortage-field incentives in North Carolina (Clotfelter et al., 2008). In addition, the mean estimate of an incentive magnitude of $1,600 found in this study is similar in magnitude to the $1,800 bonus paid to teachers in the North Carolina evaluation. The results of the retention analyses therefore serve to confirm existing evidence about the positive impact of shortage-field incentives on teacher attrition. They also expand the generalizability of the findings to include a variety of state policy settings, including those that, unlike North Carolina, require collective bargaining for public sector employees, as well as those in regions of the country that have experienced population increases and difficulty recruiting shortage-field teachers.

Moreover, the findings for the in-school teacher retention outcome support the theoretical predictions of the Roy Model, at least in part. By offering shortage-field incentives, districts seek to close the wage gap between teaching and non-teaching professions for those in shortage-fields. In doing so, the districts in this sample that offered incentives were able to reduce the economic opportunity cost to the profession for these teachers, and were successful in helping to retain their teachers.

However, the Roy Model also predicts that wages will serve to impact the selection mechanisms influencing who chooses specific professions. The fact that I was unable to identify impacts specifically for new teachers leaves this question open, and likely requires much larger samples of new shortage-field teachers in order to address satisfactorily. In addition, the lack of variation in the academic backgrounds of teachers also suggests that issues of selection into teaching are complicated earlier in the pipeline of teacher preparation. Specifically, most teachers are still prepared for the profession through traditional preparation programs, the vast majority of which seem to be training teachers who attended colleges of mediocre academic reputation. Although it is unclear from these data why this occurs, the consequence is that lack of variation in the pool of people who become teachers means that shortage-field incentives offered only to those prospective hires who have already entered the teacher labor market are ineffective for changing the selection mechanisms of workers into teaching. The lag time created between selecting into a teacher preparation program and being hired into a teaching job where an incentive may be offered complicates the selection predictions of the Roy model.

Moreover, the recruitment findings are also well understood in light of the theoretical discussion around asymmetric information, and the lack of ability of districts to identify quality teachers at the point of hire. In all likelihood the districts in this sample relied on standard mechanisms for identifying candidates, including credentials and degrees, regardless of the quality of the institution that granted them. However, since these credentials only provide weak signals of teacher quality, it is unsurprising that districts were unable to alter the kinds of teachers they hired, particularly in a market in which there is little variation, at least in terms of the academic backgrounds of those available to teach. Together these findings confirm previous research which suggests that labor market sorting is likely not affected by changes in salary incentives for teachers (Fowler, 2003; Liu et al., 2004).

Furthermore, results from the document review help further illuminate these null findings. Specifically, these analyses suggest that the current implementation and utilization of shortage-field incentives as a recruiting tool is limited to securing those who have already self-selected into the teaching labor market, for specific positions in specific schools. However, another mechanism by which incentives could work is to appeal to a wider candidate pool for teaching positions in shortage-fields. Given that so many teachers actually assigned to teach mathematics and science are teaching out of their field of expertise, it seems a missed opportunity for districts that do offer incentives, but do not currently provide them in a systematic fashion. Specific financial amounts and clear guidelines about how shortage-field incentives are to be distributed could potentially allow districts to market incentives more aggressively to college students and new graduates with degrees in these subject fields, opening a window into the profession for students who may be interesting in teaching, but who currently hesitate because of teaching's reputation for poor pay. Suggestions for future research include the investigation of whether better information about the availability of incentives serves to influence the career selection process of college students in STEM fields.

As discussed, the findings from the fixed effects analyses executed in research question three did not confirm the main findings from the IV models utilized in research questions one and two. In addition, the sample utilized for the different analyses differed, particularly in relation to the number of teachers teaching under collective bargaining contracts. More than other sample differences, this factor may have influenced the difference in results because teachers who have unions to help them address grievances may have overall higher retention rates because unions allows them exercise "voice", rather than simply "exit" as has been described as one of the main functions of unions (Hirschman, 1970). Despite controlling for collective bargaining in the models, having a large sample of teachers working under a mixed variety of contract circumstances may have nonetheless influenced the ability to detect significant findings for the outcomes of interest.

Finally, the analyses presented in this dissertation serve to contribute to some new and emerging research that seeks to provide more quantitative nuance to the understanding of the role of teachers' unions, and the heterogeneity in the kinds of contracts that govern teachers' work (Strunk, 2011; Strunk & Grissom, 2010). By utilizing the differences in specific policy stances to instrument for policy effects, this study suggests a possible new direction for the development of similar instruments, or of the same instrument to analyze the impacts of other union and/or contract policies. In addition, much of the existing policy rhetoric about teachers' unions fails to make distinctions between the differing policy outlooks of the nation's two largest teachers' unions. All too often teachers' unions are painted with one brush as monolithic obstacles to reform initiatives. By parsing the differing policy stances of the AFT and the NEA this dissertation demonstrates that differences in policy stances have real causal implications for the functioning of teacher labor markets.

Policy Implications

Although this dissertation examined the role of shortage-field incentives in the retention and recruitment of teachers in shortage sub-fields, the policy implications apply more broadly to the teaching profession as a whole. This is because although I have demonstrated that shortage-field incentives are effective for increasing retention of shortage-field teachers, the findings also beg the question of whether we want to retain the teachers we have. In the short-term, all evidence affirms that retaining current teachers in areas where we already have shortages makes good educational as well as economic and policy sense. As discussed in chapter 1, constantly replacing teachers jeopardizes student achievement, and costs districts resources that could be better invested. Shortage-field incentives are likely an effective mechanism for addressing these problems. At the same time, evidence about teacher quality presented in both the literature review and as evidence in this dissertation suggests that although improvements in shortage-field teaching are needed just to obtain quality parity with current non-shortage field teaching, a long-term prescription for improving teacher quality overall is not limited to teachers in shortage-fields.

As Labaree discusses, the policy environment is currently in the process of coalescing around a view of education as a consumer commodity, and a private good. One of the implications of this ideal, as well as the data-driven environment it has given rise to, is that there is an ever growing need for school districts to find ways to recruit and retain teachers who are more highly qualified to help students enhance their individual academic performance in order to ultimately become more competitive in the labor market. In addition, as the labor market itself becomes more and more global the pressure to create students who can compete, not just with other Americans, but with students all over the world, many of whom have superior technical skills to American students, is increasing as well. Globalization means that not only increasing policy pressure to improve student performance, but increased pressure to improve teacher performance as well. This convergence of policy ideals and labor market changes means that school inputs, which include the instructional capacity of teachers, must be more tightly coupled with school output, or student performance. Since teachers are the most influential school-based resource for student achievement, the changing policy environment has greatly increased the pressure to both improve teacher quality, and to retain the best teachers. However, findings also indicate that districts still need better tools for identifying high quality teachers who will both enter and remain in teaching.

Thus, improving teacher quality is a long-term prospect. Based on the lack of variation in teacher backgrounds, and the complications that are created in the hiring process due to poor signal quality, it seems that improving the self-selection of teachers into the profession should begin not in the district personnel office, but rather through the reform of teacher preparation programs, both in terms of quality of candidates recruited into them and the relevance of teacher preparation curricula for classroom teaching. In order to meet the demands of an education system that is striving toward the ideal of developing workers who are competitive in an economy that meters out rewards for increased skills, teachers themselves must reasonably possess both deep knowledge of the level of skills that students will be expected to master, as well as the pedagogical acumen to transmit that knowledge to students in meaningful ways. By recruiting teachers who themselves potentially possess mediocre levels of academic ability, it is unlikely that students of these teachers will be able to develop skills of excellent quality. Thus, recruiting more academically able teachers, particularly in fields that require strong academic knowledge, should be a priority.

At the same time, given the lack of connection that has been found between teacher credentials and student performance, in addition to changing the recruitment of teachers in teacher preparation programs, change in the content of how teachers are prepared to teach is also warranted. There is likely a reciprocal relationship between the quality of teachers who enter teaching programs, and the rigor of content provided by those programs. Logically, programs that seek to attract and prepare excellent teachers should begin by offering challenging, engaging, and relevant preparation curricula. That said, while these changes in teacher preparation are necessary, they are likely not sufficient to change the pool of potential teachers because high ability people likely seek to be both intellectually engaged, and financially rewarded.

In order for high ability individuals to opt into teaching, they also likely need to feel that the increased investment involved in a more challenging preparation program will be met with some increase in payoff when they actually enter teaching. Thus, in order to achieve overall improvements in the pipeline for quality teachers, there must be both better signals for districts, and increased rewards for teaching. Districts may only be willing or able to raise salaries if they are better able to predict which job candidates will both perform to high standards in the classroom, and remain in the job for enough time to reduce the negative fiscal and academic costs associated with their attrition. By the same token, increasing the supply of the kinds of teachers who, from the district perspective, will be worth the increased rewards will only occur if the payoff for teaching is increased.

In other words, although teacher compensation needs to increase, from a systemic perspective the increases are not needed to continue to recruit and retain teachers meeting the current qualification criteria. Rather they are needed to recruit and retain teachers who surpass current requirements. From a policy perspective, teachers who surpass expectations and who can be recruited in the numbers necessary to staff the nations schools must be both prepared and compensated in new ways, and at scale. This is because improvements in signal quality between teachers and district hiring managers is a symbiotic relationship where change can only occur if improvements on both ends are made simultaneously.

Although achieving these goals will no doubt involve complex processes that engage many actors along the teacher development pipeline and hiring, one outcome may be that if districts can obtain better signals for identifying teachers, they will likely be able to both recruit and retain teachers who will be better qualified to both develop and enhance student learning in schools of all kinds.

References

Akerlof, G. (1970). The market for lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500.

Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009). *The Schools Teachers Leave: Teacher Mobility in Chicago Public Schools*. Consortium on Chicago School Research.

Angrist, J. D., & Guryan, J. (2004). Teacher Testing, Teacher Education, and Teacher Characteristics. *The American Economic Review*, *94*(2).

Angrist, J. D., Imbens, G. ., & Rubin, D. B. (1993). *Identification of Causal Effects Using Instrumental Variables* (No. 136). NBER Technical Paper. National Bureau of Economic Research.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.

Ballou, D., & Podgursky, M. (2001). Let the Market Decide. *Education Next*.

Barnes, G., Crowe, E., & Schaefer, B. (2007). *The Cost of Teacher Turnover in Five School Districts: A Pilot Study*. National Commission on Teaching and America's Future. Retrieved from http://eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED497176&ERICExtSearch_SearchType_0=no&accno=ED497176

Beaudin, B. Q. (1993). Teachers Who Interrupt Their Careers: Characteristics of Those Who Return to the Classroom. *Educational Evaluation and Policy Analysis*, *15*(1), 51–64.

Bellas, M. L. (1997). Disciplinary Differences in Faculty Salaries: Does Gender Boas Play a Role? *The Journal of Higher Education*, *68*(3), 229–321.

Billingsley, B. S., & McLeskey, J. (2004). Critical Issues in Special Education Teacher Supply and Demand. *Journal of Special Education*, *38*(1), 2–4.

Borman, G. D., & Dowling, N. M. (2008). Teacher Attrition and Retention: A Meta-Analytic and Narrative Review of the Research. *Review of Educational Research*, *78*(3), 367–409.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). *Who Leaves? Teacher Attrition and Student Achievement* (Working Paper No. 23). National Center for Analysis of Longitudinal Data in Education Research.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J. E., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management*, *27*(4), 793–818.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *The American Economic Review*, *95*(2), 166–171.

Bradely, K. D., & Loadman, W. E. (2005). Urban secondary educators' views of teacher recruitment and retention. *NASSP Bulletin*, *89*(644).

Carnoy, M., & Levin, H. M. (1985). *Schooling and Work in the Democratic State*. Stanford, CA.: Stanford University Press.

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

Clotfelter, C., Ladd, H., & Vigdor, J. (2011). Teacher Mobility, School Segregation, and Pay-Based Policies to Level the Playing Field. *Education Finance and Policy*, *6*(3), 399–438.

Corcoran, S. P., Evans, W. N., & Schwab, R. S. (2002). *Changing labor market opportunities for women and the quality of teachers 1957-1992*. National Bureau of Economic Research Working Paper 9180.

Darling-Hammond, L. (2004). Inequality and the right to learn: Access to qualified teachers in California's public schools. *Teachers College Record*, *106*(10), 1936–1966.

Diaz-Rico, L. T., & Smith, J. (1994). Recruiting and Retaining Bilingual Teachers: A Cooperative School Community-University Model. *The Journal of Educational Issues of Language Minority Students*, *14*, 255–268.

Dolton, P., & van der Klaauw, W. (1999). The Turnover of Teachers: A Competing Risks Explanation. *The Review of Economics and Statistics*, *81*(3), 543–550.

Farrell, E. F., & Van Der Werf, M. (2007, May 25). Playing the Rankings Game. *The Chronicle of Higher Education*.

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

Fowler, R. C. (2003). Massachusetts Signing Bonus Program for New Teachers. *Education Policy Analysis Archives*, *11*.

Fryer, R. G. (2011). *Teacher Incentives and Student Achievement: Evidence from New York City Public Schools* (No. 16850). Working Paper. National Bureau of Economic Research.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Heirarchical Models*. New York: Cambridge University Press.

Goldhaber, D. (2009). The Politics of Teacher Pay Reform. In M. G. Springer (Ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*. Washington, DC: The Brookings Institute.

Goldhaber, D., DeArmond, M., Liu, A. Y.-H., & Player, D. W. (2008). *Returns to skills and teacher wage premiums* (No. 8). Working Paper. Center on Reinventing Public Education, University of Washington.

Goldhaber, D., Gross, B., & Player, D. (2010). *Teacher Career Paths, Teacher Quality, and Persistence in the Classroom: Are Public Schools Keeping their Best?* (No. 2010-2). Working Paper. Center for Education Data and Research.

Goldhaber, D., & Hansen, M. (2010). Race, Gender, and Teacher Testing: How Informative a Tool is Teacher Licensure Testing? *American Educational Research Journal*, *47*(1), 218–251.

Goldhaber, D., & Liu, A. Y.-H. (2005). *Teacher Salary Structure and the Decision to Teach in Public Schools: An Analysis of Recent College Graduates* (Working Paper No. 2005-3). Center on Reinventing Public Education.

Goldhaber, D., & Player, D. (2005). What different benchmarks suggest about how financially attractive it is to teach in public schools. *Journal of Education Finance*, *30*(3), 211–230.

Goodman, S., & Turner, L. (2009). *Group Incentives for Teachers: The Impact of the NYC School-Wide Bonus Program on Educational Outcomes* (No. 0910-05). Department of Economics Discussion Paper. Columbia University.

Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher Recruitment and Retention: A Review of the Recent Empirical Literature. *Review of Educational Research*, *76*(2), 173–208.

Guin, K. (2004). Chronic Teacher Turnover and Urban Elementary Schools. *Education Policy Analysis Archives*, *12*(42), 2–25.

Hannaway, J., & Mittleman, J. (2011). Education Politics and Policy in an Era of Evidence. *Shaping Education Policy: Power and Process*. New York: Routledge.

Hanushek, E. A., Kain, J. F., O'Brien, D., & Rivkin, S. G. (2005). *The Market for Teacher Quality* (No. 11154). NBER Working Paper.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). *Teachers Schools and Academic Achievement*. National Bureau of Economic Research Working Paper 6691.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *Journal of Human Resources*, 326–354.

Hanushek, E. A., Rivkin, S. G., Rothstein, R., & Podgursky, M. (2004). How to improve the supply of high-quality teachers. *The Bookings Paper on Education Policy*, *7*.

Harris, D. N., & Adams, S. J. (2007). Understanding the Level and Causes of TeacherTurnover: A comparison with Other Professions. *Economics of Education Review*, *26*, 325–337.

Henke, R. R., Zahn, L., & Carroll, C. D. (2001). *Attrition of New Teachers Among Recent College Graduates: Comparing Occupational Stability Among 1992-93 Graduares Who Taught and Those Who Worked in Other Occupations*. National Center for Education Statistics.

Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Hoxby, C. M., & Leigh, A. (2004). Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States. *The American Economic Review*, *94*(2), 236–240.

Ingersoll, R. (1999). The problem of underqualified teachers in American secondary schools. *Educational Researcher*, *28*(2), 26–37.

Ingersoll, R. M. (2001). Teacher Turnover and Teacher Shortages: An Organizational Analysis. *American Educational Research Journal*, *38*(3), 499–534.

Ingersoll, R. M. (2003). *Is There Really a Teacher Shortage?* (No. R-03-4). Center for the Study of Teaching and Policy.

Ingersoll, R., & Perda, D. (2010). Is the Supply of Mathematics and Science Teachers Sufficient? *American Educational Research Journal*, *47*(3), 563–594.

Kahlenberg, R. (2007). *Tough Liberal: Albert Shanker and the Battles Over Schools, Unions, Race, and Democracy.* New York: Columbia University Press.

Kaiser, A., & Cross, F. (2011). *Beginning Teacher Attrition and Mobility: Results From the First Throigh Third Waves of the 2007-08 Beginning Teacher Longitudinal Study.* National Center for Education Statistics.

Kane, T., Rockoff, J. E., & Staiger, D. O. (2006). What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City. *NBER w12155*.

Kelly, P. C., Tejada-Delgado, C., & Slate, J. R. (2010). Financial and Non-Financial Incentives on Teacher Recruitment and Retention: Teachers' Perspectives. *International Journal of Educational Leadership*, *5*(1).
Kershaw, J. A., & McKean, R. N. (1962). *Teacher Shortages and Salary Schedules*. New York: McGraw-Hill Book Company.
Labaree, D. F. (1997). Public Goods, Private Goods: The American Struggle Over Educational Goals. *American Educational Research Journal*, *34*(1), 39–81.
Lakdawalla, D. (2001). *The declining quality of teachers*. National Bureau of Economic Research working paper 8263.
Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62.
Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *The Journal of Political Economy*, *110*(6), 1286–1317.
Levin, H. M. (1985). Solving the shortage of mathematics and science teachers. *Educational Evaluation and Policy Analysis*, *7*(4), 371–382.
Lindy, B. A. (2011). The Impact of Teacher Collective Bargaining Laws on Student Achievement: Evidence from a New Mexico Natural Experiment. *Yale Law Journal*, (120), 1132–1183.
Liu, E., Johnson, S. M., & Peske, H. G. (2004). New Teachers and the Massachusetts Signing Bonus: The Limits of Inducements. *Educational Evaluation and Policy Analysis*, *26*(3), 217–236.
Loeb, S., & Darling-Hammond, L. (2005). How Teaching Conditions Predict Teacher Turnover in California Schools. *Peabody Journal of Education*, *80*(3), 44–70.
Loewenberg Ball, D. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, *59*(5), 389.
Loprest, P. J. (1992). Gender Differences in Wage Growth and Job Mobility. *The American Economic Review*, *82*(2), 526–532.
MacLeod, W. B., & Urquiola, M. (2009). *Anti-Lemons: School Reputation and Educational Quality* (No. 15112). NBER Working Paper.
Milanowski, A. (2003). An Exploration of the pay levels needed to attract students with Mathematics, Science and Technology skills to a career in K-12 teaching. *Education Policy Analysis Archives*, *11*(50).
Milanowski, A., & Odden, A. (2007). *A New Approach to the Cost of Teacher Turnover* (Working Paper No. 13). Center on Reinventing Public Education.
Murnane, R. J., & Olsen, R., J. (1989). The effect of Salaries and Opportunity Costs on Duration in Teachinf: Evidence from Michigan. *The Review of Economics and Statistics*, *71*(2), 347–352.
Murnane, R. J., & Olsen, R. J. (1990). The effects of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. *Journal of Human Resources*, *25*(1), 106–124.
Murphy, M. (1990). *Blackboard Unions: The AFT and The NEA 1900-1980*. Ithica and London: Cornell University Press.
Murphy, P., DeArmond, M., & Guin, K. (2003). A National Crisis or Localized Problems? Getting Perspective on the Scope and Scale of the Teacher Shortage. *Education Policy Analysis Archives*, *11*(23).

Obama, B. (2011, January). Remarks by the President in State of the Union Address. Retrieved from http://www.washingtonpost.com/wp-srv/politics/documents/Obama_SOTU_2011_transcript.html

Player, D. (2010). Nonmonetary Compensation in the Public Teacher Labor Market. *Education Finance and Policy*, *5*(1), 82–103.

Podgursky, M. (2006). Is there a "Qualified Teacher" Shortage? *Education Next*.

Podgursky, M., Monroe, R., & Watson, D. (2004). The Academic Quality of Public School Teachers: An Analysis of Entry and Exit Behavior. *Economics of Education Review*, *23*(5), 507–518.

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

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *The American Economic Review*, *94*(2), 247–252.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2008). *Can You Recognize and Effective Teacher When You Recruit One?* (Working Paper No. 14485). National Bureau of Economic Research.

Rockoff, J. E., & Speroni, C. (2011). Subjective and objective evaluations of teacher effectiveness: Evidence from New York City. *Labour Economics*, *18*(5), 687–696.

Roekel, D. V. (2008). *Alternative compensation: Create a solid pay system and then add enhancements*. NEA Policy Brief.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2012). *How Teacher Turnover Harms Student Achievement* (Working Paper No. 70). National Center for Analysis of Longitudinal Data in Education Research.

Roy, A. D. (1951). Some Thoughts in the Distribution of Earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Royalty, A. B. (1998). Job-to-Job and Job-to-Nonemployment Turnover by Gender and Education Level. *Journal of Labor Economics*, *16*(2), 392–433.

Rumberger, R. W. (1987). The impact of salary differentials on teacher shortage and turnover: The case of mathematics and science teachers. *Economics of Education Review*, *6*(4), 389–399.

Smith, R. S. (1979). Compensating Wage Differentials and Public Policy: A Review. *Industrial and Labor Relations Review*, *32*(3).

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Springer, M. G., Hamilton, L., McCaffrey, D. F., Ballou, D., Le, V.-N., & Stecher, B. M. (2010). *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching*. National Center of Performance Incentives.

Stinebrickner, T. R. (2001). A Dynamic Model of Teacher labor Supply. *Journal of Labor Economics*, *19*(1), 196–230.

Stinebrickner, T. R. (2002). An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Leave. *The Journal of Human Resources,*, *37*(1), 192–216.

Strunk, K. O. (2011). Are Teachers' Unions Really to Blame? Collective Bargaining Agreements and Their Relationship with District Resource Allocation and

Student Performance in California. *Education Finance and Policy*, *6*(3), 354–398.
Strunk, K. O., & Grissom, J. A. (2010). Do Strong Unions Shape District Policies? *Educational Evaluation and Policy Analysis*, *32*(3), 389–406.
Tuckman, B. H., & Tuckman, H. P. (1976). The Structure of Salaries at American Universities. *The Journal of Higher Education*, *47*(1), 51–64.
Tyack, D., & Cuban, L. (1995). *Tinkering Toward Utopia*. Cambridge, MA and London, UK: Harvard University Press.
Vigdor, J. (2008). Scrap The Sacrosanct Salary Schedule. *Education Next*, (Fall).
Wayne, A., J. (2000). Teacher Supply and Demand: Surprises from Primary Research. *Education Policy Analysis Archives*, *8*(47).
Weick, K. E. (1976). Educational Organizations as Loosley Coupled Systems. *Administrative Science Quarterly*, *21*(1), 1–19.
West, M. R., & Chingos, M. M. (2009). Teacher Effectiveness, Mobility, and Attrition in Florida. *Performance Incentives: Their Growing Impact on American K-12 Education*. Washington, DC: The Brookings Institute.
Where We Stand: Teacher Quality. (2003). American Federation of Teachers.

APPENDIX

Table 1
Description of all control variable used in analytic models

Variable Name	Coding
District Variables	
Census Region	Categorical.
Total k-12 enrollment	Continuous
District offers Shortage field incentives	Indicator.
Teachers union agreement	Categorical. There were 3 categories available: collective bargaining, meet-and-confer, and no union negotiated agreement. The variable was used to identify "meet-and-confer" states in R1 and R2. In the analysis for R3, an indicator for collective bargaining was created and used as a control
Urbanicity	Categorical. Indicators for urban, suburban, and rural were created. Suburban was used as the reference category in all models
Union affiliation	Indicator in which 1=AFT, and 0=NEA
State Abbreviation	Categorical. Each state assigned a single number
Survey Year	Indicator variables for the year each teacher participated in SASS were created, and used as fixed effects.
School Variables	
Enrollment	Continuous.
% of students approved for lunch	Continuous. School percentage of students eligible for the federal lunch program.
% Hispanic students	Continuous. School percentages were not included in SASS. Counts of the number of all Hispanic students were divided by the school enrollment variable, and multiplied by 100.
% Black students	Continuous. School percentages were not included in SASS. Counts of the number of all Black students were divided by the school enrollment variable, and multiplied by 100.
% students with IEPs	Provided as a continuous percentage variable.
% students classified as LEP	Provided as a continuous percentage variable.
Charter school	Indicator. 1=charter
Teacher Variables	
Totally yearly earnings	SASS included 2 salary measures. One was

	yearly salary from teaching. The other was yearly annual earnings. The salary measure was the most applicable to this research. However it contained a fair amount of missing values. In places where annual teaching salary was missing, the annual yearly earnings value was imputed, if available.
Years of teaching experience	Continuous
New teacher flag	Indicator. 1=3 years of teaching or less.
School grade level taught	Categorical. Indicators were created for elementary, middle, high school, and combined. In all models, elementary was used as the reference category.
Field of main assignment	Categorical measure used to create indicators of shortage field teachers, and infield status.
STEM teacher	Indicator. 1= any teacher with a main teaching assignment of math or science.
Special education teacher	Indicator. 1= any teacher whose main assignment is teaching special education
Bilingual Teacher	Indicator. 1= any teacher whose main assignment is teaching bilingual education.
Teacher gender	Indicator. 1= Male. Female is the reference category.
Teacher age	Continuous
Teacher status 1 year post survey	Categorical variable indicating if a given teacher remained in the same school as a teacher, remained in the school in a non-teaching position, moved to another school, left teaching, died. Used to create retention outcome variables.
Retained in school	Indicator. 1= teacher still teaching in the same school, 1 year later.
Retained in teaching	Indicator. 1= teacher still working as a classroom teacher.
Year of BA degree	Continuous.
Type of certification	Categorical. ??? used to create variable of whether or not a teacher was fully certified.
Fully certified.	Indicator. 1= teacher holds full state certification.
Field of Bachelor's degree	Categorical. Exhaustive list of possible fields of undergraduate study by teachers. Used to create measures of infield teaching status, and STEM BA.
STEM BA	Indicator. 1=Teachers with a BA in the field of their main teaching assignment.
Infield teaching status	Indicator. 1=teacher has a BA in a math or science subject area, including math or science education; teacher has a BA in a special

	education related field; teacher has a BA in bilingual education or Spanish.
Barron's rank of undergraduate institution	Ordinal. 1= least competitive and 9=most competitive.
Frequency	Ordinal. 1=teacher taught in district the first year incentive was offer, 2= teacher taught in district the second year the incentive was offered, etc.
After.	Indicator. 1= teacher taught in district the same year incentive was offered. 0=teacher taught in district the year before incentive was offered.

Appendix Table 2: Criteria for Barron's Rankings*

Category	1972	1982	1992	2004
Most Competitive				
GPA	B+ to A	B+ to A	B+ to A	B+ to A
Class Rank	Top 10-20%	Top 10-20%	Top 10-20%	Top 10-20%
Median SAT	675-800	625-800	625-800	655-800
Median ACT	28	27+	29+	29+
Percent Admitted	Small percentage	Less than 33%	Less than 33%	Less than 33%
Highly Competitive				
GPA	B to B+	B to B+	B to B+	B to B+
Class Rank	Top 20-30%	Top 20-35%	Top 20-35%	Top 20-35%
Median SAT	600-675	575-625	575-625	620-654
Median ACT	26-28	26-27	26-28	26-28
Percent Admitted	Less than 25%	33-50%	33-50%	33-50%
+(plus)	SAT 675; ACT 27; accept less than 25%	SAT 615+; ACT 27+; accept less than 25%	SAT 616+; ACT 28+; accept less than 25%	SAT 645+; ACT 28+; accept less than 25%
Very Competitive				
GPA	B- or above	B- or above	B- or above	B- or above
Class Rank	Top 30-50%	Top 35-50%	Top 35-50%	Top 35-50%
Median SAT	550-600	525-575	525-575	573-619
Median ACT	23-26	25-26	24-26	24-26
Percent Admitted	NA	Less than 33%	Less than 33%	50-66%
+(plus)	Median SAT/ACT scores close to top of range; less than 33% admitted	Sat 565+; ACT 25+; Less than 33% admitted	Sat 565+; ACT 26+; Less than 33% admitted	Sat 610+; ACT 26+; Less than 33% admitted
Competitive				
GPA	C to B-	C to B-	C to B-	C to B-
Class Rank	Top 30-50%	Top 35-50%	Top 35-50%	Top 35-50%
Median SAT	400-550	450-525	450-525	500-572
Median ACT	21-23	19-22	21-23	21-23
Percent Admitted	75-85%	75-85%	75-85%	75-85%
+(plus)	SAT 500+; ACT 22+; Less than 50% accepted	SAT 515+; ACT 23+; Less than 50% accepted	SAT 515+; ACT 24+; Less than 50% accepted	SAT 563+; ACT 23+; Less than 50% accepted
Less Competitive				
GPA	C	Below C	Below C	Below C
Class Rank	Top 75%	Top 65%	Top 65%	Top 65%
Median SAT	Less 450	Less 450	Less 450	Less 500
Median ACT	Less 21	Less 19	Less 21	Less 21
Percent Admitted	75%	Top 85%	Top 85%	Top 85%
Non-Competitive				
GPA	HS diploma	HS diploma	HS diploma	HS diploma
Class Rank	NA	NA	NA	NA
Median SAT/ACT	May require placement exam	May require placement exam	May require placement exam	May require placement exam
Percent Admitted	All	All	98%	98%

*Reproduced from *Documentation for the Restricted-Use NCES- Barron's Admissions Competitiveness Index Data Files: 1972, 1982, 1992, 2004, and 2008*