Wage Frictions and Teacher Quality: An Empirical Analysis of Differential Effects Across Subject Areas *

D. Chris Ferguson† and Greg Gilpin
Indiana University- Bloomington
January 2009

Abstract:

Using teacher level data from public, private, and charter schools, we examine how various school level wage frictions such as salary schedules and collective bargaining agreements affect the level of teacher quality across subject levels in U.S. high schools. We find that salary schedules have a significant negative effect on the scholastic aptitude of teachers within a school, reducing aptitude by as much as 2-3%. The effect is particularly large in math and science and at the higher end of the quality distribution where the rigidity imposed by the schedules is most likely to bind. At the highest quantile of teacher aptitude we find an additional reduction of nearly 3.5% for math and science teachers in salary schedule schools. We also find evidence that average class sizes are as much as 10-12% larger in schools with rigid wage structures. We further find that a reduction in this rigidity through the use of teacher incentive pay has significant positive effects increasing quality by approximately 1.5%. Finally, we examine the impact of our measurement of teacher quality and class size on student achievement outcomes and find significant effects in both graduation and college attendance rates.

JEL classification: I2, I22, J44, J51
Keywords: Collective Bargaining, Teacher Union, Incentive Based Pay, Education Finance Reform, Math and Science Education, Teacher Quality, Student Achievement, Schools and Staffing Survey, Teacher Roles Rights and Responsibilities

---

**Acknowledgements*: Funding for this study was provided, in part, by a grant from the National Council on Teacher Quality. We would like to thank Gerhard Glomm, William Becker, Brian Peterson, Rusty Tchernis, Pravin Trivedi, the members of the public economics and DBSE working groups at Indiana University, and the Indiana University Center for Evaluation & Education Policy (CEEP) for their useful comments and assistance.

†Corresponding author: D. Chris Ferguson (dcfergus@indiana.edu), Indiana University, 100 S. Woodlawn Ave, Bloomington, IN 47405-7104. Phone 812-855-1021

I. ***Introduction***

An increased emphasis on teacher quality has been seen in the education reform debate following the passing of the No Child Left Behind Act of 2002, with particular concern over the quality of math and science instruction. Recent findings in the literature have highlighted the connection between the quality and quantity of teachers and the educational attainment of students. As a part of this discussion, there has been much concern over the effects of teacher unions and collective bargaining agreements on education quality and student achievement. The central focus of this study is to examine how the wage rigidities imposed by these agreements can have particularly important implications on the relative quality of education in math and science.

When wages are determined by a rigid salary schedule that is dependent only on experience and education levels, as is common under many union contracts, the flexibility to pay teachers based on their individual ability levels is limited. This wage friction can be particularly binding when wages are equalized across different subjects within a school. Individuals with knowledge and ability in math and science face different wage opportunities in the non-education labor market than those whose knowledge and ability is focused in the humanities. As shown in Table 1, college graduates with training in math and sciences earn a significant wage premium. The average salary for teachers is much lower than the average salary for those individuals with degrees in the sciences, but is roughly competitive with the average salary of individuals trained in the humanities. It should be expected then that if a fixed salary structure is imposed, in which teachers with the same experience are paid the same wage regardless of the subject taught, for any given wage and *quantity* of teachers the *quality* of teachers will differ across subjects due to the differences in their outside options for other jobs.

Using a new combined data set formed from the National Center for Education Statistics' Schools and Staffing Survey (SASS), the National Post-Secondary Aid Survey (NPSAS), and the National Council on Teacher Quality's Teacher Roles, Rights, and Rules (TR^3) Database we examine how teacher quality within a school is affected by the relative degrees of wage frictions imposed by salary schedules, the use of incentive or reward pay, and the presence of collective bargaining agreements. In order to gauge the potential impact of these quality differences, we also examine the effects on student achievement outcomes, as measured by graduation and college attendance rates.

The body of relevant literature that explores similar issues of teacher quality and education is rich and varied and can be categorized into several broad categories. The first of these is the branch that studies the relationship between teacher quality and student outcomes. Teacher quality itself is a broad category, which includes dimensions such as experience, subject knowledge, and scholastic aptitude. Beginning with the scholastic aptitude of teachers, several studies have examined the relationship between a teacher's test scores or the quality of their undergraduate institution and the outcomes of students on standardized tests. These studies have commonly found that teachers with high test scores or highly selective educational backgrounds are more likely to produce gains in student achievement. Ferguson (1991) and Ferguson and Ladd (1996) find significant positive effects of teacher test scores on student test scores in Texas and Alabama schools respectively. Strauss and Sawyer (1986) find that a 1% increase in the standardized test scores of teachers increases the pass rates of North Carolina high school students by 5% on math and reading proficiency tests. Ehrenberg and Brewer (1993) find that the quality of a teacher's undergraduate institution is highly related with student test outcomes and find that a one category increase in the selectivity of a teacher's institution is associated with a 1-2% increase in student test scores. The practice of using the selectivity of a teacher's undergraduate college to proxy for their scholastic ability is common in the literature, and further examples include Summers & Wolfe (1977) and Winkler (1975), who find similar positive effects on student achievement. For a general survey of the literature linking teacher academic ability and student achievement, Greenwood et al. (1996) and Hanushek (1981, 1986) provide an excellent summary of the findings. These surveys generally conclude that of all measurable school and teacher characteristics, academic ability of teachers has the largest effect on student outcomes. As Hanushek (1981) states, "The only relatively consistent finding is that 'smarter' teachers seem to do better in terms of student achievement."

A second thread in the literature focuses on specific subject knowledge of teachers. Aaronson (2007) examines 9th-grade math teachers in the Chicago public schools and finds correlation between observable teacher characteristics and student outcomes, with one of the strongest effects coming from a teacher's undergraduate major. Math and science majors were found to have a positive effect on student math scores while education majors had a negative effect. Monk (1994) examines the impact of deep subject or content-area knowledge, such as a major or minor in the subject taught, and finds that significant student learning gains in U.S. high schools occur with these types of teachers. Monk further finds these subject knowledge gains to be particularly strong in math and science subjects. Clotfelter et al. (2007) make use of detailed data on 3rd-5th-graders in North Carolina public schools and find that the effect of improved teacher credentials is especially strong in math achievement and appears to be non-linear with bigger effects found at either end of the teacher quality distribution. Goldhaber and Anthony (2007) also make use of North Carolina data to examine the effects of the National Board Certification process and also find mixed evidence that improved observable teacher credentials such as degree type and quality have positive student achievement impacts.

Several authors have explored the effect of class size on education production, although the evidence here is mixed. Rivkin et al. (2005) use Texas public school data and find that a ten student reduction in class size produces smaller benefits to student achievement than a one standard deviation increase in teacher quality. Clotfelter et al (2007) similarly find that measurable teacher credentials have stronger effects than class size in North Carolina public schools. Krueger (1999) studies the impact of class size on student achievement and finds positive effects for smaller classes in Tennessee public schools with increases in test scores ranging from 1-4%. Angrist and Lavy (1999) use data from Israeli schools and find dramatic increases in student test scores for 4th- and 5th-graders in smaller classes but not for 3rd-graders. The effect of class size, therefore, is still somewhat ambiguous, but there is at least some evidence to suggest that the quantity of teachers as well as their quality plays a role in the educational achievement of students.

If the literature therefore suggests that teacher quality and quantity are key components in educational production, a remaining question is how the unique educational labor market conditions imposed by union and non-union wage rigidities influence these dimensions. There have been several studies that looked specifically at the labor supply decision for teachers. Bonnesronnig et al (2005) take advantage of unique union hiring rules to identify the supply curve for teachers in Norway and find that when wages are rigidly structured, teachers seem to sort themselves into schools that have attractive bundles of other characteristics such as workload and student body composition. Dolton (1990) looks at the decision to become a teacher among college graduates in the United Kingdom and finds that expected outside wages play a major role in the decision and also finds substantial momentum among teachers to remain in the education sector. Using U.S. data from Missouri, Podgursky et al. (2004) examine the entry and exit behavior of individuals in the teaching profession and find that the scholastic aptitude of individuals, as measured by ACT scores, is negatively correlated with both the decision to become a teacher and the decision to remain in teaching, with the effect largest in math and science. In the United States, it has been found by Ballou (1996) and Podgursky et al. (2004) that undergraduates of high academic quality are more likely to choose teaching careers when salaries increase. However, Ballou also shows that individuals in math and science, those with high standardized scores, and those at high quality undergraduate institutions do not seem to receive significantly different wage offers in the teaching job market for public schools than their colleagues with lesser credentials.

Some authors have looked specifically at the effect of unions and incentives and have found teacher unions to have negative effects on both teacher quality and student outcomes. Hoxby (1996) looks at the effect of state collective bargaining legislature and finds that presence of teacher unions increases the expenditures of a school and decreases the productivity for an overall net negative effect on student performance, as measured by drop-out rates. If the decreased wage flexibility imposed by unions has a negative effect, Figlio (2002) examines whether this effect can be offset through the use of teacher incentive programs in public schools. He finds some evidence that higher quality teachers respond to incentive programs, although it is unclear whether this finding is the result of teachers working harder in schools with incentives or schools with incentives attracting and retaining higher quality teachers. In a subsequent study, Figlio and Kenny (2007) again look at the effect of incentives and find student outcomes are improved in schools with incentive programs. Again, it is unclear whether this is the effect of increased effort or improved quality of hires. One theoretical paper that has examined the potential impact of wage rigidity in the education labor market is provided by Gilpin and Kaganovich (2008), who present a model of the implications of changes in teacher quality over time and show that rigid wage structures can lead to inefficiently low levels of teacher quality. In addition, they show that these wage rigidities can explain the decline in teacher aptitude over time observed by Hoxby and Leigh (2004) and Corcoran et al (2002).

This body of evidence points to two main findings. First, higher quality teachers are able to contribute to better student achievement outcomes, particularly in math and science subjects.

Second, the teaching market does not seem to reflect this finding with any sort of wage premium for individuals with strong academic backgrounds or high standardized test scores. It would appear that a key factor is the influence of teacher unions and collective bargaining agreements. While some have separately examined the impact of incentive programs and teacher unions on teacher quality, to our knowledge no one has jointly explored the full range of potential wage frictions in a unified way, nor has anyone explored the differential effect such frictions may have across various academic subjects. The contribution of this paper therefore is to provide a unified analysis of the ways in which wage rigidities influence class sizes and teacher quality and to highlight the differences of these effects between math and science and humanities subjects. We do this by making use of a new combined data set and model that is able to incorporate the simultaneous relationship between the wages, quantity, and quality of teachers hired within a school. We further examine the ultimate impact of these wage policies by relating them to student achievement outcomes within the school to see the true costs and benefits on education production. An additional contribution is made by expanding the analysis to a national sample of teachers and schools, whereas most previous literature has focused on a particular state or district. Further, we incorporate public, private, religious, and charter schools rather than focusing on a single sector as is common in many studies, and thus we hope to provide a more complete picture of the U.S. education system.

II. <u>Motivating Theoretical Model:</u>

In order to provide a framework in which to examine the impact of educational wage rigidities, we first provide a theoretical model for the staffing decisions of schools. We will assume that the objective of a school is to produce the highest education quality possible given the available budget. Schools combine teacher inputs to educate students in two areas, the sciences and the humanities, and face the following problem:

$$\max_{w_s, w_h, N_s, N_h} E(Q_s, Q_h) \tag{1}$$

$$s.t.$$

$$B = w_s N_s + w_h N_h \tag{2}$$

$$Q_S = f(\quad_{N^s} A_s(w_s), N_s) \tag{3}$$

$$Q_h = f(\quad_{N^h} A_h(w_h), N_h) \tag{4}$$

where B is the budget, E is the total education quality, Q_j is the quality of education in subject j, w_j is the wage paid to a teacher in subject j, N_j is the number of teachers in subject j, and A_j is the ability of a teacher in subject j. Thus, for a given budget, schools will choose the quantity and quality of teachers in each subject area. For simplicity, we ignore the possibility of physical capital and fixed costs in the production of education and leave these issues for future extensions of the model. We assume that education quality is increasing in each input and assume diminishing returns in each input. Schools maximize their quality output by adjusting the resources expended on teachers in both subject areas. The teacher input can be increased either by hiring more teachers or by hiring "better" teachers.

In addition to this model of school hiring decisions, we will assume that teachers face a decision about the school at which they agree to teach. Potential teachers will face the following problem:

$$\max\{U_j(w_j, l_j, v_j) \;\; j \quad J, U(O)\} \tag{5}$$

where U_j is the indirect utility of a teacher at school j, which is a function of the wage (w), workload (l), and environment (v) they would face at that school, and U(O) is the indirect utility they would get from a job outside the education sector.

Given the framework outlined above for modeling the behavior of schools and teachers, we see that an equilibrium in the market for teachers is characterized by a fairly complex set of simultaneous decisions by schools and teachers. The model allows us to make some predictions about the effect of imposing a uniform salary structure across subjects. The basic structure utilized here is similar to that outlined by Gilpin and Kaganovich (2008), which is an excellent source for a more complete theoretical analysis than is provided here. To obtain our predictions, we first make a few simplifying assumptions about the structure of the school education function and assume a Cobb-Douglas form so that the school's problem becomes:

$$\max_{w_s, w_h, N_s, N_h} \left(A_s \; N_s \;\right)\left(A_h \, N_h \;\right) \tag{6}$$

$$s.t.$$

$$B = w_s N_s + w_h N_h \tag{7}$$

We further simplify the problem by assuming that both subjects are weighted equally in the school's education production function so that = and = . This produces the education quality maximizing solution that

$$\frac{w_s^*}{w_h^*} = \frac{\partial A_s \big/ \partial w_s}{\partial A_s \big/ \partial w_s} \frac{A_h}{A_s} \tag{8}$$

which implies that if the elasticity of ability with respect to wage is greater for math than science, then the optimal solution for the school is to pay science teachers more than humanities teachers and to hire a larger number of humanities teachers than science teachers. If the problem is altered so that $w_s = w_h = \overline{w}$, then the optimal solution for the school is to hire relatively more science teachers than in the case with no wage rigidity and hire science teachers of relatively lower quality.

In the next section, we will attempt to empirically test a reduced form version of the model to determine the degree to which schools respond when affected by these wage frictions. Specifically, we attempt to see which of the dimensions the schools make the greatest compensations in, and how big these differences are in schools with and without salary rigidity. We then turn to student outcomes to see how these differences may ultimately affect the level of education quality of the school.

III. Data and Methodology:

The goal of this study is to determine the effect of wage frictions on teacher quality within a school and the subsequent effect on student outcomes. To do this, we must consider how these frictions will affect both teacher supply and teacher demand at a school. Our theoretical model predicts that schools will make a simultaneous decision in choosing the number and quality of teachers according to the limits of their budget constraint and the wage that teachers of each given type are willing to accept. For their part, teachers will choose to accept a given offer based on a combination of wages and non-pecuniary school characteristics. We thus estimate a reduced form, simultaneous equations model of teacher scholastic aptitude, class size, and wage as a function of school characteristics, teacher characteristics, and community characteristics. In this section we describe the variable used to measure each of these sets of characteristics.

The primary data for this study comes from the Schools and Staffing Survey (SASS) conducted by the National Center for Education Statistics. This survey has been conducted every four years since 1985 and incorporates questionnaires from roughly 50,000 teachers in 10,000 public schools and 10,000 teachers in 3,000 private schools in each wave. In addition, a one-year follow up of teachers is conducted that interviews those who are still teaching as well as those who have left teaching since the initial survey period.

One of the advantages of the SASS data is that this is a large micro data set with information on many educational aspects including teacher demographics and school demographics (as well as principal, district, and library information). Furthermore, this data includes public, private, religious, and charter schools. Using this data, we obtain a set of control variables intended to describe both the school's decision to hire a particular teacher and a teacher's decision to accept a position at a particular school. Our theoretical assumptions are that a teacher will choose a school based on wage, workload, and non-pecuniary amenities available at a school. The SASS provides exact wage and benefit data and allows us to construct several workload measurements including the number of students a teacher teaches, their number of classes or preparations, and the length of the school day and year. To describe the non-pecuniary environment of a teaching position, we use control variables such as the racial and gender diversity of the school, poverty measures such as the number and percentage of school lunch recipients, and measures of the teaching environment, which include safety measures such as the number of attacks on teachers and the presence of metal detectors, as well as academic factors such as team teaching, block scheduling, and the opportunity for AP and honors classes. We also construct a "troubled school" dummy variable, which takes a value of one if teachers have been attacked or threatened in the past year or if there are metal detectors present in the school.

There are also limited student outcomes available in the SASS data, including the graduation rates of the school and the percentage of students who go on to two- and four-year colleges. While these are not teacher-specific student outcome variables and may be an imprecise measurement of student achievement, we feel that they are representative of whether or not a school has achieved its assumed goal of educating its students.

One empirical issue we face in this model is in quantifying teacher quality. In this paper we will focus on the aspects of quality that relate to a teacher's scholastic aptitude. While we recognize that there are other components that contribute to the quality of a teacher's performance in the classroom, previous findings in the literature such as Greenwood et. al (1996) and Hanushek (1981, 1986, 2005) have suggested that the largest effects on student achievement among measureable teacher credentials come from a teacher's scholastic ability. We do, however, also attempt to control for other factors such as experience and subject knowledge. In addition to the SASS data, we use data from the restricted version of the National Post-Secondary Aid Survey (NPSAS) in order to form a measurement of a teacher's scholastic ability. We do this by determining the average quality of majors within undergraduate institutions in the NPSAS. We use five years of data from this survey covering the period from 1989-2004. This is a large micro-data survey of college students that includes data on individual ACT and SAT scores as well as student majors. Using this information we construct an average ability score of each category of majors (sciences, arts, humanities, education, and business) within each institution. This information is then matched with the teacher data in the SASS survey to construct a measurement of the quality of each teacher's undergraduate degree. As a measurement of a teacher's scholastic ability, we believe that that the inclusion of this data provides a significant improvement over previous attempts in the literature. Most studies opt to use either the selectivity level of a college or a college ranking value such as those found in U.S. News and World Report or Barron's. While our measure is similar, it allows a more quantitative measurement and also addresses the concern that there may be significant differences in the quality of majors even within the same college. Our data tends to support this view as we find the variance of scholastic aptitude to be higher across majors than within majors in our sample.

We further augment our data by using community information from the 2000 U.S. Census. We use the Census data to construct information on the population, median and per capita income, MSA status, average education level, and local school funding level for each community. This information is then matched with the SASS data at the district and zip-code level. These data serve two purposes in our model. First, they are used as a proxy for school budgets and second, as possible factors in the utility of teachers that may impact their decision to accept a position in a particular community.

Finally, we use information from a new source, the National Council on Teacher Quality's Teacher Roles, Rights, and Responsibilities (TR^3) database, which provides detailed information regarding the collective bargaining agreements used by teachers and schools across the nation. This is the first such data to be available and provides detailed provisions regarding the permissibility and/or requirement of various bargaining topics including teacher wages and class sizes in all fifty states and the District of Columbia. A detailed description of the data construction is outlined in an appendix at the end of the paper. Summary statistics for the teacher, school, and community variables are found in Tables 2-4.

Particular attention should be drawn to a few important variables, the first of which are our measures of wage frictions. There are several possible ways to measure the degree of unionization or wage friction in each school. These variables include the presence of a salary schedule, a collective bargaining agreement, and the use of incentive pay. Each of these categories represents a different type of wage friction.

Salary schedules impose the restriction that each teacher with a given experience level and education level will earn a particular wage. These schedules typically hold for all teachers in a school or district, regardless of subject area. Some schools have incentive programs that allow some flexibility to pay teachers differently even within a salary schedule system. Incentives may be given to teachers of particularly high quality or to hire and retain teachers in a particular field. In our analysis we will assume that a school with incentives is one with a less rigid wage structure than those without incentives. Some caution must be taken in this assumption due to the fact that, in some schools, it is common to give the incentive pay to a very large proportion of the teachers equally, in which case this is more of a shift in the schedule than a decrease in rigidity. Due to the limitations of the data, we are not able to make a distinction between selective incentive pay and less selective incentive pay, however a more detailed study of these differences is available in Figlio et al. (2007) who construct a unique data set based on their own survey of schools. To the extent that the schools in our sample use less selective incentives, our results would tend to underestimate the true effect of more selective incentive programs.

The third possible rigidity comes through collective bargaining agreements. Collective bargaining by public school teachers is legally required to at least some degree in thirty-six of the fifty states. A list of those states in which there does not exist mandatory collective bargaining legislature is found in Table 5. If teachers within a school are part of a collective bargaining agreement, they will collectively agree on a given salary schedule as well as a set of working conditions. Collective bargaining does exist in some of the states in Table 5, although it is voluntarily negotiated at the district level rather than being mandated by state law. We assume that the largest effect of collective bargaining agreements comes through the impact of the salary schedules that are imposed. However, since schools without collective bargaining can also have salary schedules, we explore the possibility that collective bargaining may have additional effects. One possible effect on teacher quality could be seen if teachers within a union engage in a type of self-policing or have positive peer effects in improving the teaching quality of their colleagues. The effects could also be negative if the additional power gained by the agreement leads to inefficiency. We expect this effect could be large in the determination of class size if bargaining agreements also include components that impose restrictions on teacher workloads. By utilizing the detailed data on the requirements of these agreements found in the TR^3 database, we explore each of these possibilities in a way that was not possible with previous data.

Contained in Table 6 is a summary of each of the categories of wage frictions across school types. It should be noted that the breakdown of wage frictions does not perfectly math the public vs. private school breakdown. Many private schools have salary schedules, and many public schools are prohibited from entering into collective bargaining agreements. Therefore, in our data schools in each sector may be subject to each of these possible frictions.

IV. <u>Econometric models and results:</u>

As described above, the primary purpose of this study is to estimate the treatment effect of union and non-union wage frictions on the quality and quantity of teachers within a given school. Because our theoretical model assumes that wage, quality, and quantity are jointly determined, we will make use of a simultaneous equations framework that will account for this potential endogeneity. The model that we will estimate is a recursive structural treatment effects model and is described as follows:

$$
\begin{aligned}
Y_{SAT} &= {}_{0} + X_T\ {}_{1} + X_S\ {}_{2} + X_C\ {}_{3} + T\ {}_{4} + Y_{Wage}\ {}_{1} + Y_{ClassSize}\ {}_{2} + u_{SAT} \\
Y_{Wage}(\ ) &= {}_{0} + X_T\ {}_{1} + X_S\ {}_{2} + X_C\ {}_{3} + T\ {}_{4} + Y_{CS} + Z_W\ {}_{W} + v_W \\
Y_{ClassSize}(\ ) &= {}_{0} + X_T\ {}_{1} + X_S\ {}_{2} + X_C\ {}_{3} + T\ {}_{4} + Z_{CS}\ {}_{CS} + v_{CS}
\end{aligned} \qquad (9)
$$

where Y_{SAT} is the scholastic aptitude of a teacher at a given school, Y_{Wage} is the wage of the teacher, and $Y_{ClassSize}$ represents average number of students in classes taught by the teacher. X_T, X_S, and X_C represent vectors of teacher, school, and community characteristics, T is the vector of treatment effects for our measures of wage frictions, and Z_W and Z_{CS} are the instrumental variables associated with the wage and class size equations. We initially assume that the error terms u_{SAT}, v_W, and v_{CS} are stochastically independent and identically distributed and will later relax this homogeneity assumption to allow for possible correlation in errors within schools. If there are unobserved factors that are common to teachers in a given school, we would expect to see this type of heterogeneity, and we estimate a version of the model that is robust to this type of clustered errors. It should be noted that we have assumed a recursive triangular structure for our system of equations, as outlined originally in Strotz and Wold (1960) and more recently in Ma and Koenker (2004) and several other papers. In this structure we assume a "causal chain" in which class size, or defined differently, the number of teachers hired, affects the wage, and both the wage and number of teachers affect the quality of the teachers hired. This assumption has both a practical and a theoretical purpose. Practically, such a structure greatly simplifies the error structure and makes identification much easier. Theoretically, we believe this model accurately describes the hiring process of a school. First, school administrators determine their staffing needs and decide on the number of teachers to be hired. Second, the wage that can be offered is determined by the budget and number of teachers. Finally, the actual quality of the teachers is determined as an ultimate outcome given the wage being offered and the number of teachers hired.

Our first results, shown in Table 7, present the baseline OLS regressions under the assumption that there is no endogeneity between the three equations. Reported standard errors were computed using a robust estimation clustered by school. The coefficients of particular interest are the wage friction variables. We find that salary schedules reduce the average aptitude of teachers in a school by 18 SAT points, or roughly 2%. We find that this effect can be offset positively by the use of incentive pay, which increases aptitude by approximately 11 SAT points or 1%. Each of these coefficients is significant at the 95% level or more. In this case, contrary to our theoretical model, we do not find evidence that the presence of salary schedules leads to a significant difference between the quality of science vs. humanities teachers. Nor do we find that they lead to a difference in the sizes of science vs. humanities classes as our estimate is small and insignificant. We do find, however, that schools with salary schedules have an average of 2.2 more students per class, or roughly 10% larger classes.

Other coefficients of note show that private schools have significantly higher aptitude teachers on average (54 SAT points), but religious private schools have significantly lower average aptitude than secular private schools (59 SAT points). We find that the coefficient on teacher experience is negative, which could have two possible explanations. First, it could be that newer or younger teachers have higher aptitude than those who entered teaching before them. Second, it could be that as attrition of teachers occurs, it is the highest aptitude teachers who are most likely to leave teaching. With our cross-sectional data, we cannot sort out these possible effects and in future research we will need to make use of the teacher follow-up panel data to determine which effect is dominating. To test the validity of these OLS results, we first use an Anderson-Rubin test to check that the endogenous regressors are significant in the equation for teacher ability, and we reject at the 99% significance level the hypothesis that they are not significant. We then make use of the Wu-Hausman, Durban-Wu-Hausman, and Anderson-Rubin tests of endogeneity, and each case we reject at the 99% level that the hypothesis that class size and wage are exogenous variables in the scholastic ability regression. These statistics are summarized in Table 8. Because of this, we next turn to our recursive structural model and instrumental variables approach, which will account for this endogeneity.

In Table 9 we present the results of the Three Stage Least Squares (3SLS) regression with robust clustered standard errors, this time allowing for endogeneity among the three dependent variables. We find the OLS coefficients are generally robust to this specification, with only minor changes in the magnitude and significance of our coefficients, which in general become more precise. For our instruments in this specification, we use the collective bargaining legislation for each state. In the equation for wages, our instrument is an indicator that is equal to one for schools in states that have mandatory collective bargaining on the topic of wages. Similarly, for our class size equation our instrument is an indicator for schools in states with mandatory collective bargaining on the topic of class sizes. We test the validity of these instruments in several ways. The first is a test of orthagonality in the equation for teacher ability in which we fail to reject the hypothesis that our instruments are uncorrelated to teacher ability. Next, we test whether our model is identified by the instruments using a variety of tests. The Anderson Canonical Correlation statistic and Cragg-Donald statistic both reject, at the 95% level, the hypothesis that the model is under-identified. However, these calculations are based on an assumption of homogeneity, and thus neither of these statistics is fully valid for the cluster robust formulation of the model. We therefore also calculate the robust Kleibergen-Paap rk statistic, which again rejects, at the 99% level, the hypothesis that the model is under-identified. To test for weak identification we use the Stock and Yogo test and the procedures for robust testing outlined in Baum et al (2007) and reject the hypothesis of weak instruments at the 99% level. A summary of these tests is included in Table 10.

In each of these models, we fail to find that wage frictions played a significant role in causing any differences in teacher ability or class size across different academic subject areas. However, our theoretical model predicts that such frictions would be most binding for the highest ability teachers or in schools that hire the greatest number of teachers. We therefore turn our attention to a quantile regression framework that will allow us to examine the effects of wage rigidities at different points within the distribution of ability and class sizes.

The econometric model we use for this estimation is a recursive structural quantile treatment effects (QTE) model, which is given by the following form:

$$
\begin{aligned}
Y_{SAT}(\ ) &= {}_0(\ ) + X_T\ {}_1(\ ) + X_S\ {}_2(\ ) + X_C\ {}_3(\ ) + T\ {}_4(\ ) + Y_W\ {}_1(\ ) + Y_{CS}\ {}_2(\ ) + u_{SAT}(\ ) \\
Y_W(\ ) &= {}_0(\ ) + X_T\ {}_1(\ ) + X_S\ {}_2(\ ) + X_C\ {}_3(\ ) + T\ {}_4(\ ) + Y_{CS}\ (\ ) + Z_W\ {}_W(\ ) + v_W(\ ) \qquad (10) \\
Y_{CS}(\ ) &= {}_0(\ ) + X_T\ {}_1(\ ) + X_S\ {}_2(\ ) + X_C\ {}_3(\ ) + T\ {}_4(\ ) + Z_{CS}\ {}_{CS}(\ ) + v_{CS}(\ )
\end{aligned}
$$

where the variables are as described in the first recursive treatment effects model (equation 9) but where $\boldsymbol{\tau}$ represents a given quantile of the distribution for the dependent variable. We follow the three stage control variate approach as oulined by Ma and Koenker (2004), which is shown to be the most asymptotically efficient estimator. The main difference in the control variate approach is that in every stage of the IV estimation, the fitted value is computed at each quantile, which Ma and Koenker show provides a dramatic improvement in bias over traditional quantile regressions and multi-stage quantile regressions estimated at the mean or median value only. An excellent analysis and comparison of alternative methods can be found in their paper.

The control variate quantile treatment effects regression is estimated using the same set of control variables and robust standard errors are estimated using 500 bootstraps at the 25%, 50% and 75% quantiles. The abbreviated results of key coefficients are reported in Table 11 and 12. Using the quantile framework, we find that salary schedules have the strongest negative effect at the high end of the teacher aptitude distribution and additionally find that math and science teachers under a salary schedule on average have significantly lower ability levels than their peers who are not under a salary schedule. We find that among the teachers at the 75% percentile of scholastic ability, the presence of a salary schedule reduces the ability of teachers by 26 points or roughly 2.5%, and reduces the ability of math and science teachers by an additional 3.5%. For class sizes, we find that salary schedules tend to increase class sizes by between 2 and 2.5 students per class, or roughly 10% for all quantiles, with an additional 0.5 students or 2.5% increase in states with mandatory collective bargaining on class size. We find that the additional effect for science classes is only seen for schools with relatively large class sizes, or to put it another way, relatively few teachers. At the 75^{th} quantile, science classes in schools with rigid wages have an additional 0.32 students, meaning that they are approximately 2% larger than in schools with flexible wage structures. In the quantile regressions, we therefore find evidence to support our hypothesis that the market for math and science teachers is differently affected by wage frictions than the market for arts and humanities teachers, but this seems to occur only at the high end of the ability distribution. This is to be expected because it is the highest ability individuals who would face outside wage options that are most significantly different than those wage opportunities in the education sector.

Finally, in order to gauge the ultimate impact of these effects on teacher quality and class size we turn to student outcomes to see how the various levels of wage rigidity might impact the graduation and college attendance rates of a school. First, in Table 13 we show the results of a regression of graduation rates on school and community explanatory variables as well as average teacher aptitudes and class sizes. This is done using a Tobit regression framework with limits between 0 and 100 and bootstrapped standard errors. We use the predicted values from the simultaneous 3SLS regression in order to eliminate the potential endogeneity of these variables. It is possible that a school with a high graduation rate may be attractive to teachers and may thus influence a teacher's decision to accept a position at the school. If this is the case, there would be an endogeneity bias, and so we instead use the predicted values to eliminate this potential problem. We find that graduation rates do not seem to be significantly impacted by either class sizes or teacher abilities. This result is also robust to the quantile estimation shown in Table 14, with the finding that increased class sizes have a small negative effect in schools that have relatively large classes. At the 50th and 75th quantiles, a 10% increase in class sizes leads to a decrease in graduation rates of approximately 2%, however these estimates are only significant at the 90% level.

While class sizes and teacher quality seem to have little effect on graduation rates, they do seem to have significant effects on our other measure of student achievement, the attendance rates of four-year colleges after high school. We repeat our bootstrapped Tobit regression with college attendance as the dependent variable in Table 15 and find that, on average, a 10% increase in teacher aptitude results in an increase of college attendance rates of roughly 2%. To put the wage rigidity effects in perspective, recall that salary schedules lower average teacher aptitude by approximately 2%. We can then predict that through the channel of teacher ability, salary schedules may lower college attendance rates as much as 0.5-1%, which is a relatively large magnitude when compared with other coefficients in the regression. We also examined the student outcomes in the quantile framework and found that for college attendance rates this effect is nearly doubled at the 25th percentile, which represents schools with relatively low rates. The quantile results are shown in Table 16. For class size, we find no significant effects on college attendance in the standard Tobit regression, although we find large and significant decreases on attendance rates at the 50th and 75th quantiles. Thus, the results suggest that an increase in class size for schools with relatively large classes has a particularly strong adverse effect to the extent that at the 75th quantile, a 10% increase in class size results in a 1.5% decrease in college attendance rates.

V. <u>Conclusion & Future Extensions</u>

In this study we have use of a new data set and new econometric techniques to examine the effects of union and non-union wage rigidities on teacher quality and student achievement outcomes. We find evidence to support the hypothesis that when schools are faced with rigid wage structures that limit their ability to optimally choose teacher quality and quantity, they are more likely to make adjustments in the quality dimension than the quantity dimension. Average scholastic aptitude of teachers in schools with salary schedules was found to be roughly 2.5%

lower than those in schools without salary schedules. The use of incentive pay and its associated increased wage flexibility results in 1% higher average ability teachers. In addition, we find that the use of salary schedules has the strongest negative impact on ability for teachers at the high end of the ability distribution, where we see an additional 3.5% reduction in scholastic ability among math and science teachers. Wage rigidity also seems to have an effect on class sizes, where we see 10% larger classes in schools with salary schedules and an additional 2.5% in schools where the class size is a mandatory subject of collective bargaining.

We further find that our measures of teacher quality and class size have significant effects on student outcomes as measured by graduation and college attendance rates. We estimate that the 2.5-5.5% reduction in teacher ability and 10-12% increase in class size caused by salary schedules could translate to reductions in graduation rates of over 2% in schools with large classes. We also estimate that the effects of salary schedules could translate to even larger reductions in college attendance rates, perhaps as much as 2.5-3%. These results suggest an interesting policy implication in which schools with different student achievement goals might be best served by undertaking very different policy initiatives. Schools with low graduation rates may improve by focusing on reducing class sizes while schools hoping to increase college attendance may want to focus on improving teacher quality. These questions require further study, however these results suggest this may be an important area for extension and policy analysis.

This study presents several other opportunities for extension, and we hope to enhance our analysis in several dimensions in future work. First, by utilizing additional data from subsequent survey periods, we plan to explore possible changes in the effects of wage policies over time. Some theoretical models such as Gilpin and Kaganovich (2008) predict that increasing outside wage opportunities will lead to a decrease in teacher quality over time. We hope to be able to test this prediction using some of the panel and repeated cross-section opportunities afforded by the SASS data set. With the larger data set we will also be able to test possible endogeneity of teacher and school characteristics by restricting our sample to only new teachers. In addition, we plan to extend the theoretical and empirical model to include issues of physical capital. One might expect that if teacher wages are becoming relatively costly as capital and technology prices are becoming relatively cheap, schools may choose to substitute capital for labor in the production of education. This question, in addition to the question of whether this new technology tends to be a complement or substitute to teacher quality, is one we hope to address in future extensions. Finally, we hope to extend our analysis to questions of teacher retention and attrition and examine the effects of union and collective bargaining policies in these areas.

Appendix A: Tables and figures

Table 1: NACE Fall 2007 Salary Survey (Select degrees)

Discipline (bachelor's degree level)	Average
Chemical Engineering	$59,218
Computer Science	$53,051
Bioengineering & Biomedical Engineering	$51,044
Economics (Business/Managerial)	$47,782
Mathematics / Statistics	$46,547
Political Science / Government	$35,261
History	$35,092
Education	$33,679
English Language & Literature / Letters	$31,924
Sociology	$32,161

Source: National Association of Colleges and Employers (NACE) Fall 2007 Salary Survey

Table 2: Summary Statistics -- Teacher Variables

Variable	Obs	Mean	Std. Dev.
Scholastic Aptitude	12034	1018.116	117.601
Experience (yrs)	12034	13.989	10.357
Avg. Class size	12034	22.047	9.690
Number of Classes (per day)	12034	5.012	2.002
Number of Preparations (per day)	12034	1.991	1.104
Wage+Benefits (Thousands)	12034	48.387	17.688
Non-school earnings (Thousands)	12034	1.274	4.688
BA degree in subject	12034	0.309	0.462
Education BA degree in subject	12034	0.168	0.374
Advanced degree in sub (MA, Phd)	12034	0.071	0.257
Advanced ed degree (MA, Phd)	12034	0.307	0.461
Full certification in subject	12034	0.430	0.495
Partial certification in subject	12034	0.049	0.217
African American	12034	0.046	0.210
Hispanic	12034	0.037	0.189
Asian	12034	0.020	0.139
Native American	12034	0.020	0.141

Male	12034	0.463	0.499

Table 3: Summary Statistics -- School Variables			
Variable	**Obs**	**Mean**	**Std. Dev.**
Number of students (thousands)	12034	1.006	0.772
Private (secular and religious)	12034	0.092	0.289
Religious	12034	0.071	0.256
Charter	12034	0.036	0.187
Full certification in subject required	12034	0.743	0.437
Major/minor in subj required	12034	0.624	0.484
% African-American	12034	0.123	0.217
% Hispanic	12034	0.092	0.180
% Asian	12034	0.031	0.088
% Male	12034	0.494	0.103
% Eligible for School Lunch	12034	0.278	0.263
Gifted program	12034	0.672	0.469
Team Teaching	12034	0.445	0.497
Block Scheduling	12034	0.533	0.499
Year-round classes	12034	0.036	0.185
Troubled School	12034	0.116	0.320

Table 4: Summary Statistics: Community Variables			
Variable	**Obs**	**Mean**	**Std. Dev.**
Total Population (1000's)	12026	176.7	594.7
% BA degree	12026	0.47	3.13
Median Income (1000's)	12026	49.02	15.71
Per-capita Income (1000's)	12026	20.31	6.83
% local funding	12026	17.82	150.82
Metropolitan Statistical Area (MSA)	12026	0.66	0.48

Table 5: States without mandatory collective bargaining laws for teacher unions		
Arizona	Mississippi	Utah
Arkansas	Missouri	Virginia
Colorado	New Mexico	West Virginia
Georgia	North Carolina	Wyoming
Kentucky	South Carolina	

Louisiana	Texas	

Table 6: Degree of Wage Friction by School Type – (# of teacher observations)

Salary Schedule	Incentives or Rewards	Public	Private - Secular	Private - Religious	Charter	Total	Percent
No Collective Bargaining							
Yes	No	3,159	101	675	146	4,081	34%
Yes	Yes	888	30	47	59	1,024	9%
No	Yes / No	78	130	127	152	487	4%
Collective Bargaining							
Yes	No	5,417	--	--	74	5,491	46%
Yes	Yes	945	--	--	6	951	8%
		10,487	261	849	437	12,034	100.0

Table 7: Separate OLS Regressions									N=12,034
	Scholastic Ability (SAT)			**Wage ($1000s)**			**Class size (#students)**		
		Clst robust			**Clst robust**			**Clst robust**	
Variables	**Coeff**	**Std Err**		**Coeff**	**Std Err**		**Coeff**	**Std Err**	
Salary Schedule	**-17.235**	**8.122**	**	**2.904**	**1.028**	***	**2.192**	**0.647**	***
Incentives	**11.411**	**3.543**	***	**-0.493**	**0.630**		**-0.581**	**0.310**	*
Math/Science	**33.433**	**4.819**	***	**-0.871**	**0.562**		**-0.787**	**0.385**	**
Sal Sch*Math/Sci	**-1.130**	**5.171**		**0.354**	**0.584**		**-0.155**	**0.427**	
Teacher Variables:									
Experience	-0.549	0.117	***	0.713	0.016	***	0.021	0.008	***
# of Preparations	0.030	1.010		0.280	0.137	**	-0.861	0.085	***
Non-school income	0.641	0.211	***	-0.122	0.042	***	-0.018	0.018	
Degree in subject	73.336	3.003	***	-0.305	0.361		-0.282	0.220	
Ed Deg in subject	-12.914	3.018	***	-1.025	0.379	***	0.359	0.232	
Adv deg in sub	11.959	4.504	***	4.858	0.586	***	-1.263	0.314	***
Adv ed degree	3.652	2.305		5.057	0.321	***	-0.914	0.185	***
Cert in Subject	-8.876	2.987	***	-0.254	0.355		0.589	0.225	***
Partial Cert in Subj	-13.750	5.368	***	-1.357	0.580	***	0.673	0.378	*
African-American	-83.691	7.236	***	0.356	0.749		0.918	0.454	**
Hispanic	-17.568	6.310	***	-0.468	0.629		0.950	0.569	*
Male	9.610	2.091	***	4.095	0.270	***	1.985	0.171	***
School Variables:									
School Size	0.137	1.931		2.041	0.339	***	3.743	0.192	***
Private	54.237	10.729	***	-3.093	1.409	**	-5.947	0.597	***
Religious	-59.410	11.250	***	-6.861	1.411	***	5.822	0.612	***
Charter	5.388	7.468		-2.732	1.043	***	2.474	0.820	***
Cert Required	-4.319	2.796		-0.175	0.480		0.001	0.242	
Maj/Min in sub req	-4.157	2.451	*	0.403	0.429		0.007	0.201	
School: % Afr-Am	-13.535	6.946	**	7.218	1.147	***	0.310	0.525	
School: % Hispanic	-27.316	8.714	***	8.584	1.377	***	0.557	0.718	
School: % male	-25.635	12.330	**	-2.404	2.002		-2.645	0.879	***
School: % School Lunch	-16.780	6.196	***	-0.496	0.975		-1.906	0.459	***
Gifted programs	-5.792	2.559	**	-0.915	0.448	**	0.377	0.210	*
Block Scheduling	7.296	2.313	***	-0.067	0.404		0.165	0.189	
Team Teaching	3.686	2.361		0.922	0.414	**	-0.195	0.192	
Year round school	-14.007	6.587	**	-1.042	1.080		-0.167	0.554	
Troubled School	-3.189	3.416		0.915	0.468	**	-0.308	0.300	
Community Vars:									
% local funding	-0.040	0.028		-0.004	0.009		-0.003	0.002	
Median Income	0.516	0.089	***	0.282	0.020	***	-0.003	0.008	
% with BA	2.596	1.452	*	0.089	0.447		0.119	0.095	
Total Population	0.004	0.002	*	0.001	0.000	***	0.000	0.000	***
Class size	-0.119	0.107		0.034	0.016	**	---	---	
Wage	-0.129	0.076	*	---	---		---	---	
Col Barg: Wages	---	---		3.431	0.427	***	---	---	
Col Barg: Class Sz	---	---		---	---		0.630	0.259	***
Constant	1025.35	11.746	***	13.916	1.858	***	18.575	0.952	***
		R-sq= .2751			R-sq= .4051			R-sq=.2603	
*= 90% significance level, **=95% level, and ***=99% level									

Table 8: Tests of Endogeneity of Class Size, Wage in Regression of Scholastic Ability		
H0: Regressors are exogenous		
	Statistic	**P-value**
Wu-Hausman F test:	12.099	0.00001
Durbin-Wu-Hausman chi-sq test:	24.229	0.00001
Anderson-Ruben F test	8.84	0.00001
Anderson-Ruben Chi-sq test:	17.74	0.00001

Table 9: 3SLS Regression						N=12,034
	Scholastic Ability (SAT)		**Wage ($1000s)**		**Class size (#students)**	
		Clst robust		**Clst robust**		**Clst robust**
Variables	**Coeff**	**Std Err**	**Coeff**	**Std Err**	**Coeff**	**Std Err**
Salary Schedule	**-17.902**	**5.957** ***	**2.979**	**0.761** ***	**2.192**	**0.495** ***
Incentives	**11.608**	**3.190** ***	**-0.511**	**0.409**	**-0.582**	**0.266** **
Math/Science	**33.659**	**4.717** ***	**-0.896**	**0.603**	**-0.787**	**0.392** **
Sal Sch*Math/Sci	**-1.196**	**5.086**	**0.359**	**0.650**	**-0.155**	**0.423**
Teacher Variables:						
Experience	-0.644	0.101 ***	0.714	0.013 ***	0.021	0.008 ***
# of Preparations	0.087	0.987	0.251	0.126 **	-0.861	0.082 ***
Non-school income	0.659	0.211 ***	-0.123	0.027 ***	-0.018	0.018
Degree in subject	73.398	2.829 ***	-0.315	0.361	-0.281	0.235
Ed Deg in subject	-12.825	3.109 ***	-1.013	0.397 ***	0.359	0.258
Adv deg in sub	11.477	4.115 ***	4.815	0.526 ***	-1.263	0.342 ***
Adv ed degree	3.088	2.236	5.025	0.286 ***	-0.914	0.186 ***
Cert in Subject	-8.911	2.847 ***	-0.234	0.364	0.589	0.237 **
Partial Cert in Subj	-13.695	4.984 ***	-1.337	0.637 **	0.673	0.414 *
African-American	-83.850	5.194 ***	0.387	0.663	0.918	0.432 **
Hispanic	-17.624	5.463 ***	-0.436	0.698	0.950	0.454 **
Male	8.816	2.016 ***	4.161	0.258 ***	1.985	0.168 ***
School Variables:						
School Size	-0.604	1.598	2.167	0.204 ***	3.743	0.133 ***
Private	55.313	7.709 ***	-3.292	0.985 ***	-5.945	0.642 ***
Religious	-59.264	7.990 ***	-6.667	1.021 ***	5.822	0.664 ***
Charter	5.412	5.887	-2.647	0.752 ***	2.475	0.490 ***
Cert Required	-4.338	2.499 *	-0.177	0.320	0.001	0.208
Maj/Min in sub req	-4.177	2.160 **	0.405	0.277	0.007	0.179
School: % Afr-Am	-14.341	5.665 ***	7.236	0.730 ***	0.309	0.471
School: % Hispanic	-28.457	6.843 ***	8.613	0.875 ***	0.560	0.573
School: % male	-25.062	10.298 ***	-2.495	1.316 **	-2.644	0.856 ***
School: % School Lunch	-16.491	5.077 ***	-0.560	0.648	-1.906	0.422 ***
Gifted programs	-5.683	2.214 ***	-0.900	0.284 ***	0.377	0.184 **
Block Scheduling	7.303	2.003 ***	-0.061	0.256	0.165	0.167
Team Teaching	3.567	2.036 *	0.914	0.260 ***	-0.195	0.169
Year round school	-13.855	5.382 ***	-1.046	0.687	-0.167	0.447
Troubled School	-3.282	3.189	0.903	0.407 **	-0.308	0.265
Community Vars:						
% local funding	-0.039	0.026	-0.004	0.003	-0.003	0.002
Median Income	0.477	0.076 ***	0.282	0.010 ***	-0.003	0.006
% with BA	2.565	1.269 **	0.093	0.162	0.119	0.106
Total Population	0.003	0.002 *	0.001	0.000 ***	0.000	0.000 ***
Class size	-17.902	5.957 ***	3.456	0.267 ***	---	---
Wage	11.608	3.190 ***	---	---	---	---
Col Barg: Wages	---	---	2.979	0.761 ***	---	---
Col Barg: Class Sz	---	---	---	---	0.630	0.159 ***
Constant	1021.21	9.609 ***	14.533	1.229 ***	18.573	0.801 ***
		R-sq= .2751		R-sq= .4051		R-sq=.2603

*= 90% significance level, **=95% level, and ***=99% level

Table 10: Tests of Identification and Validity of Instruments

H0: The model is under-identified	**Statistic**	**P-value**
Anderson Canon. Corr LR test (not robust)	4.19	0.0407
Cragg-Donald test (not robust)	4.19	0.0407
Kleibergen-Paap rk LM test (wage)	130.42	0.0000
Kleibergen-Paap rk LM test (class size)	7.69	0.0055
H0: Instruments are weak		
Cragg-Donald test (not robust)	5.09	0.0032
Stock & Yogo test	130.42	0.0000

Table 11: Recursive Quantile Treatment Effects Regression with bootstrapped standard errors:*

Dependent variable: Scholastic Aptitude **

Variable	**Q25 Coef.**	**Q50 Coef.**	**Q75 Coef.**
Salary Schedule	-7.71	-17.32 **	-26.73 ***
Incentives	9.64 ***	17.62 ***	9.91 **
Math/Sci Teacher	27.68 ***	33.63 ***	49.25 ***
Math/Sci * Sal Sch interaction	-0.44	-29.92 **	-34.36 ***

*200 bootstraps
** The regression also includes the teacher, school, and community control variables used in the 3SLS model. Full results are omitted for space but are available upon request from the authors

Table 12: Recursive Quantile Treatment Effects Regression with bootstrapped standard errors:*

Dependent variable: Class Size **

Variable	**Q25 Coef.**	**Q50 Coef.**	**Q75 Coef.**
Salary Schedule	1.982 ***	2.586 **	2.295 ***
Incentives	-0.600 ***	-0.428 ***	-0.357
Math/Sci Teacher	-0.723 ***	-0.328 ***	-0.696
Math/Sci * Sal Sch interaction	-0.038	-0.431 **	-0.321 *
Mandatory Collective Bargaining on class size	0.458 ***	0.504 ***	0.612 ***

*200 bootstraps
** The regression also includes the teacher, school, and community control variables used in the 3SLS model. Full results are omitted for space but are available upon request from the authors

Table 13: IV Tobit regression of Graduation Rate N=3389

Variable	Observed Coef.	Bootstrap Std. Err.	
Predicted avg teacher aptitude	**-0.008**	**0.011**	
Predicted avg class size	**-0.282**	**0.206**	
% Local school funding	0.016	0.168	
Median community income	-0.093	0.051	*
% BA degree in community	-0.355	2.498	
Total community population	0.000	0.001	
School Size	-0.513	0.992	
Private	17.103	3.499	***
Religious	-0.758	3.136	
Charter	-7.935	3.400	***
Full Certification required	-1.246	0.781	*
Major/minor in sub req	0.942	0.963	
% African-American	-12.750	2.374	***
% Hispanic	-10.729	3.230	***
% Male	-4.301	6.287	
% School Lunch	-3.060	2.371	
Gifted	4.774	0.881	***
Block Scheduling	-1.077	0.746	
Team Teaching	0.819	0.899	
Yr-round Schooling	-7.710	2.577	***
Troubled School	-1.884	1.280	
Constant	117.571	12.753	***

Table 14: Quantile IV Tobit regression of graduation rate

	Q25			Q50			Q75		
Variable	**Coef.**	**Bootstp Std. Err.**		**Coef.**	**Bootstp Std. Err.**		**Coef.**	**Bootstp Std. Err.**	
Predicted avg teacher aptitude	**-0.001**	**0.005**		**0.001**	**0.003**		**0.000**	**0.001**	
Predicted avg class size	**-0.234**	**0.049**	***	**-0.160**	**0.061**	**	**-0.103**	**0.043**	*
% Local school funding	0.001	0.005		0.000	0.001		0.000	0.000	
Median community income	-0.005	0.017		0.002	0.009		0.002	0.002	
% BA degree in community	-0.052	0.283		0.010	0.067		0.004	0.016	
Total community population	-0.001	0.002		0.000	0.001		0.000	0.000	
School Size	-1.061	0.665	*	-1.336	0.495	***	-0.951	0.024	***
Private	4.847	0.487	***	1.731	0.391	***	0.450	0.181	***
Religious	-0.221	0.693		0.212	0.281		0.192	0.174	
Charter	-13.867	9.912		-3.286	0.766	***	0.070	0.120	
Full Certification required	-0.148	0.203		-0.401	0.238	*	-0.084	0.094	
Major/minor in sub req	0.736	0.596		0.347	0.114	***	0.112	0.063	*
% African-American	-8.655	1.556	***	-4.424	1.013	***	-2.144	0.378	***
% Hispanic	-8.718	0.879	***	-1.874	0.849	**	0.184	0.195	
% Male	-1.313	1.071		-0.398	0.395		0.093	0.437	
% School Lunch	-1.686	1.689		-0.039	1.011		0.022	0.179	
Gifted	1.251	0.213	***	0.487	0.227	**	0.086	0.069	
Block Scheduling	-0.440	0.584		-0.322	0.197	*	-0.162	0.067	***
Team Teaching	0.350	0.814		0.072	0.358		0.056	0.079	
Yr-round Schooling	-8.969	7.476		-0.774	0.412	**	0.077	0.221	
Troubled School	-1.666	1.473		-0.414	0.616		0.021	0.190	
Constant	97.002	6.305	***	97.802	2.558	***	99.843	0.509	***

Table 15: IV Tobit regression of College Attendance Rate N=3389

Variable	Coef.	Bootstrap Std. Err.	
Predicted avg teacher aptitude	**0.059**	**0.009**	***
Predicted avg class size	**-0.076**	**0.291**	
% Local school funding	-0.003	0.021	
Median community income	0.182	0.045	***
% BA degree in community	0.044	1.046	
Total community population	0.002	0.001	*
School Size	5.604	1.474	***
Private	36.814	5.411	***
Religious	-4.847	5.051	
Charter	-15.357	3.214	***
Full Certification required	-1.656	1.432	
Major/minor in sub req	3.564	1.025	***
% African-American	-1.073	2.695	
% Hispanic	-20.510	3.799	***
% Male	-21.418	8.072	***
% School Lunch	-15.979	2.995	***
Gifted	5.920	1.102	***
Block Scheduling	-0.786	0.965	
Team Teaching	1.222	1.056	
Yr-round Schooling	-4.610	3.383	
Troubled School	-4.723	1.651	***
Constant	27.703	1.314	**

Table 16: Quantile IV Tobit regression of college attendance rate

Variable	Q25 Coef.	Q25 Bootstp Std. Err.		Q50 Coef.	Q50 Bootstp Std. Err.		Q75 Coef.	Q75 Bootstp Std. Err.	
Predicted avg teacher aptitude	**0.054**	**0.016**	***	**0.046**	**0.018**	***	**0.023**	**0.012**	**
Predicted avg class size	**0.142**	**0.469**		**-0.417**	**0.164**	***	**-0.545**	**0.176**	***
% Local school funding	0.033	0.026		-0.008	0.014		-0.005	0.023	
Median community income	0.156	0.070	**	0.275	0.086	***	0.236	0.061	***
% BA degree in community	-2.095	1.483		0.159	0.747		0.091	1.371	
Total community population	0.000	0.000		0.001	0.003		0.004	0.003	
School Size	6.484	2.169	***	5.948	0.562	***	4.528	0.827	***
Private	49.261	8.921	***	43.473	2.184	***	31.419	1.039	***
Religious	-10.863	11.671		-7.338	2.380	***	-2.028	2.343	
Charter	-9.613	2.912	***	-17.988	3.996	***	-18.682	6.307	***
Full Certification required	-0.601	1.375		0.057	1.596		-0.547	1.065	
Major/minor in sub req	4.166	1.207	***	2.870	0.734	***	1.984	1.022	**
% African-American	0.419	3.262		0.560	3.718		-1.312	3.909	
% Hispanic	-18.045	3.096	***	-16.758	3.328	***	-21.299	5.293	***
% Male	-19.182	5.698	***	-17.866	6.633	***	-11.566	3.663	***
% School Lunch	-12.143	1.026	***	-20.407	3.016	***	-15.980	2.421	***
Gifted	7.850	0.955	***	4.744	0.579	***	2.676	0.212	***
Block Scheduling	-0.715	0.517		0.098	0.486		-1.229	0.379	***
Team Teaching	1.428	0.590	***	1.243	1.030		1.550	1.534	
Yr-round Schooling	-6.286	3.003	**	-0.653	2.313		1.365	2.554	
Troubled School	-1.861	1.682		-1.802	2.178		-3.394	2.576	
Constant	-17.284	27.844		35.630	7.301	***	49.152	7.600	***

Appendix B: Data Construction

1) Major and field groupings:

Because of our interest in differential effects of frictions across subjects, we chose to group different subjects and degree majors into distinct categories based on what we felt were groups with similar non-education sector job opportunities. We grouped majors into the following categories: math & science, humanities, arts, education, and business/other. A breakdown of the individual majors in each broad category is given in the first table below. We followed a similar algorithm for assigning high school classes into each of these similar categories: math & science, humanities, arts, vocational and other. These classes can be found in the second table below.

Our most contentious decision was in assigning the social science group for which a compelling argument could be made to include in the math & science group. An examination of the characteristics of these teachers suggested they had more in common with the humanities group than the math & science group and so they were thus placed. However, we estimated alternative specifications as a check of robustness that included them in the science group, as well as their own separate group, and these specifications had little or no significant effects on our results.

Data Appendix Table A1: College major subject groupings

Humanities = 1	**Science = 2**	**Education = 3**	**Arts = 4**
English literature or composition	Mathematics	Early childhood education	Art, fine and applied
Communications or journalism	Statistics	Prekindergarten	Drama or theater
French	Biology/Life science	Secondary education	Visual/performing arts
German	Chemistry	Elementary education	Music
Latin	Geology/Earth science	Kindergarten	
Russian	Physics	Agricultural education	**Other = 0**
Spanish	Other natural sciences	Art education	Business and mangmt
Other languages	Ag and nat'l resources	Bilingual education	Family consumer sci
Native American studies	Computer science	Business education	All Other Areas
Humanities	Engineering	Cross-cultural education	Military science
Law		ESL education	General studies
Library and information science		English/language arts educ	Health professions
Multi- or interdisciplinary studies		Home econ education	Architecture
Philosophy		Foreign languages education	Environmental design
Religion or theology		Health education	
Other area or ethnic studies		Native American educ	
Public administration or service		Mathematics education	
Economics		Music education	
History		Physical education	
Political science and government		Reading education	
Psychology		Religious education	
Sociology		Science education	
Other social sciences		Social studies education	
		Trades and industry/industrial arts education	
		Special education, general	
		Autism	
		Deaf and hard-of-hearing	
		Developmentally delayed	
		Early childhood special educ	
		Emotionally disturbed behavior disorders	
		Learning disabilities	
		Mentally retarded	
		Mildly or moderately disabled	
		Orthopedically impaired	
		Severely disabled	
		Speech or language impaired	
		Traumatically brain injured	
		Visually impaired	
		Other special education	
		Counseling and guidance	
		Curriculum and instruction	
		Educational administration	
		Educational psychology	
		Other education	

Data Appendix Table A2: High school class groupings

Math & Science	**Humanities & Social Sciences**	**Art**
Algebra, elementary	Literature	Arts and crafts
Algebra, intermediate	Composition/journalism/creative writing	Filmmaking and photography
Algebra, advanced	English as a Second Language	Chorus
Analytic geometry	Reading	Band
Basic and general mathematics	Other English/language arts courses	Drama/theater/dance
Business and applied math	French	Music
Calculus	German	Other visual/performing arts
Geometry	Latin	
Integrated math	Russian	**Other**
Pre-algebra	Spanish	Accounting
Pre-calculus	Other foreign languages	Agriculture or natural resources
Statistics and probability	Philosophy	Business/office
Trigonometry	Religion	Career education
Other math		Child care
Biology or life science	Social studies	Communications technologies
Chemistry	Civics	Cosmetology
Integrated science	Economics	Food services
Geology/earth science/space science	Geography	Health occupations
Physics	History	Keyboarding
Other physical science	Political science/government	Trades and industry
Other natural science	Psychology	Vocational family and consumer science
	Sociology/social organization	Other vocational-technical education
	World civilization	Computer awareness/applications
	Other social science	Computer programming
		Other computer science
		Driver education
		Health education
		Home economics
		Physical education
		Other courses not elsewhere classified

2) Data merging & cleaning:

The 1999-2000 Schools and Staffing Survey contains data on the following school types:

Public	Private	Charter
District		
School	School	School
Teacher	Teacher	Teacher

The district survey contains 4,690 public school districts. This is merged with the public school data, which contains 8,432 public schools, using the Common Core District identification number common to both data sets. After dropping non-matching observations, this yields 7,467 public district-school unique observations. There were 541 schools that did not have an associated district, resulting in 6.4% of the sample of public schools being dropped. Examination of survey response rates of schools and districts suggests this is consistent with the expected number of non-matches. Further attempts to match the remaining public schools to districts using zip codes or district control numbers yielded no additional matches.

We then merge the school level data from public, private, and charters schools with the respective teacher level data sets. The 42,086 public school teachers yield 34,323 teacher-public school pairs (81.6% match rate). The 7,098 private school teachers and 2,611 private schools yield 6,372 private teacher-school pairs (89.8% match rate). The 2,847 charter school teachers and 870 charter schools yield 2,555 charter teacher-school pairs (89.7% match rate). Again, these match rates are consistent with expectations given the respective survey response rates. Teachers with no matching school information are dropped from the sample. There does not appear to be any systematic difference between matched and unmatched teachers.

Combining the three teacher-school pair data sets yields 43,250 unique teachers with corresponding schools characteristics. Of these, 41,161 remain after dropping observations with no college reported. We then merge this data with the NPSAS data to obtain measures of the quality of each teacher's degree. These degree quality averages are constructed using a pooled data set made up of 68,929 observations from 1989 NPSAS, 56,857 observations from 1995 NPSAS, and 63,172 observations from 1999 NPSAS. This pooled data yields a total of 42,187 undergraduate BA students with an observed ACT or SAT score. For students with an ACT score, we convert this to the SAT scale using the algorithm outlined by Schneider and Dorans (1999). If more than one score was reported, we averaged the available scores. These observations were then used to determine the average SAT score for each major group within each university. The average SAT scores are then merged with the SASS data to assign each teacher a degree quality measurement.

The data is further restricted to high school teachers only, reducing the sample to a total of 20,141 teacher-school pairs. If teachers were not classified as "regular" (full time, part-time, or long term substitutes), these were dropped from the sample, leaving a total of 19,271. Additional observations were dropped if no SAT score was available in the NPSAS data set. This further reduced the sample to 13,735. Of all the data cleaning conducted, this was the only step we found to be potentially biasing. Although the NPSAS was a representative sample of undergraduate institutions, an examination of the teachers who attended colleges not featured in the NPSAS suggests that these omitted teachers were more likely to be from small, less selective institutions and more likely to teach in schools with salary schedules. To the extent that this was true, our remaining sample of teachers will be made up of higher quality teachers than the true population. This should have the effect of potentially weakening our results somewhat. In order to address this issue, we hope to find a more comprehensive source of college data than the NPSAS and test the robustness of our results.

In our final data construction step, we merge the SASS-NPSAS teacher observations with community level data by matching district IDs and zip codes. In this step, 11,758 public teacher observations are matched at the district level, and 1,977 private and charter teachers are matched at the zip code level (of which 966 are able to be assigned the weighted average of the five-digit level community data and 1011 are assigned the weighted average at the three-digit level community data).

References:

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

2. Angrist, J.D. and Lavy, V., "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114 (1999), 533-575.

3. Baker, Bruce D., and Jill L. Dickerson, "Charter Schools, Teacher Labor Market Deregulation, and Teacher Quality: Evidence From the Schools and Staffing Survey," *Educational Policy*, 20 (2006), 752-778.

4. Ballou, Dale, "Do Public Schools Hire the Best Applicants?," *Quarterly Journal of Economics*, 111 (1996), 97-144.

5. Bonesronning, H., T. Falch, and J. Strom, "Teacher sorting, teacher quality, and student composition," *European Economic Review*, 49 (2005), 457-483.

6. Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor, "How and Why do Teacher Credentials Matter for Student Achievement?," *National Bureau of Economic Research Working Paper Series*, No. 12828 (2007).

7. Corcoran, S., Evans, W., and Schwab, R., "Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957-2000," *American Economic Review*, 94 (2007), 230-235.

8. D. Scheider, and N.J. Dorans, "Concordance between SAT and ACT scores for Individual Students," *College Entrance Examination Board* 1999).

9. Dolton, Peter, "The Economics of UK Teacher Supply: The Graduate's Decision," *The Economic Journal*, 100 (1990), 91-104.

10. Dorans, N.J., "Correspondences between ACT and SAT I Scores," *College Board Research Report* (New York: College Entrance Examination Board, 1999).

11. Duplantis, Richard, Chandler, Timothy and Terry Geske, "The growth and impact of teachers' unions in states without collective bargaining legislation," *Economics of Education Review*, 14 (1995), 167-178.

12. Eberts, R., K. Hollenbeck, and J. Stone, "Teacher performance incentives and student outcomes," *Journal of Human Resources*, 37 (2002), 913-927.

13. Ehrenberg, Ronald and Dominic Brewer, "Do School and Teacher Characteristics Matter? Evidence from Highs School and Beyond," *Economics of Education Review*, 13 (1994), 1-17.

14. Ferguson, R.F., "Paying for public education: New evidence on how and why money matters," *Harvard Journal on Legislation*, 28 (1991), 465-498.

15. Ferguson, R.F., and Ladd, H.F, ed., *How and why money matters: An analysis of Alabama schools* (Washington, D.C.: Brookings Institution, 1996).

16. Figlio, David N., "Can Public Schools Buy Better-Qualified Teachers?," *Industrial and Labor Relations Review*, 55 (2002), 686-699.

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

18. Freeman, Richard and Casey Ichniowski, "When Public Sector Workers Unionize," (Chicago and London: National Bureau of Economic Research Project Report, 1988).

19. Gilpin, Greg and Kaganovich, Michael, "The Quantity and Quality of Teachers: A Dynamic Trade-off," *Working Paper*2008).

20. Goldhaber, D., and E. Anthony, "Can teacher quality be effectively assessed? National board certification as a signal of effective teaching," *Review of Economics and Statistics*, 89 (2007), 134-150.

21. Greenwood, Rob, Larry Hedges, and Richard Laine, "The Effect of School Resources on Student Achievement," *Review of Educational Research*, 66 (1996), 361-396.

22. Guarino, C. M., L. Santibanez, and G. A. Daley, "Teacher recruitment and retention: a review of the recent empirical literature," *Review of Educational Research*, 76 (2006), 173-208.

23. Hanushek, Eric A., "Throwing Money at School," *Journal of Policy Analysis and Management*, 1 (1981), 351-388.

24. Hanushek, Eric A., "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, 24 (1986), 1141-1177.

25. Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steven G. Rivkin, "The Market for Teacher Quality," *National Bureau of Economic Research Working Paper Series*, No. 11154 (2005).

26. Hoxby, Caroline, "How Teachers' Unions Affect Education Production," *Quarterly Journal of Economics*, 111 (1996), 671-718.

27. Hoxby, C. M., and A. Leigh, "Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States," *American Economic Review*, 94 (2004), 236-240.

28. Krueger, Alan B., "Experimental Estimates of Education Production Functions," *The Quarterly Journal of Economics*, 114 (1999), 497-532.

29. Lakdawalla, D., "The economics of teacher quality," *Journal of Law & Economics*, 49 (2006), 285-329.

30. Ma, Lingjie and Roger Koenker, "Quantile regression methods for recursive structural equation models " *Journal of Econometrics*, 134 (2006), 471506.

31. Monk, David, "Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement," *Economics of Education Review*, 13 (1994), 125-145.

32. N.J. Dorans, C.F. Lyu, M. Pommerich, and W.M. Houston, "Concordance between ACT Assessment and Recentered SAT I Sum Scores," *College and University*, 73 (1997), 24-31.

33. Podgursky, M. J., and M. G. Springer, "Teacher performance pay: A review," *Journal of Policy Analysis and Management*, 26 (2007), 909-949.

34. Pudgursky, Michael, Ryan Monroe, and Donald Watson, "The academic quality of public school teachers: an analysis of entry and exit behavior," *Economics of Education Review*, 23 (2004), 507-518.

35. Rivkin, Steven G., Hanushek, Eric A., and Kain, John F., "Teachers, Schools, and Academic Achievement," *Econometrica*, 73 (2005), 417-458.

36. Strauss, Robert and Elizabeth Sawyer, "Some New Evidence on Teacher and Student Competencies," *Economics of Education Review*, 5 (1986), 41-48.

37. Strotz, R. and Wold, H., "Recursive vs. Non-recursive Systems: an Attempt at Synthesis," *Econometrica*, 28 (1960), 417-427.

38. Summers, A.A. and Wolfe, B.L., "Do Schools Make a Difference?," *The American Economic Review*, 67 (1977), 639-652.

39. Winkler, D.R., "Educational Achievement and School Peer Group Composition," *Journal of Human Resources*, 10 (1975), 189-204.