Analyzing the Determinants of the Matching of Public School Teachers to Jobs

Don Boyd
Rockefeller Institute of Government

Hamp Lankford
University at Albany

Susanna Loeb
Stanford University

Jim Wyckoff
University at Albany

July, 2002

We are grateful to the Smith Richardson Foundation and the Office of Educational Research and Improvement, U.S. Department of Education for financial support. They do not necessarily support the views expressed in this paper. All errors are attributable to the authors.

I. Introduction

Differences across schools in the qualifications of teachers are likely contributors to the substantial gaps in academic achievement among income and racial/ethnic groups of students. Such sorting of teachers across schools and districts is the result of a range of decisions made by individual teachers and school officials. These include the decisions of individuals and school officials that determine initial job matches and subsequent decisions that affect job quits, transfers and terminations. This paper focuses on the initial matching of teachers to jobs and introduces an empirical model that employs the method of simulated moments to estimate the effect of various characteristics of teachers and schools on this matching.

Low-income, low-achieving and non-white students, particularly those in urban areas, often are taught by the least skilled teachers. For example, in schools in the highest quartile of student performance on the New York State 4th Grade English Language Arts Exam only three percent of teachers are uncertified, only ten percent earned their undergraduate degree from least competitive colleges, and only nine percent of those who have taken a general knowledge teacher certification exam have failed.[1] In contrast, in schools in the lowest quartile of student performance, 22 percent are uncertified, 26 percent come from least competitive colleges, and 35 percent have failed a certification exam (Lankford, Loeb and Wyckoff, 2002).

Inefficient hiring and district assignment may contribute to the disparities observed in the data, however teacher preferences are likely to be particularly influential.[2] Teachers differ fundamentally from other school resources. Unlike textbooks, computers, and facilities, teachers have preferences about whether to teach, what to teach, and where to teach. Potential teachers

[1] Teachers in New York have had the option of taking the NTE General Knowledge Exam or the NYSTCE Liberal Arts and Science Exam. Throughout the paper “failure” refers to failing one of these exams on the first attempt.

[2] Few studies have explored district-hiring practices, though Pflaum & Abramson (1990), Ballou (1996) and Ballou and Podgursky (1997) do provide evidence that many districts are not hiring the most qualified candidates. Schools also vary in the political power they exert, which may lead to differences in teacher qualifications among schools within the same district. Bridges (1996) found that when parents and students complained about poor teachers, the teachers were likely to be transferred to schools with high student transfer rates, large numbers of students receiving free or reduced-price lunches, and large numbers of minority students.

prefer one type of district to another; and within districts, they prefer one school to another. There has been much discussion about the role that compensation plays in the ability of schools to attract and retain high-quality teachers. A large literature suggests that teachers respond to wages though research on the compensating wage differentials needed to attract teachers with particular characteristics to schools with particular characteristics has not produced consistent results.[3] Salaries are one job attribute that likely affects sorting, but non-pecuniary job characteristics, such as class size, preparation time, facilities, or characteristics of the student body, are important as well.[4]

This paper models the allocation of teachers to schools based on the preferences of both employers and employees. Our long-term goal is to identify policies that are effective for attracting and retaining teachers in low-performing schools. As we discuss further below, such identification has many difficulties, not the least of which is the endogeneity of any policy we observe. The goal of this paper is more limited. We introduce our model for the matching of teachers to schools and estimate this model with a limited set of school and teacher measures.

Our data comes from administrative records in New York State that allow us to follow *all* teachers in the state over the past 30 years. The breadth of the data (all teachers in all public schools) allows analytical flexibility not possible with smaller datasets. For example, we can look at the sorting of teachers separately for each region of the state. In addition, we can see how the preferences and decisions of one teacher in a labor market affect the matching of other teachers to jobs. The length of the data will allow us to assess how sorting and preferences have changed over time and also, eventually, will allow us to use fixed-effect approaches to identify policy impact.

[3] As a group, these studies show that individuals are more likely to choose to teach when starting teacher wages are high relative to wages in other occupations (Baugh and Stone, 1982; Brewer, 1996; Dolton, 1990; Dolton and van der Klaaw, 1999; Dolton and Makepeace, 1993; Hanushek and Pace, 1995; Manski, 1987; Mont and Reece, 1996; Murnane, Singer & Willett, 1989; Rickman and Parker, 1990; Stinebrickner, 1998, 1999, 2000; Theobald, 1990; Theobald and Gritz, 1996). Baugh and Stone (1982), for example, find that teachers are at least as responsive to wages in their decision to quit teaching, as are workers in other occupations.

The data is richer in its descriptions of teachers than other administrative datasets used to date, including teachers' test scores and undergraduate institutions. It also allows us to match teachers to characteristics of the schools in which they teach in a way that most national longitudinal surveys, such as High School and Beyond or the National Longitudinal Survey of Youth, do not.

In what follows, we start with a description of the data and then provide background on the teacher labor market in New York State -- including the extent of systematic sorting of teachers across schools; the importance of initial job matches, compared to transfers and quits, in determining this sorting; and evidence concerning the geographical extent of teacher labor markets. Section IV discuss traditional models for assessing the sorting of teachers across schools within regions, including the difficulties that arise from the two-sided matching of teachers to employers. Section V presents our model of the initial job match and estimation strategy. The final sections give empirical results and conclude.

II. The Data

Our database links seven administrative datasets and various other information characterizing districts, communities, and local labor markets in New York State. It includes information for every teacher and administrator employed in a New York public school at any time from 1969-70 through 1999-2000. (See the table in Appendix A.) The core data comes from the Personnel Master File (PMF), part of the Basic Education Data System of the New York State Education Department. In a typical year there are 180,000+ teachers identified in the PMF. We have linked these annual records through time, yielding detailed data characterizing the career history of each individual.

---

[4] In Texas, Hanushek, Kain and Rivkin (1999) found teachers moving to schools with high-achieving students and, in New York City, Lankford (1999) found experienced teachers moving to high-socioeconomic status schools when positions became available.

Several other databases that contain a range of information about the qualifications of prospective and actual teachers, as well as the environments in which these individuals make career decisions, substantially enrich this core data. For teachers this information includes age, gender, race/ethnicity, salary, course subject and grade taught, experience (in the district, in NYS public schools, and total), years of education and degree attainment, and teacher certification exam scores of individual teachers and whether they passed on their first attempts. In addition, we identify the institutions from which individual teachers earned their undergraduate degrees and combine it with the Barron's ranking of college selectivity to construct variables measuring the selectivity of the college from which each teacher graduated and the location of the institution. Measures of schools and districts include enrollment, student poverty, racial composition, limited English proficiency composition, student test results for recent years, dropout rates, district wealth, district salary schedules, crime, spending in numerous categories, number of employees in numerous categories, as well as many other measures. Using information on the zip code of residence when the teacher applied for certification and the zip code of each school, we create a "distance from home" measure for each school-teacher combination in our sample. For a sub-sample of teachers we know where they lived while in high school.

While we utilize much of this data for the background discussion, we use a small subset for the estimation of the matching model -- first year elementary school teachers in the Albany-Schenectady-Troy, Buffalo, Syracuse, Rochester, and Utica-Rome metropolitan areas for the years 1994-95 through 1999-2000.

III. Background

This section describes the teacher labor market in New York State, providing evidence of substantial sorting of teachers across schools and of the importance of the initial (first-job) match of teachers to schools in determining this sorting. The main analysis in this paper will model the initial sorting, and thus it is worth establishing that this sorting exists. In addition, this section describes how far teachers travel from the high schools they attended to their first teaching jobs. The proximity of these locations suggests that distance is important in teachers' assessment of job benefits. Different teachers, having grown up in different areas, then may have very different assessments of the relative benefits of teaching in different schools. This preference heterogeneity is difficult to incorporate in traditional models of sorting but fits easily into the model that we use. In addition, the close proximity of teachers' home towns to their jobs suggests that it is worth looking at a number of smaller labor markets within New York State, instead of assuming that the state is a single market.

Characterizing the Sorting of Teachers Across Schools:[5] The characteristics of teachers differ substantially and in systematic ways between schools. This is true across a wide range of teacher attributes. For example, at least ten percent of schools in New York State have no teacher who: is new to teaching, is teaching out of their certification area, failed a certification exam on their first attempt, or graduated from a "least competitive" undergraduate colleges. Yet, at least another ten percent of schools have a substantial portion of teachers who have no prior teaching experience (18 percent), only are teaching courses for which they are not certified (24 percent), or failed a certification exam (about one third). In ten percent of schools less than half of the teachers are permanently certified in all of the courses they teach while in another ten percent of schools this figure is nearly 90 percent.

School-level teacher attribute measures are correlated; schools that have poorly qualified teachers as measured by one attribute are more likely to have poorly qualified teachers based on all other measures.[6] Because of this correlation across measures, we use principal components

[5] This section is based on work reported in Lankford, Loeb, and Wyckoff (2002).
[6] For example, schools with high proportions of teachers who failed exams are more likely to have teachers from less competitive colleges (correlations of approximately 0.45); schools with a high proportion of teachers who are not certified to teach any of the courses that they currently teach are much more likely to have graduated from the less competitive colleges (correlation of .40).

analysis at the school level to combine a number of characteristics and create a composite of average teacher qualifications. Appendix B describes the components of this measure. It has a reliability of 0.86 and explains 52 percent of the variation in its component measures.

To investigate the geographical nature of this sorting, we characterize nine labor markets in the state, consisting of six individual metropolitan statistical areas (MSAs) and three remaining rural areas.[7] Figure 1 shows the distribution of the composite index across the State's labor markets in 2000. The distribution of teachers looks similar across regions. The notable exception is that the New York City Region has fewer skilled teachers than the other regions. This is true at the 10th, 50th, and 90th percentiles, with the differences at the 10th percentile being by far the greatest. The figure also shows that the variation in teacher qualifications within regions is greater than the variance between regions.

Within regions, urban schools systematically employ less-qualified teachers than suburban schools. For each of the metropolitan labor markets in New York State, Table 1 separately shows the distribution of school-level teacher attributes for urban and suburban schools. In ten percent of urban schools in the Buffalo Region one third of the teachers had failed the exam, whereas in suburban schools, only one fifth had. Similar trends are evident across multiple measures and across the multiple metropolitan areas. The results for the New York City region are again most striking. Ten percent of New York City urban schools have an average teacher qualification measure that is five standard deviations lower than the state average.

Within urban areas, the qualifications of teachers are sorted with respect to the racial and economic attributes of students.[8] Table 2 shows that in urban Rochester and Syracuse white students attend schools with teachers with .20 to .35 standard deviations higher skills than non-

---

[7] The urban regions are Albany-Schenectady-Troy, Buffalo-Niagara Falls, New York City (including Putnam, Rockland, Westchester Nassau, and Suffolk counties), Rochester, Syracuse, and Utica-Rome. The rural regions are Mid-Hudson, North Country, and the Southern Tier.

[8] LEP students also receive less qualified teachers when compared to non-LEP students, although the differences are not as great as those comparing non-whites to whites and poor to non-poor.

white students as measured by the teacher qualifications factor; and non-poor students attend schools with teachers with .20 to .27 standard deviations higher skills than poor students.[9] Yonkers, though not shown, fits this same distribution. The disparities within the New York City School District are substantially greater. In New York City, 21 percent of those teaching non-white students are not certified in any subject taught, compared to 15 percent of those teaching white students. Twenty six percent of those teaching non-white students failed either the General Knowledge or Liberal Arts and Science certification exam, compared to 16 percent of those teaching white students.[10]

In summary lesser-qualified teachers teach low-income, low-achieving and non-white students, particularly those in urban areas. While some of these differences are due to differences in average characteristics of teachers across districts, not within urban districts; differences among schools within urban districts are important as well. The New York City school district, in particular, exhibits large differences among student groups in the qualifications of their teachers.[11]

The Relative Importance of Initial Matches, Transfers and Quits: [12] What are the dynamics that lead to the systematic sorting of teachers across schools? At any point in time, the distribution of teachers across schools will depend upon how starting teachers are matched to schools and on the patterns of subsequent transfers, quits and terminations. To assess the importance of the initial match, we follow a cohort of new teachers, looking at the distribution of those who remain from this cohort in each year after entry into the New York State system. Figure

[9] Poverty status is more accurately reported for students in kindergarten through sixth grade. Because of this, we only include schools that have some of these grades in the poor / non-poor comparison. The race comparisons are estimated over the full set of schools. These measures are based on school averages weighted by the student composition of schools.

[10] Some of these differences may be driven by differences in the preferences of residents over unobserved attributes. As an example, schools with a high percent of minority students may benefit from having teachers with similar racial and ethnic backgrounds. These teachers may have attended lower ranked undergraduate institutions and may score lower on certification exams than other teachers of similar quality. If this were the case, we may see teachers with poorer qualifications as measured by test score and school ranking in high percent-minority schools yet the teachers in these schools would have important unmeasured skills that makes their overall quality higher. We know of no studies that systematically examine this issue.

[11] Note, this analysis only assesses differences in the average characteristics of schools. Additional systematic sorting of teachers to students may occur within schools.

2 plots the percent of teachers from the 1995 entering cohort that failed a teacher certification exam by whether the teachers taught in New York City urban schools, schools in New York City suburbs, other urban schools, or other suburban schools. The conclusion is evident. Most of the sorting on this measure occurs at the initial match. The results for other cohorts and other teacher measures are similar (approximately 83 percent of differences across these four groups in the percent of teachers from highly competitive colleges and 91 percent of differences in the percent of teachers from least competitive colleges are due to the initial match rather than to quits and transfers, based on the six entering cohorts from 1990-1995).

Within urban areas, sorting at the initial match also is evident, though additional sorting occurs through transfers and quits. Figure 3 shows evidence of this for New York City, plotting the percent of teachers who failed a general knowledge certification exam on their first attempt by school quartiles for the percent of black or Hispanic students in the school. Across the 1990 through 1995 cohorts 53 percent of differences in the percent of teachers who failed between the lowest and highest quartile schools is due to the initial match rather than quits or transfers. Thus, while differences in exit and transfer behavior do affect the disparities of teachers across urban schools, initial matches are important within urban school systems, between urban and suburban areas, and among regions. The remainder of the paper explores this initial match in more detail.

The geographic extent of teacher labor markets: [13] The data include the location of New York teachers at several points in their careers. For those who were awarded certification to teach in New York State since 1985, we know where they attended college and where they took their first teaching job. For many of these teachers we also know their hometowns during high school. Based on this information, we examine the relationship between the place of a teacher's first

---

[12] This section is based on work reported in Boyd, Lankford, Loeb and Wyckoff (2002a).
[13] This section is based on work reported in Boyd, Lankford, Loeb and Wyckoff (2002b).

teaching job, and the locations of his/her college and 'hometown.'[14] Teachers take their first job strikingly close to where they grew up. Table 3 shows that 59 percent of teachers teach within 15 miles of the place where they grew up and more than 82 percent teach within 40 miles of their hometown. This proximity to home is similar across urban and suburban districts (see Table C1 in Appendix C). Even of those that travel over 100 miles to college, most return home to teach.

Much of the difficulty in modeling the initial match of a teacher to a school results from the two-sided choice characteristic of this match; that is, both teachers and schools have to agree upon the match. However the choice of the region in which to teach and, perhaps, whether to teach in the urban or suburban part of a region, is more likely to be a one-sided choice on the part of the teacher, making a simple choice framework more appropriate. While much of the sorting of teachers across schools occurs within regions, an analysis of sorting between regions can identify factors that may be important for within region sorting and can help define the geographic extent of local labor markets for the two-sided analysis.

A conditional logit model estimating the effect of distance on teachers' choice of region indicates that distance from hometown is important to teachers' regional location decisions (Table C2 in Appendix C).[15] As shown in Figure 4, an individual is more than twice as likely to locate in a region that is within five miles of her hometown as one 20 miles away and more than four times as likely to locate in a region within five miles of her hometown as one 40 miles away. She is more

[14] Unfortunately we know the location of an individual's hometown only if they applied to a SUNY campus since 1990. Individuals are more likely to be SUNY applicants north of New York City. For example, among first time teachers in 2000, 32 percent applied to a SUNY college. In New York City this figure is closer to 20 percent but often approaches 50 percent in many upstate regions. The key analysis in this paper is based on an alternative distance measure available for all teachers.

[15] We model choice as a function of: (1) the distance from hometown to each region, including squared and cubed terms; (2) distance interacted with gender, SAT, and grew up in urban, suburban or rural area; (3) dummy variables indicating whether a region includes the hometown and this variable interacted with urbanicity; (4) same urbanicity (e.g., urban) as the region of hometown, and this variable interacted with urbanicity; (5) other portion of hometown metropolitan area; (6) the distance from the college, including squared and cubed terms and interactions with gender and urbanicity; and (7) college in the same region. The estimation employs all teachers whose first teaching job occurred between 1998 and 2000. We define regions as in Figure 1 by seven metropolitan areas and three rural areas. The metropolitan areas are then split between urban and suburban schools. A number of specification checks produce similar results including: combining urban and suburban areas, looking at the choice of metropolitan region of teaching; excluding New York City; and using all teachers with distance from college instead of distance from home.

than three times as likely to locate in a region 25 miles from her hometown as one 80 miles away. Teachers appear to place a premium on jobs close to their hometowns.[16,17]

The simple conclusion of this analysis is that distance is an important factor in teachers' decision making and, because of this, labor markets are geographically quite limited. Thus, similar schools located in areas that produce different amounts of high school graduates interested in teaching may face substantially different supplies of potential teachers.

IV. Common Approaches for Modeling Sorting

Before describing our sorting model in detail it is worth reviewing several literatures pertinent to the study of the sorting of teachers across jobs. These include the hedonic wage literature and at least two literatures concerned with two-sided matching.

Hedonic Wage Equations: Most previous studies of teacher labor markets (such as Antos and Rosen, 1975) use a hedonic wage model to estimate teacher preferences for various school or district characteristics. These models maintain a competitive framework in which there are large numbers of employers and potential employees having vectors of attributes. There are sufficiently large numbers of employers and employees and the distributions of worker and job attributes are such that employers' choices of various combinations of employee attributes and employees' choices of various combinations of employer attributes are continuous. A market clearing price function, determined by the joint distribution of the attributes and preferences of employees and the joint distribution of the job attributes and the preferences of employers, characterizes the competitive equilibrium.

These price functions are generally estimated based on the following equation:

---

[16] New teachers have strong preferences to locate in the region of their hometowns, but also prefer to locate in regions similar to those of their hometown, other things (including distance) equal. For example, a new teacher whose hometown is in a suburban area is 4.5 times as likely to locate in that suburban area, as she is to locate in the urban portion of the same metropolitan area (Appendix C, Table C3).

[17] The importance of distance to an individual varies only slightly by the individual's attributes (Figure C1 in Appendix C).

$$(1) \quad W_{kj} = X_k\theta + Y_j\beta + \eta_{kj}$$

where W_{kj} is the log wage of worker j in job k, X_k is a vector of job characteristics, Y_j is a vector of worker characteristics, and η_{kj} is a random disturbance. β is interpreted as the percent increase in wages needed to attract a worker of a given quality to job k when the value of Y_j increases by one unit. While the simplicity of the approach is appealing, in practice the estimates have proven inconsistent. A number of reasons have been posited for the inconsistencies including omitted variables (Brown, 1980; Lucas, 1977), simultaneity (McLean et al. 1978), measurement error, and labor market frictions (Hwang, Mortensen and Reed, 1998; Lang and Majumdar, 2001). In the case of teacher labor markets, omitted variables characterizing schools and students and the endogenous determination of policy have been the focus of much of the discussion surrounding hedonic wage results. However, there are other problems with these models as well.

First, a problem arises in estimating a hedonic wage equation for public school teachers as a result of set wage schedules which imply that all teachers in a district with the same education and experience earn the same salary, regardless of their other attributes or the characteristics of the schools in which they teach. Because salaries are set at the district level and generally do not vary with teacher and school attributes, the error term in Equation 1 is correlated with the explanatory variables by construction. Thus, parameter estimates will be biased.

Second, in a market categorized by contracts that set wages for three or more years and by social decision-making practices that are likely to limit both the variation and the flexibility of the wage, teacher quality and not wage is likely to clear the market. In this context a "quality hedonic" model, such as Equation 2, in which the wage is just considered another characteristic of the work environment may make more sense.

$$(2) \quad Y_{jk} = X_k\alpha + W_k\delta + \upsilon_{kj}$$

Yet, this model is not satisfying either. It is likely to require multiple equations because there are many characteristics of teachers and not just one overall quality measure that districts care about.

Such a multiple equation framework will not capture employers' tradeoffs in choosing among those attributes. More importantly, this model is subject to many of the same problems as the traditional hedonic models.

Third, because hedonic equations measure price for the marginal worker and workers are likely to be heterogeneous in their preferences, hedonic equations do not provide good estimates of the impact of large changes, arising, for example, from policy or demographic shifts. If there is heterogeneity in teachers' subjective evaluations of various school attributes, an estimated hedonic equation will only provides implicit prices for those individuals currently "at the margin". With low-income and minority students in urban setting being taught by teachers who, on average, are less qualified, a hedonic wage equation by itself would not provide estimates of the salary premiums that would have to be paid to attract more highly qualified teachers to those same environments. Similarly, an estimated reduced-form hedonic equation loses much of its relevance when either supply or demand side factors change significantly. For example, ongoing shifts in student demographics and significant changes in the relative scarcity and composition of teachers due to large numbers of retirements can significantly change the sorting of teachers across schools and the implicit prices on various school and student-body attributes.

The more general point is that an estimated hedonic wage equation by itself does not identify demand-and supply-side preferences of agents. As an illustration, changes in the relative number of buyers and sellers could change the hedonic estimates even if the underlying preferences did not change. Under some circumstances, a second stage estimation can yield estimates of such preference parameters. Distance measures have been introduced into some structural hedonic models, such as rent-gradient models in which all individuals are assumed to prefer living as close as possible to the central city. However, more general specifications of the impact of distance are problematic for this framework. For example, the estimation of a hedonic equation and the second-stage estimation of preference parameters is difficult when there is extreme heterogeneity arising from individuals having differing assessments of the value of particular job locations because these assessments depend upon the candidates own location.

Finally, hedonic models assume a large enough market that the distributions of employer and employee attributes are continuous. This is unlikely to be the case in teacher labor markets, especially because of the apparent small size of these markets. The "effective size" of the market is limited by the importance of distance in teachers' evaluation of jobs. In such contexts, discrete choice models such as random utility models are likely to be more accurate (Freeman 1979; Palmquist 1991).

To deal with these and other consideration, we estimate a structural model, based on the two-sided matching literature that accounts for pertinent features of teacher labor markets as well as the factors affecting the separate, but interdependent, choices made by job candidates, teachers and school officials. The model can easily allow for variation in preferences by including measurable characteristics of teachers or schools as preference shifters. Preferences with respect to distance also enter in a straightforward manner.

Two-sided matching: The two-sided matching literature is applicable to a broad range of settings having the common feature that individuals in one group are matched with individuals, agents or firms in a separate, second group. Examples include models of marriage, employment and college attendance.[18] In all of these cases, the matching is two-sided in that whether a particular match occurs depends upon separate choices made by the two parties. Furthermore, these choices are not made in isolation. "A worker's willingness to accept employment at a firm depends not only on the characteristics of the firm but also the other possible options open to the worker. The better an individual's opportunities elsewhere, the more selective he or she will be in evaluating a potential partner" (Burdett and Coles, 1999).

[18] These cases differ from the roommate problem where those being matched come from the same group. In two-sided match models all agents fall into one of two distinct groups and seek a match with one or more agents in the other group.

Within the two-sided matching literature, there are now a large number of papers that build upon the work of Gale and Shapley (1962) and are concerned with the allocation (matching) of fixed numbers of agents from two disjoint sets. This game-theoretic research has considered both one-to-one matching such as marriage and many-to-one matching such as employment and college-admission, the former being a special case of the latter.[19] While a growing number of papers allow utility to be transferable so that the division of match surplus is determined endogenously at the time partners match, most game-theoretic models have assumed that utility is *nontransferable*; that is, how the surplus from any given match is split between the matching pair is predetermined. This more traditional assumption is applicable to teacher labor markets since salaries (set through collective bargaining for three to five year periods), other conditions of work, and the attributes of teacher candidates are fixed in the short-run.

In addition to the game-theoretic studies, there is a large literature in labor economics employing two-sided matching models with search. This research distinguishes itself in a number of respects. First, whereas almost all the game-theoretic models assume full information and no market frictions, such frictions are central to the labor-search models of marriage and job match. A second difference is that the demand side of the labor-search models often is characterized by free entry of profit maximizing firms so that the number of jobs to be filled is not fixed as in the game-theoretic match literature. A third difference that is especially pertinent for our empirical analysis concerns the extent and nature of agent heterogeneity allowed in the models. Game-theoretic two-sided match models typically only require that each agent's ranking of match partners is complete and transitive, with no restrictions regarding the extent of preference heterogeneity. In contrast, the search models either maintain homogeneity of preferences or allow for only limited heterogeneity.

[19] In addition to the papers focusing on decentralized allocation mechanisms, extensive research has addressed centralized mechanisms such as those used to assign medical interns to hospitals. Roth and Sotomayer (1990) provide a clear synthesis of both the theoretical literature to date and how the theoretical findings provide important insights regarding implications of the institutional features characterizing the centralized matching algorithms used, as well as factors that have contributed to the evolution of those features.

Some models maintain *match heterogeneity*, where agents in each group are ex-ante identical but some matches are relatively more productive, with the productivity of each possible match determined by a random draw from some known distribution. Other models maintain *ex-ante heterogeneity* where there are systematic differences across agents independent of the partners to whom they are matched, with all agents in one group having the same ranking of the potential partners in the other. For example, some workers may be more productive than others and some jobs may be more or less attractive. Limitations on the degree of heterogeneity are needed in order to solve for the search equilibriums (Burdett and Coles, 1999). Such limited heterogeneity would be quite restrictive if maintained in our analysis. For this reason, our model builds on the game-theoretic approach.

V. The Model

Consider an environment in which $C=\{c_1,\cdots,c_J\}$ represents the set of J individuals seeking teaching jobs and $S=\{s_1,\cdots,s_K\}$ represents the set of K schools having jobs to be filled, $J\geq K$. (For now assume that each school has one job opening though this is relaxed in the empirical analysis.) We assume that each agent has a complete and transitive preference ordering over the agents on the other side of the market and that these orderings arise from job candidates' preferences over job attributes and hiring authorities' preferences over the attributes of candidates.

Let u_{jk} represent the utility of working in the k^th school as viewed from the perspective of the j^th candidate where $u_{jk}=u(x_k^1,d_{jk}\,|\,y_j^2,\beta)+\delta_{jk}$. x_k^1 is a vector of observed attributes of the k^th school pertinent to the j^th individual and d_{jk} is the distance to the k^th job for the candidate. Vector y_j^2 represents observed attributes of the j^th candidate that affect the individual's assessment of the k^th alternative and β is a vector of parameters. δ_{jk} is a random variable reflecting unobserved heterogeneity in the attractiveness of a particular school for different individuals. If no job match is entered, the individual's utility is u_{j0} which depends upon observed and unobserved attributes of the individual. Thus, the individual will always turn down a job offer if $u_{jk} < u_{j0}$. Here we assume that $u_{jk} > u_{j0}$, for all k and j but plan to allow for the more general case when we extend the model to consider all candidates, not just those actually obtaining jobs.

The hiring authority for the k[th] school is assumed to have preferences over the attributes of job candidates. Let $v_{jk} = v(y_j^1 \mid x_k^2, \alpha) + \omega_{jk}$ represent the attractiveness of the j[th] candidate from the perspective of the hiring authority for school k. The vector y_j^1 represents pertinent observed attributes of the j[th] candidate. The vector x_k^2 represents the observed attributes of the k[th] school that might affect the authority's assessment of the j[th] candidate. α is a vector of parameters. The random error ω_{kj} reflects unobserved factors. To simplify the analysis, we assume hiring authorities prefer all of the candidates to the alternative of leaving job vacancies unfilled. This assumption, combined with the assumption that there are sufficient numbers of willing candidates, implies that all job openings will be filled.

Consider a case where the sets C and S are known, as are the values of $y_j = (y_j^1, y_j^2)$ for each candidate and $x_k = (x_k^1, x_k^2)$ for each job. Given the vector of parameters β and a particular set of random variable draws for the δ_{jk}, the formula $u_{jk} = u(x_k^1, d_{jk} \mid y_j^2, \beta) + \delta_{jk}$ implies the matrix of candidates' benefits represented in panel (A) of Figure 5. Each row shows the benefits that a particular candidate attributes to being employed in each of the K school alternatives. These rows of benefit values, in turn, imply candidates' complete rankings of school alternatives shown in panel (C). r_{jk}^c is the jth candidate's ranking of the kth school alternative. In a similar way, the vector of parameters α and a particular set of random variable draws for the ω_{kj}, together with the formula

$v_{jk} = v(y_j^1 | x_k^2, \alpha) + \omega_{jk}$, imply the matrix of school benefits represented in panel (B) of Figure 5 and the complete rankings of candidates by hiring authorities shown in panel (D). Each column of panel B shows the benefits to a particular school of having an opening filled by each of the alternative candidates. r_{jk}^s is the ranking of the jth candidate from the perspective of the kth employer.

Figure 5

(A) Candidates' benefits from alternative employment

$$\begin{array}{cccccc} & s_1 & s_2 & \cdots & s_K \\ c_1 & u_{11} & u_{12} & \cdots & u_{1K} \\ c_2 & u_{21} & u_{22} & \cdots & u_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_J & u_{J1} & u_{J2} & \cdots & u_{JK} \end{array}$$

(B) Schools' benefits from alternative candidates

$$\begin{array}{cccccc} & s_1 & s_2 & \cdots & s_K \\ c_1 & v_{11} & v_{12} & \cdots & v_{1K} \\ c_2 & v_{21} & v_{22} & \cdots & v_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_J & v_{J1} & v_{J2} & \cdots & v_{JK} \end{array}$$

(C) Candidates' rankings of employers

$$\begin{array}{cccccc} & s_1 & s_2 & \cdots & s_K \\ c_1 & r_{11}^c & r_{12}^c & \cdots & r_{1K}^c \\ c_2 & r_{21}^c & r_{22}^c & \cdots & r_{2K}^c \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_J & r_{J1}^c & r_{J2}^c & \cdots & r_{JK}^c \end{array}$$

(D) Employers' rankings of candidates

$$\begin{array}{cccccc} & s_1 & s_2 & \cdots & s_K \\ c_1 & r_{11}^s & r_{12}^s & \cdots & r_{1K}^s \\ c_2 & r_{21}^s & r_{22}^s & \cdots & r_{2K}^s \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_J & r_{J1}^s & r_{J2}^s & \cdots & r_{JK}^s \end{array}$$

If each of the candidates unilaterally were able to choose the school in which to teach, the framework summarized in panel (A) would imply that β in $u_{jk} = u(x_k^1, d_{jk} | y_j^2, \beta) + \delta_{jk}$ could be estimated using data characterizing those choices and a standard multinomial probit or logit random utility model. Similarly, α could be estimated easily using the same type model if each hiring authority unilaterally chose among candidates. However, the empirical model we employ is more complex for two reasons. First, it is the interaction of decisions made by a candidate and a hiring authority for a school that determines whether the two are matched. Second, even though any such interaction would complicate the model, the decisions made by the two parties considering whether to match crucially depend upon the choices made by all other candidates and employers. In particular, a candidate's willingness to accept a particular match depends upon her own preferences as well as her "effective" choice set, i.e., the set of schools willing to hire her given their own "effective" alternatives. In turn, whether employers make the candidate an offer will depend upon whether they prefer to employ alternative candidates who are willing to fill their positions, and so on.[20]

Because our framework is an example of the standard two-sided matching model extensively studied by game theorists, many of the theoretical findings in that literature directly apply to our analysis (Roth and Sotomayer, 1990). As is common in the literature, we assume that there is a decentralized job-match mechanism having the following characteristics. Each employer makes an offer to its highest ranked prospect. Job candidates receiving offers reject those that are dominated either by remaining unemployed or by better job offers, and "hold" their best offers if they dominate being unemployed. Employers whose offers are rejected make second round offers to their second highest ranked choices. Employers whose offers remain open stay in communication with these candidates but otherwise take no action. Job candidates receiving better offers inform employers that they are rejecting the less attractive positions previously held. In subsequent steps each employer having an opening with no outstanding offer makes an offer to its top candidate among the set of job seekers who have not already rejected an offer from the employer. Employees in turn respond. This *deferred acceptance procedure* continues until firms have filled all their positions with their top choices among those not having a better offer or have made unsuccessful offers to all their acceptable candidates. As shown by Gale and Shapley (1962),

[20] To see how one can have a model with joint decisions that avoid this complexity, one need only consider a two-sided search model in which candidates and employers randomly meet and individually decide whether they are willing to match based upon reservation-wage decision rules, with a match occurring only if both agree. The relative simplicity of this model comes from the underlying assumptions of the model that imply the reservation-threshold for any agent is not affected by the choices made by any other agent. In contrast, the model we use explicitly allows for complex

such an allocation mechanism always will yield a stable matching, in the sense that there will be no candidate and employer currently not matched who both would prefer to be matched to each other rather than to the agents to whom they are matched. Furthermore, if the rankings are strict (i.e., no agent is indifferent between any two alternatives), the resulting stable matching will be both unique and employer-optimal (i.e., all employers weakly prefer this match to all other stable matches). Alternatively, a deferred acceptance procedure in which candidates made offers to hiring authorities would result in an employee-optimal match.

The equilibrium employer-optimal stable matching corresponding to the alternatives and rankings characterized in Figure 5 is represented in the left side of Figure 6. The right side of Figure 6 characterizes this matching in terms of the resulting relationship between the attributes of candidates and the schools where they are employed.

Figure 6: Resulting Matching of teachers and Jobs

School-teacher matched pairs	Joint distribution of school and teacher attributes
$\begin{Bmatrix} (s_1, c_j) \\ (s_2, c_{j'}) \\ \cdots \\ (s_2, c_{j'''}) \end{Bmatrix}$	$\begin{bmatrix} x_1 & y_j \\ x_2 & y_{j'} \\ \vdots & \vdots \\ x_K & y_{j'''} \end{bmatrix} = [x \ y]$

The matching of candidates to schools represented in Figure 6 corresponds to particular values of the model's random variables (δ_{jk} and ω_{kj}; j = 1,2,…,J and k = 1,2,…,K), the explanatory variables (e.g., y_j and x_k) and the parameters $(\theta = (\alpha, \beta))$ of the model. Given the implied rankings for candidates and jobs, deriving such a stable matching is relatively easy using the Gale-Shapley matching algorithm. However, deriving closed-form expressions for the likelihood of

---

interactions. Furthermore, our model allows for unobserved heterogeneity in agents' rankings of alternatives whereas search models typically assume there are common rankings or only very limited heterogeneity.

observing any particular candidate-job matching or the probability distribution of any particular distribution of worker and job attributes is impossible.[21] To compute the likelihood of a particular stable matching one would need to identify the set of all possible combinations of the random errors that would lead to that same stable matching. This would entail determining all possible combinations of the rankings of candidates and employers that would yield a particular matching and, in turn, all the combinations of random variable values that would lead to each of those sets of rankings. This is an impossible task, especially since it would have to be done repeatedly for various parameter values. Even if the ranges of the various random errors could be identified, computation of the corresponding likelihood would be impossible given that the implied integrals would have high dimensions and very complex regions of integration.[22] These complexities motivate our use of a method of simulated moments (MSM) estimation strategy.

Before discussing the MSM approach, it is first necessary to generalize the notation and framework. Whereas the above discussion was for a single market at one point in time, our empirical analysis considers M local labor markets, m = 1,2,…,M, and T years, t = 1,2,…,T. To account for this generalization, we need only add the subscripts "m" and "t" to the explanatory and random variables defined above. For example, y_{jmt} represents the attributes of the jth candidate first employed in the mth market during time period t. An assumption is needed to allow for multiple job openings in a single school in any given year. With our empirical analysis focusing on elementary schools where there is a large degree of homogeneity across teaching jobs, we assume that all job openings within a school are identical. As shown in the two-sided match literature, the pertinent theoretical underpinning for a many-to-one match parallel the one-to-one matches discussed above.

Let y_{mtki} represent the attributes of the teacher newly employed during period t to fill the ith vacancy of school k in labor market m where i = 1,2,…, n_{mtk} and n_{mtk} is the total number of job

[21] Note that multiple worker-job matchings will yield the same distribution of matched attributes if either multiple candidates or multiple jobs have the same observed attributes.

openings in the kth school for that year. (Reflecting the two-sided match, y_{mtki} from the perspective of this employer is the same as y_{mtj} defined above from the perspective of the employee where j is the individual employed to fill the kth firm's ith position.) The structure of the two-sided matching model, given values of parameters α and β and the distributions of sets of random variables δ_{jk} and ω_{kj}, imply the joint distribution of x_{mtk} and y_{mtki}, j = 1,2,…,J and k = 1,2,…,K. This in turn implies the expected value of y_{mtki} for the kth school, $E\left(y_{mtki} \middle| x_{mtk};\theta\right)$. Subscript i in this expression can be dropped as a result of the assumption that all the job openings within the school are identical and thus the expected values for all positions within the schools are identical. The above expression implies that $E\left[y_{mtki} - E\left(y_{mtk} \middle| x_{mtk};\theta\right) \middle| x_{mtk}\right] = 0$; for a school having attributes x_{mtk}, the difference between the attributes of the ith newly hired teacher, y_{mtki}, and the expected mean attributes, given x_{mtk}, is zero in expectation. In turn, this implies that $E\left(x_{mtk}\left[y_{mtki} - E\left(y_{mtk} \middle| x_{mtk};\theta\right)\right]\right) = 0$; across schools, the difference between the actual and expected attributes of the new teachers hired by a school is orthogonal to the school's own attributes.

The sample analog of the last expression is $\sum_t \sum_k \sum_i x_{mtk}\left[y_{mtki} - E(y_{mtk} \mid x_{mtk};\theta)\right] = 0$ which can be rewritten $\sum_t \sum_k n_{mtk} x_{mtk}\left[\bar{y}_{mtk} - E(y_{mtk} \mid x_{mtk};\theta)\right] = 0$, where $\bar{y}_{mtk}$ is the mean attributes of the new teachers employed by the kth school. We employ this moment condition in our estimation. Similarly, we employ $\sum_t \sum_k n_{mtk} x_{mtk}\left[\bar{d}_{mtk} - E(d_{mtk} \mid x_{mtk};\theta)\right] = 0$ which relates the average distance for those newly employed in a school, $\bar{d}_{mtk}$, to the corresponding expectation. We also employ the moment condition $\sum_t \sum_k n_{mtk}\left[\bar{d}_{mtk} - E(d_{mtk} \mid x_{mtk};\theta)\right] = 0$, which relates to the

[22] Berry (1992) makes a similar point in a game-theoretic model of entry in the airline industry.

overall average distance traveled by new teachers in a market. Note that the three moment equations are defined at the market level, implying that there is a set of such conditions for each of the five markets included in the analysis. We do not employ the moment condition $\sum_t \sum_k n_{mtk}[\bar{y}_{mtk} - E(y_{mtk} \mid x_{mtk};\theta)] = 0$. This condition holds for all θ since our analysis only includes candidates who obtained jobs and, thus, the mean attributes of teachers are fixed.

An issue that arises in implementing our estimation strategy concerns the fact that $E\left(y_{mtk} \middle| x_{mtk};\theta\right)$ and $E\left(d_{mtk} \middle| x_{mtk};\theta\right)$ are not easily computed; it is not possible to write out, much less compute, analytical expressions for these expected values. We instead compute values for these expressions using simulation. Let $F(y_{mtk} \mid x_{mtk};\theta)$ represent the approximation of $E(y_{mtk} \mid x_{mtk};\theta)$ obtained through simulation. Similarly, defining $F(d_{mtk} \mid x_{mtk};\theta)$ to be the simulator for $E(d_{mtk} \mid x_{mtk};\theta)$.

Our method for calculating the simulated moments is as follows. (1) A standard-normal random number generator generates H sets of independent draws for the random variables in the model. In each draw, random numbers are generated corresponding to the random variable in each candidate's benefit equation for every school alternative. We denote these values in the h[th] draw using the notation δ_{jk}^h, j = 1,2,…,J and k = 1,2,…,K. Similarly the h[th] draw includes randomly generated values for the random error terms (ω_{jk}^h) in the equations characterizing the benefits to each employer associated with hiring each candidate. These randomly generated values are held constant throughout the estimation, as are the observed attributes of candidates and schools. (2) For a given set of parameter values $(\theta = (\alpha, \beta))$ the simulated moments are obtained as follows. The values of δ_{jk}^h and ω_{jk}^h for a particular draw (h) are used to infer the rankings of candidates and jobs discussed above. In turn, these rankings are used with the Gale-Shapley matching algorithm to determine the school-optimal stable matching and the resulting distribution of teacher and job attributes. In turn, $\bar{y}^h_{mtk} = \frac{1}{n_{mtk}} \sum_{i \in S^h_{mtk}} y^h_{mtk}$ and $\bar{d}^h_{mtk} = \frac{1}{n_{mtk}} \sum_{i \in S^h_{mtk}} d^h_{mtk}$ are computed for each of the K schools hiring in the hth simulation of the outcome in market m during period t. S^h_{mtk} is the set of teachers in school k in draw h. Repeating this step for each of the draws yields the following approximations of the pertinent expected values.

$$F(y_{mtk} | x_{mtk}; \theta) = \frac{1}{H} \sum_h \bar{y}^h_{mtk} \approx E(y_{mtk} | x_{mtk}; \theta)$$

$$F(d_{mtk} | x_{mtk}; \theta) = \frac{1}{H} \sum_h \bar{d}^h_{mtk} \approx E(d_{mtk} | x_{mtk}; \theta)$$

We substitute these expressions into the above moment conditions to get the simulated moment conditions summarized by Equations 1 and 2:

$$\begin{aligned} \psi^a_{mtk} &\equiv x_{mtk} [\bar{y}_{mtk} - F(y_{mtk} \mid x_{mtk}; \theta)] \\ \psi^b_{mtk} &\equiv x_{mtk} [\bar{d}_{mtk} - F(d_{mtk} \mid x_{mtk}; \theta)] \\ \psi^c_{mtk} &\equiv [\bar{d}_{mtk} - F(d_{mtk} \mid x_{mtk}; \theta)] \end{aligned} \qquad \textbf{(1)}$$

$$\psi_m = \sum_t \sum_k n_{mtk} \begin{bmatrix} \psi^a_{mtk} \\ \psi^b_{mtk} \\ \psi^c_{mtk} \end{bmatrix} = \sum_t \sum_k n_{mtk} \psi_{mtk} = 0 \qquad \textbf{(2)}$$

Defining $\psi(\theta)$ to be a column vector containing the stacked values of $\psi_1, \psi_2, \ldots, \psi_5$ for the five markets, the method of simulated moment (MSM) estimator is defined by:

$$\hat{\theta}(W) = \arg\min_{\theta} \; \psi(\theta)' W \psi(\theta) .$$

where W is a symmetric, positive semidefinite weighting matrix. In general, the optimal weighting matrix is $W = \Omega^{-1}$ where $\Omega = AsyVar\,[\psi]$. Given our framework, Ω simplifies to the following block diagonal matrix where the mth diagonal block can be approximated using the formula $\tilde{\Omega}_m = \frac{1}{n_m} \sum_t \sum_k \psi_{mtk}(\tilde{\theta})\, \psi'_{mtk}(\tilde{\theta})$ evaluated at some consistent estimate of θ, $\tilde{\theta}$.

$$\Omega = E[\overline{\psi}\overline{\psi}'] = \begin{bmatrix} E\psi_1\psi_1' & 0 & \cdots & 0 \\ 0 & E\psi_2\psi_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E\psi_5\psi_5' \end{bmatrix} = \begin{bmatrix} \Omega_1 & 0 & \cdots & 0 \\ 0 & \Omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Omega_5 \end{bmatrix}$$

Thus, the efficient MSM estimator in our case will be $\hat{\theta}(\tilde{\Omega}) = \arg\min_{\theta} \sum_{m} \psi_m(\theta)' \tilde{\Omega}_m^{-1} \psi_m(\theta)$.

In the empirical analysis below, we obtain the consistent, but inefficient, estimate of θ, $\tilde{\theta}(I) = \arg\min_{\theta} \sum_{m} \psi_m(\theta)' \psi_m(\theta)$ for the case of an identity weighting matrix, I. In later work, these estimates will be used to compute the $\tilde{\Omega}_m$ used to obtain the second-stage estimates $\hat{\theta}(\tilde{\Omega})$. However, the first-stage estimates are of interest in themselves, since they are consistent estimates of the parameters of interest and can give us a sense of the fruitfulness of the estimation strategy.

The asymptotic covariance matrix of the estimator $\tilde{\theta}$ is $V(\tilde{\theta}) = \frac{1}{n}[D'D]^{-1} D'\Omega D [D'D]^{-1}$ where $D = E_0\left[\frac{\partial\psi(\theta_0)}{\partial\theta'}\right]$ and Ω is defined above. We approximate D using the formula $\tilde{D} = \sum_m \sum_t \sum_k n_{mtk} \frac{\partial \psi_{mtk}(\tilde{\theta})}{\partial\theta'}$ and approximate the block diagonal elements of Ω using the formula for $\tilde{\Omega}_m$ shown above to obtain the standard errors of the first-stage parameter estimates.

Within the burgeoning set of papers employing the method of simulated moments, we know of four papers that have substantial overlap with our application. Epple and Sieg (2001) and Bayer, McMillan and Rueben (2002) employ the method of simulated moments approach to estimate Tiebout equilibrium models of residential choice. Their moment conditions relate to the equilibrium, one-sided sorting of households to local communities. Berry (1992) has employed a simulation estimator to estimate an equilibrium game-theoretic model of market entry in the airline industry, with the simulated moments based on the equilibrium number of firms operating at each airport each year. Sieg (2000) has estimated a bargaining model of medical malpractice disputes.

Even though this analysis focuses on bilateral interactions between individual plaintiffs and defendants, rather than a market-level analysis, the paper is pertinent in that the simulated moments are obtained by repeatedly solving a game-theoretic model for each of a large number of draws of the model's random variables, as is the case in Berry's analysis.

VI. Estimates of Several Models

As the first test of this model we look at the initial sorting of first through sixth grade teachers across schools in the Albany-Schenectady-Troy, Buffalo, Rochester, Syracuse, and Utica-Rome metropolitan areas for school years 1994-95 through 1999-2000. We estimate the following utility functions.

$$\begin{aligned} u_{jk} &= \beta_1(\text{salary}) + \beta_2(\%\text{minority}) + \beta_3(\%\text{poor}) + \beta_4(\text{urban}) + \beta_5(\text{distance}) + \delta_{jk} \\ v_{jk} &= \alpha_1(\text{teacherquality}) + \omega_{jk} \end{aligned} \qquad \textbf{(3)}$$

Thus, teachers' utility is assumed to be a function of salary, the percent of students in the school who are black or Hispanic, the percent of poor students in the school as measured by eligibility for free lunch, whether the school is in an urban area, and distance. Distance is measured from the address given when the individual applied for certification, a point in time typically prior to when individuals apply for teaching jobs. (An alternative distance measure could be based on their location when in high school.) If the distance to all the districts in the labor market where the individual took their first job was greater than 50 miles, the distance measures for all job alternatives were set equal, so that distance would not be a factor in the candidate's choice of jobs. Employers' utility is given solely as a function of teacher qualifications. Teacher qualifications is a composite of (1) whether the teacher ever failed a certification test; (2) the test score on the certification exam; (3) the Barron's rating of his/her undergraduate institution; and (3) whether or not he/she has attained more than a Bachelor's degree. Both equations have normal random errors that are normalized, with no loss of generality, to have standard deviations of one. We then run a number of alternative models as well.

Table 4 presents the sample statistics. Starting salaries average \$32,458 with a small standard deviation of \$2,607. On average 21 percent of students in a school were black or Hispanic and 29 percent were poor. Many more new teachers were hired in recent years. Few (6.4 percent) were black or Hispanic, and for those traveling less than 100 miles to their job, the average distance was only ten miles. For the estimation, salary and distance were normalized to standard deviation units.

The MSM estimations rely on 45 moment conditions. For each of the five labor markets these correspond to teacher quality interacted with each of the four school characteristics (salary, percent minority, percent free lunch, and urban), distance interacted with each of the four school characteristics, and overall average distance. We use 25 draws of the random errors to calculate the simulators and a combination of grid search and derivative techniques to estimate the parameters. We then use 250 draws of the random errors in the simulations used to calculate the derivatives of the moments needed to compute the standard errors of the point estimates.

Before considering the simulation estimates it worth looking at the hedonic results. Table 5 gives the results for the following two equations:

$$\begin{aligned} \text{salary} &= \beta_0 + \beta_1(\text{teacherqualifications}) + \beta_2(\%\text{minority}) + \beta_3(\%\text{poor}) + \beta_4(\text{urban}) + \varepsilon \\ \text{teacherquality} &= \alpha_0 + \alpha_1(\text{salary}) + \alpha_2(\%\text{minority}) + \alpha_3(\%\text{poor}) + \alpha_4(\text{urban}) + \eta \end{aligned} \qquad \textbf{(4)}$$

Fixed effects for years and for metropolitan areas are included in columns II and IV of each panel. Estimates in column III include a dummy variable for whether or not the teacher is non-white and an interaction of non-white with the percent of minority students. Column IV estimates include measures of distance to job: both a continuous measure of distance for those who travel 100 miles or less to their job and a dummy variable for traveling farther.

The hedonic models produce typically inconsistent results. In the traditional specification, with salary on the left-hand-side, wages are higher in schools with higher proportions of minority students (which we might predict, especially if racial composition proxies for other school or neighborhood characteristics that are not appealing to teachers). Yet, there appears to be no premium for better teacher qualifications, and teachers are willing to take lower salaries to teach in schools with high proportions of children in poverty and in urban schools. In the "quality" hedonic, there is again no relationship between quality and salary; but at the same wage, schools with higher proportions of poor students appear to attract less-qualified teachers. This specification shows no relationship between qualifications and either urban or the percent of minority students. The one exception to this is for non-white teachers whose qualifications are lower in high proportion minority schools.[23] Clearly, it would be difficult to draw policy implications from these results.

Table 6 gives the method of simulated moments results. The results corresponding to Equation 3 are in the first panel. Note that all the estimated coefficients are of the expected signs and standard errors are quite small. Teacher qualifications have a positive effect on employer utility. Salary has a positive effect on teacher utility; while percent minority and distance both have negative effects. The coefficients on percent poor and urban are smaller but also negative and statistically significant at traditional levels.

To interpret the size of these effects we can compare the coefficient estimates across variables or compare the size of the effect to the variance of the error (signal to noise). Salary and distance were measured in standard deviation units, $2,607 and 16.25 miles, respectively. Teachers appear to strongly value a school's proximity. The estimated coefficients for distance and salary indicate that an individual would have to be paid an additional $5,758 in order to be willing to take a job five (direct-line) miles further in distance rather than work at a closer, but otherwise identical, school. The utility loss associated with teaching in a school having 30 percent more minority

[23] Market level hedonics produce similarly unintuitive results.

students (approximately one standard deviation) is 0.46, an effect that could be offset by roughly a $3,475 increase in salary.

Teacher qualifications as measured by test scores and college attended contributes to schools' assessments of potential teachers. A one standard deviation increase in qualifications increases utility by 0.35 points. With both the error in this equation and the teacher qualifications factor both having standard deviations equal to one, the overall variance in utility is 1.119 (alpha squared times the variation in qualifications plus the variation of the random error), assuming that qualifications are orthogonal to the error. Thus, our qualifications measure appears to account for somewhat more than ten percent of the total variance in utility.

Why do these results differ from the hedonic estimates? Clearly distance is one factor. The second panel of Table 6 reports results without distance. The coefficient on qualifications drops by more than half though it remains significant. Salary, percent minority, percent poor and urban also continue to significantly affect utility though the relative importance of percent minority increases relative to salary and the overall proportion of variances explained decreased markedly. Overall, distance provides important identification in the standard model but does not explain all of the difference between the MSM results and hedonic results. The structure of the MSM including the matching mechanism is also important.

The third model in Table 6 introduces the race/ethnicity of candidates into the utility function of the employers. This does not substantially change the coefficients on the other variables but does show that employers value minority candidates. They appear to be willing to tradeoff approximately one-half a standard deviation in the quality index for a non-white teacher. Model IV adds an interaction between the measure of school racial composition and whether or not a teacher is non-white. The estimates for the teachers' utility do not show a difference between white and non-white teachers in the effect of the proportion of non-white students. Both sets of teachers prefer schools with lower proportions of minority students. This result could easily arise if this measure of student body composition were proxying for unmeasured characteristics of neighborhoods and schools. Distance continues to play an important role in this specification. When distance is removed from the equation in Model V, the results are qualitatively different. Without adjustments for distance it appears that non-white teachers favor higher percent minority schools. This change is likely a result of non-white teachers geographically clustered near schools with higher proportions of non-white students. The final model in Table 6 adds a squared term for distance. As might be expected the linear is negative and the squared term, positive. Indicating that the effect of distance is stronger when the distance is short.

VII. Conclusion

Descriptive analyses point to a high degree of systematic sorting of teachers across schools. Hedonic wage models have not produced consistent estimates for understanding this sorting. Our first method of simulated moments estimates of the two-sided matching model suggests that this may be a useful route to explore further. Unlike the hedonic models, this matching model produces estimates in keeping with the hypotheses that schools prefer high ability teachers and teachers prefer both higher wages and schools with fewer poor or minority students.

Clearly the model presented here is limited. The negative estimate of the effect of minority students on the utility of both minority and non-minority teachers, for example, suggests that the proportion of minority students in a school may be proxying for other characteristics of the school. Similarly the large estimated effects of distance suggest that some omitted variables may, again, be biasing the results. For example, if most teachers live near schools that are appealing to teach in for unmeasured reasons, the distance measure may pick up some of the effects of these unmeasured attributes. A more complete model would include many more characteristics of both schools and teachers. The specifications in this paper also do not include measures of many potential policy levers -- such as class size, teacher preparation time, school facilities, and other non-instructional resources; and thus, do not provide a roadmap for effective policy. Yet, the estimates suggest that the approach may be useful for providing policy evaluations in the future.

Aside from inclusion of other variables to deal with omitted variables bias and provide estimates of the effect of specific policies, there are a number of other expansions of the model worth pursuing. First, the current sample does not include New York City. Yet New York City is a particularly interesting case. Because there are many more new teachers in the region than in the regions in the current analysis, such inclusion requires substantially more computational power. Second, while the two-sided matching model more easily incorporates heterogeneity of preferences and the discrete nature of choice set than the hedonic wage approach, it does not address the endogeneity of school characteristics. In further work we hope to address this with an instrumental variables approach. For example, because New York State districts have the ability to raise local dollars for schools, we may be able to instrument for salaries using the proportion of non-residential property either in the district in question or in surrounding districts.

The model may also be expanded to address questions of who becomes a teacher and who quits or transfers. The framework allows us to include all potential teachers in the matching process and not just those who took teaching jobs. Similarly, instead of assuming that the only openings are for new teachers in the jobs that new teachers fill, we can allow for vacancy chains. That is, when an opening becomes available because a teacher leaves the system or because the number of teachers in a school increases, we can allow current teachers to move into those spots, creating vacancies in their old schools. Finally, there are market frictions worth dealing with, including information and timing of offers. This paper is clearly only the first step toward understanding the full dimensions of the teacher labor market and the factors that influence teachers' decisions about whether and where to teach and schools' decisions about which teachers to hire.

References

Antos and Rosen (1975). "Discrimination in the Market for Public School Teachers," *Journal of Econometrics, 3*(2) 123-50.

Ballou, D. (1996). "Do public schools hire the best applicants?" *Quarterly Journal of Economics, 111(*1), 97-133.

Ballou, D., & Podgursky, M. (1997). *Teacher pay and teacher quality*. Kalamazoo, MI: E. Upjohn Institute for Employment Research.

Baugh, W.H., & Stone, J.A. (1982). "Mobility and wage equilibration in the educator labor market," *Economics of Education Review, 2*(3), 253-274.

Bayer, P., R. McMillan and K. Rueben (2002) "The Causes and Consequences of Residential Segregation: An Equilibrium Analysis of Neighborhood Sorting," working paper.

Berry, Steven T. (1992) "Estimation of a Model of Entry in the Airline Industry," *Econometrica* 60: 889-917.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2002a). "Initial Matches, Transfers, and Quits: The Role of Teachers' Career Decisions in the Disparities in Average Teacher Qualifications Across Schools," Working Paper.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff (2002b). "The Geography of Teacher Labor Markets: Implications for Policy," Working Paper.

Brewer, D.J. (1996). "Career paths and quit decisions: Evidence from teaching." *Journal of Labor Economics, 14*(2) (April), 313-339.

Bridges, E.M. (1996). "Evaluation for tenure and dismissal." In Millman, J. & Darling-Hammond, L. (Eds.), *The New Handbook of Teacher Evaluation* (pp.147-157). Newbury Park: Sage Publications.

Brown, C (1980). "Equalizing Differences in the Labor Market," *Quarterly Journal of Economics, 94*(1): 113-34.

Burdett and Coles (1999). "Long-term partnership formation: marriage and employment," *The Economic Journal 109*, F306-F334.

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

Dolton, P.J., & Makepeace, G.H. (1993). "Female labour force participation and the choice of occupation," *European Economic Review, 37*: 1393-1411.

Dolton, P.J., & van der Klaaw, W. (1999). "The turnover of teachers: A competing risks explanation," *Review of Economics and Statistics, 81*(3): 543-552.

Epple, Dennis (1999) “Estimating Equilibrium Models of Local Jurisdictions,” *Journal of Political Economy, 107*(4): 645-681.

Epple & Sieg (2001). “Interjurisdictional Sorting and Majority Rule: An Empirical Analysis,” *Econometrica, 69*(6): 1437-65

Freeman, A. M. (1979). “Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues,” *Scandinavian Journal of Economics, 81*(2):154-173.

Gale, D. & L. Shapley (1962). “College Admissions and the Stability of Marriage,” *American Mathematical Monthly 69*: 9-15.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (1999). “Do higher salaries buy better teachers?” Working Paper.

Hanushek, E.A., & Pace, R.R. (1995). “Who chooses to teach (and why)?” *Economics of Education Review, 14*(2), 101-117.

Hwang, H., D.T. Mortensen & W.R. Reed (1998). “Hedonic Wages and Labor Market Search,” *Journal of Labor Economics, 16*(4): 815-47.

Lang, K. & S. Majumdar (2001). “The Pricing of Job Characteristics When Markets Do Not Clear: Theory and Policy Implications,” Working Paper.

Lankford, H. (1999). *A descriptive analysis of the New York State and New York City teaching force. Report prepared for the New York Supreme Court case Campaign for Fiscal Equity v. New York State.*

Lankford, H., S. Loeb and J. Wyckoff (2002). “Teacher Sorting and the Plight of Urban Schools,” *Education Evaluation and Policy Analysis, 24*(1): 37-62.

Lucas (1977). “Hedonic Wage Equations and Psychic Wages in the Returns to Schooling,” *America Economic Review, 67*(4): 549-58.

Manski, C.F. (1987). “Academic ability, earnings, and the decision to become a teacher: Evidence from the National Longitudinal Study of the High School Class of 1972.” In D.A. Wise (Ed.), *Public Sector Payrolls*. Chicago: University of Chicago Press: 291-312.

McLean (1978). “Compensating Wage Differentials for Hazardous Work: An Empirical Analysis,” *Quarterly Review of Economics and Business, 18*(3): 97-107.

Mont, D., & Rees, D.I. (1996). “The influence of classroom characteristics on high school teacher turnover,” *Economic Inquiry, 34*(1): 152-167.

Murnane, R.J., Singer, J.D., & Willett, J.B. (1989). “The influences of salaries and "opportunity costs" on teachers' career choices: Evidence from North Carolina,” *Harvard Educational Review, 59*(3), 325-346.

Palmquist, Raymond B. (1991). “Chapter IV: Hedonic Methods,” in John B. Braden and Charles D. Kolstad (Eds). *Measuring the Demand for Environmental Quality*, Elsevier Science Publishers B.V.: 77-120.

Pflaum, S.W. & Abramson, T. (1990). "Teacher Assignment, Hiring, and Preparation: Minority Teachers in New York City," *Urban Review 22*(1): 17-31.

Rickman, B.D., & Parker, C.D. (1990). “Alternative wages and teacher mobility: A human capital approach,” *Economics of Education Review*, *9*(1): 73-79.

Rosen, S. (1974). “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy 82*(1): 34-55.

Rosen, S. (1986). “The Theory of Equalizing Differences,” in Ashenfelter, E. and R. Layard, *Handbook of Labor Economics Volume I*. Elsevier Science Publishers: 642-692.

Roth, A.E. and M.A. Sotomayer (1990) *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge: Cambridge University Press.

Sieg, H. (2000) “Estimating a Bargaining Model with Asymmetric Information: Evidence from Medical Malpractice Disputes,” *Journal of Political Economy, 108*(5): 1006-1021.

Stinebrickner, T.R. (2000). An analysis of occupational change and departure from the labor force: Evidence of the reasons that teachers quit. Working paper.

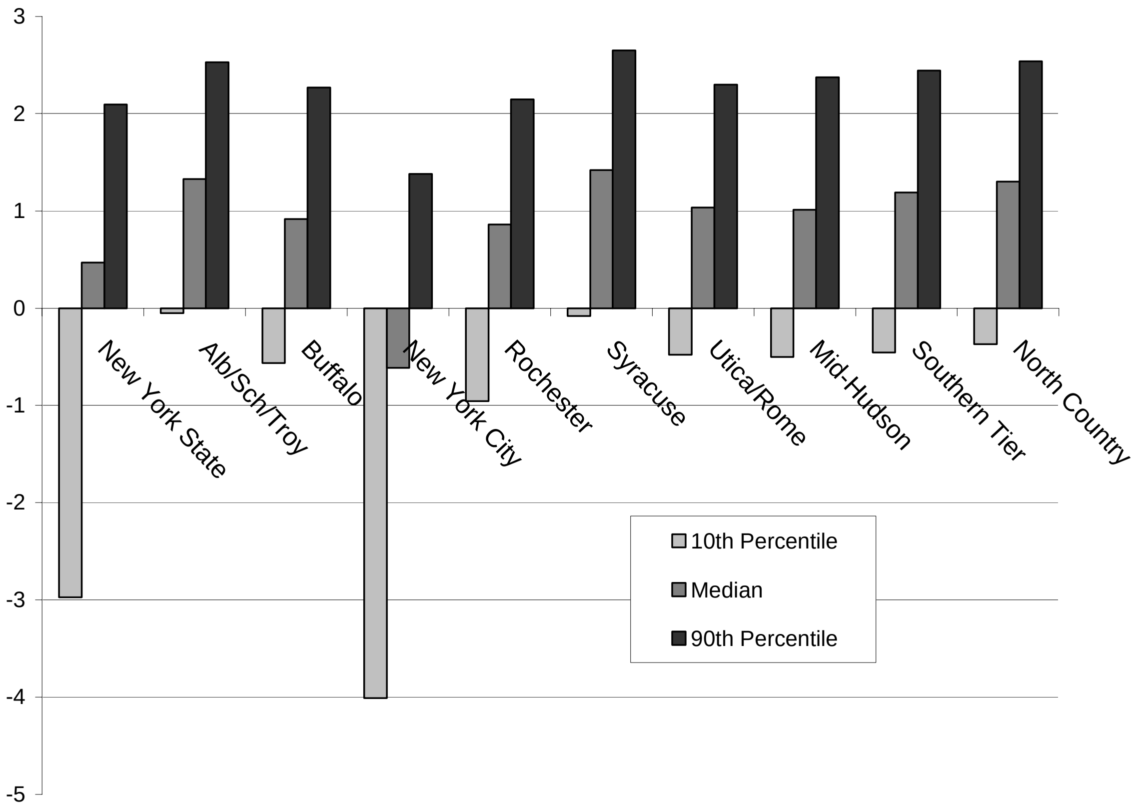
Stinebrickner, T.R. (1999). “Estimation of a duration model in the presence of missing data.” *Review of Economics and Statistics, 81*(3), 529-542.

Stinebrickner, T.R. (1998). “An empirical investigation of teacher attrition,” *Economics of Education Review, 17*(2): 127-136.

Theobald, N.D. (1990). “An examination of the influences of personal, professional, and school district characteristics on public school teacher retention,” *Economics of education review, 9*(3), 241-250.

Theobald, N.D., & Gritz, R.M. (1996). “The effects of school district spending priorities on the exit paths of beginning teachers leaving the district,” *Economics of Education Review, 15*(1), 11-22.

Figure 1: The Distribution of Composite Teacher Qualifications By Region



* Buffalo reflects the schools in the Buffalo MSA outside of the Buffalo City School District. The Buffalo City School District has a certification program that differs from that in the remainder of the State and therefore certification data is not comparable and the composite measure could not be computed.

Figure 2: Percent of Teachers from the 1995 Entering Cohort Who Failed a Teacher Certification Exam, 1995-2000

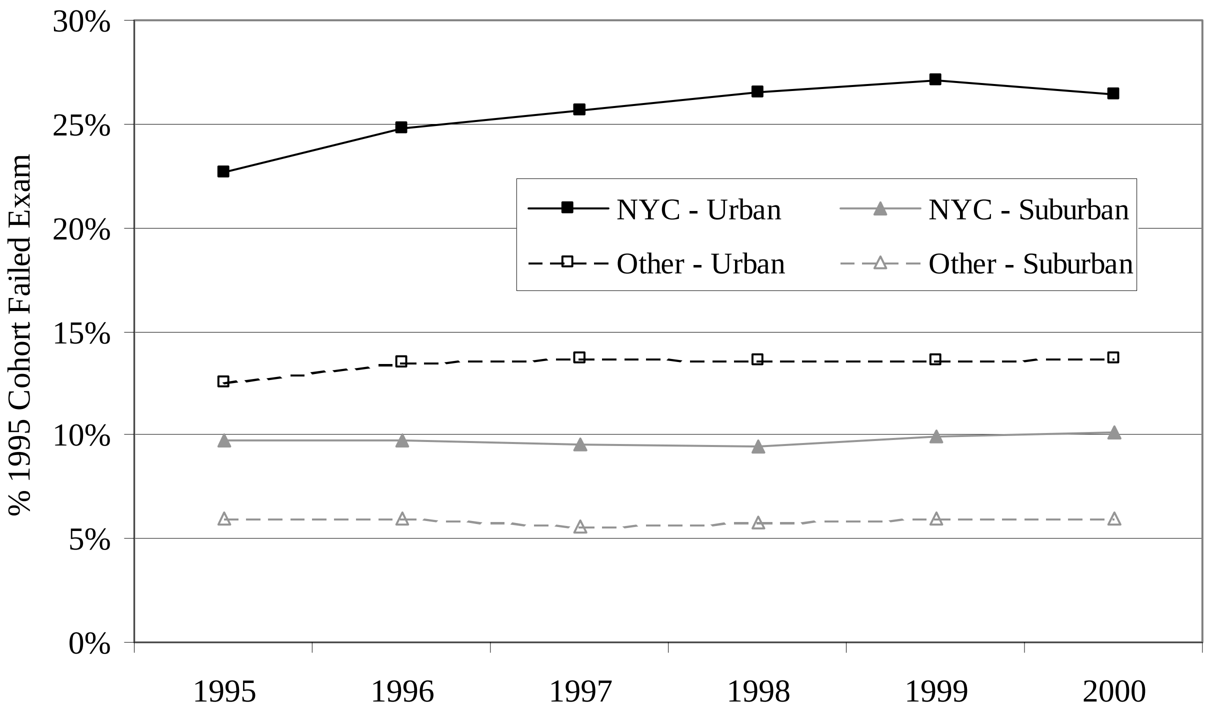


Figure 3: Percent of New York City Teachers from the 1995 Cohort Who Failed a Teacher Certification Exam by Percent Minority Students in the Schools, 1995-2000

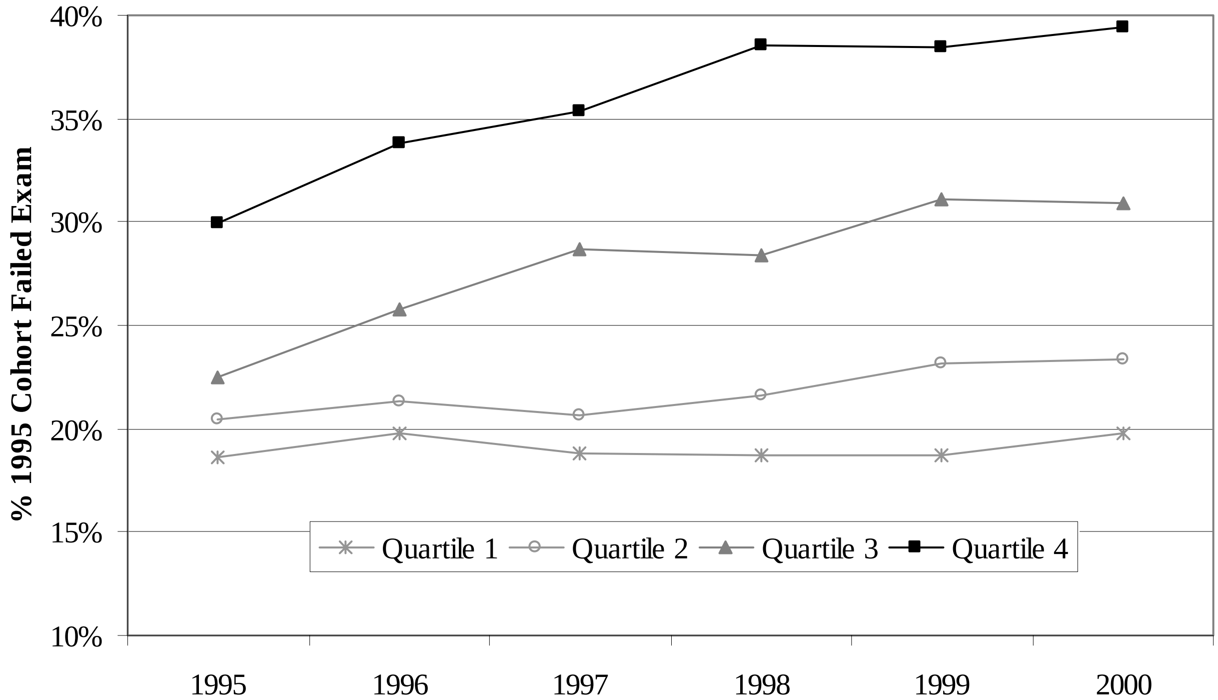


Figure 4: Effect of Distance from Hometown on Employment Location in Region 1 Relative to Region 2

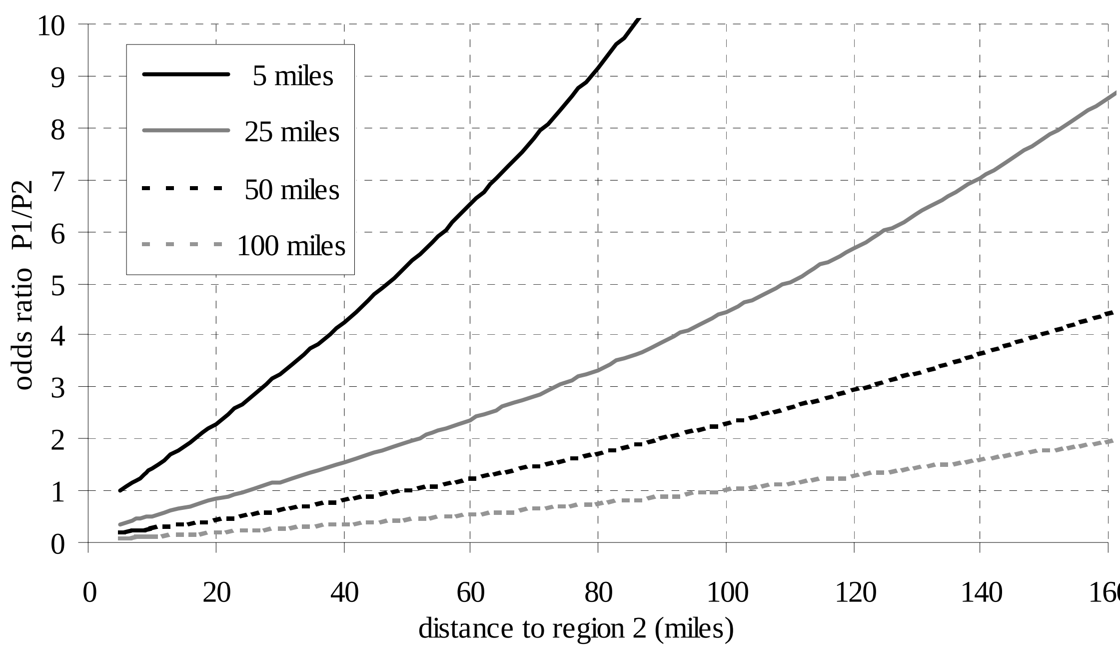


Table 1: New York State Teacher Attributes by MSA, All Teachers 2000 (All teachers FTE > .5)

		Alb/Sch/Troy		Buffalo		New York City		Rochester		Syracuse		Utica/Rome	
		urban	suburb	urban	suburb	urban	suburb	urban	suburb	urban	suburb	urban	suburb
Teacher Qualifications	10th	-0.19	0.04	na	-0.56	-4.99	-1.47	-2.00	-0.55	-0.39	0.03	-0.80	-0.38
Factor	median	1.20	1.40	na	0.92	-1.97	0.70	0.07	1.02	0.87	1.44	0.90	1.13
	90th	1.75	2.62	na	2.27	0.15	1.93	1.45	2.30	2.10	2.70	2.93	2.18
Percent having no prior	10th	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.02	0.00	0.00	0.00
teaching experience	median	0.06	0.06	0.08	0.06	0.10	0.05	0.09	0.06	0.06	0.05	0.06	0.06
	90th	0.17	0.13	0.22	0.15	0.24	0.15	0.18	0.14	0.14	0.14	0.14	0.14
Percent not certified in	10th	0.00	0.00	na	0.00	0.09	0.00	0.06	0.00	0.00	0.00	0.00	0.00
any assignment	median	0.03	0.01	na	0.00	0.20	0.03	0.14	0.02	0.05	0.02	0.02	0.00
	90th	0.10	0.06	na	0.06	0.38	0.10	0.26	0.08	0.11	0.09	0.10	0.06
Percent who failed	10th	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
General Knowledge or	median	0.10	0.00	0.10	0.04	0.27	0.09	0.13	0.00	0.10	0.00	0.12	0.00
Liberal Arts exam	90th	0.20	0.18	0.33	0.20	0.53	0.32	0.25	0.17	0.24	0.19	0.29	0.21
Percent having BAs from	10th	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.14	0.05	0.07	0.00
most competitive colleges	median	0.06	0.08	0.03	0.06	0.07	0.11	0.19	0.22	0.21	0.15	0.14	0.13
	90th	0.13	0.17	0.08	0.13	0.23	0.24	0.25	0.36	0.29	0.29	0.21	0.23
Percent having BAs from	10th	0.00	0.00	0.00	0.00	0.11	0.06	0.00	0.00	0.00	0.00	0.00	0.00
least competitive colleges	median	0.05	0.03	0.06	0.03	0.24	0.15	0.10	0.03	0.05	0.03	0.06	0.03
	90th	0.08	0.10	0.14	0.09	0.42	0.28	0.16	0.09	0.09	0.07	0.11	0.09

Table 2: Teacher Attributes for the Average Student with Given Characteristics

	Overall quality factor	No Teaching Experience	Not Cert in any subject taught	Failed Gen Know or Lib Arts Exam	B.A. from Least Compet. College
New York State					
Non-White	-1.484	0.099	0.166	0.212	0.214
White	0.847	0.067	0.040	0.071	0.102
Poor	-2.393	0.118	0.207	0.279	0.250
Non-Poor	-1.223	0.098	0.159	0.202	0.239
New York City SD					
Non-White	-2.183	0.109	0.212	0.256	0.247
White	-0.726	0.078	0.150	0.161	0.254
Poor	-2.562	0.120	0.215	0.296	0.268
Non-Poor	-1.341	0.100	0.167	0.212	0.258
Rochester City SD					
Non-White	-0.302	0.105	*0.148*	0.107	0.103
White	0.051	0.089	*0.147*	0.099	0.107
Poor	-0.418	0.108	*0.173*	0.120	0.097
Non-Poor	-0.221	0.111	*0.171*	0.111	0.096
Syracuse City SD					
Non-White	1.029	0.080	0.058	0.100	0.045
White	1.254	0.063	0.054	0.095	0.043
Poor	0.970	0.081	0.056	0.109	0.046
Non-Poor	1.194	0.069	0.046	0.103	0.040

* Differences between Non-Whites and Whites and between Poor and Non-Poor significant at the p<.01 level except for those in italics.

Table 3: Distance from High School to First Job, by MSA, 1997-2000

		Distance High School to Job				
Distance College to Job		**0 to 15 miles**	**15 to 40**	**40 to 100**	**100+ miles**	**All**
0 to 15 miles	Row – percent	71.5	17.5	5.5	5.5	100.0
	Col – percent	39.5	24.1	19.3	22.4	32.6
15 to 40 miles	Row – percent	52.8	32.1	8.5	6.7	100.0
	Col – percent	19.7	29.8	20.1	18.4	22.0
40 to 100 miles	Row – percent	50.6	23.8	18.2	7.4	100.0
	Col – percent	15.0	17.7	34.3	16.3	17.6
100 or more miles	Row – percent	54.6	24.2	8.8	12.3	100.0
	Col – percent	25.7	28.4	26.3	42.9	27.8
All	Row – percent	59.0	23.7	9.3	8.0	100.0
	N	10600	4254	1674	1432	17960

Table 4: The Sample: 5028 First Year K-6 Teachers, 2443 Employers

Variable	Mean	Std Dev	Variable	Mean	Std Dev
Qualific. Index	0.00	1.00	Percent Poor, K-6	0.293	0.265
Salary	32,458	2,607	Urban	0.217	
Percent Minority	0.210	0.293	Distance to Job (miles)	24.61	115.27
Minority Teacher	0.064		Distance if < 100 miles	10.29	13.18
Year					
1995	0.109		1998	0.139	
1996	0.123		1999	0.211	
1997	0.151		2000	0.267	
MSAs/Regions					
Albany	0.178		Syracuse	0.167	
Buffalo	0.251		Utica-Rome	0.055	
Rochester	0.350				

Note: Salaries are for 2000. If the 2000 salaries were not available due to districts operating out of contract, we used salary information for the most recent prior year and inflated the value using the average percent change across districts with salaries in both years. Only 4 percent of the sample traveled more than 100 miles to their job.

Table 5: Hedonic and "Quality Hedonic" Results

Variable	Salary				Quality			
	I	II	III	IV	I	II	III	IV
Salary					-0.013 (0.014)	-.018 (0.017)	-0.016 (0.017)	-0.018 (0.017)
Qualification Index	-0.012 (0.014)	-0.012 (0.011)	-0.011 (0.011)	-0.012 (0.011)				
% Minority	***1.30*** ***(0.11)***	***1.37*** ***(0.10)***	***1.35*** ***(0.10)***	***1.37*** ***(0.10)***	0.090 (0.113)	-0.16 (0.12)	-0.030 (0.124)	-0.16 (0.12)
% Poor, K-6	***-1.16*** ***(0.11)***	***-1.12*** ***(0.09)***	***-1.12*** ***(0.09)***	***-1.12*** ***(0.09)***	***-0.62*** ***(0.11)***	***-0.52*** ***(0.12)***	***-0.51*** ***(0.12)***	***-0.52*** ***(0.12)***
Urban	***-0.21*** ***(0.08)***	***-0.18*** ***(0.07)***	***-0.18*** ***(0.07)***	***-0.18*** ***(0.06)***	0.025 (0.079)	0.12 (0.08)	0.11 (0.08)	0.12 (0.08)
Non-white teacher			-0.058 (0.073)				0.13 (0.09)	
% Min * Non-White			0.14 (0.12)				***-0.79*** ***(0.15)***	
Distance				-3.4E-4 (8.8E-5)				*2.0E-3* *(1.1E-3)*
Dist > 100 miles				-0.081 (0.057)				*0.12* *(0.70)*
Year fixed effects MSA fixed effects	No No	Yes Yes	Yes Yes	Yes Yes	No No	Yes Yes	Yes Yes	Yes Yes
R^2	0.0328	0.3593	0.3595	0.3596	0.0189	0.0409	0.0495	0.0410

Note: Standard errors reported in parentheses

Table 6: Estimated Parameters in Employers' and Employees' Criterion Functions

	Model I	Model II	Model III	Model IV	Model V	Model VI
Employers' Criterion Function						
Quality Index	0.345	0.156	0.346	0.346	0.198	0.334
	(0.015)	(0.014)	(0.011)	(0.015)	(0.013)	(0.015)
Non-White			0.163	0.161	0.166	0.169
			(0.007)	(0.008)	(0.009)	(0.006)
Candidates' Criterion Function						
Salary	0.342	0.139	0.343	0.342	0.0902	0.346
	(0.017)	(0.006)	(0.013)	(0.021)	(0.006)	(0.014)
Percent Minority	-1.533	-1.150	-1.532			
	(0.066)	(0.047)	(0.047)			
% Poor, K-6	-0.187	-0.0633	-0.192	-0.192	-0.188	-0.187
	(0.008)	(0.011)	(0.010)	(0.013)	(0.011)	(0.009)
Urban	-0.280	-0.0845	-0.269	-0.274	-0.025	-0.275
	(0.013)	(0.003)	(0.012)	(0.011)	(0.002)	(0.011)
% Minority for Non-White				-1.629	0.945	-1.541
				(0.091)	(0.045)	(0.099)
% Minority for White				-1.535	-1.085	-1.547
				(0.067)	(0.057)	(0.062)
Distance	-2.455		-2.454	-2.432		-3.457
	(0.089)		(0.089)	(0.098)		(0.097)
Distance Squared						0.400
						(0.015)

Note: Standard errors reported in parentheses.

Appendix A

Workforce Database

	Personnel data	*Certification and exam data*	*SUNY student data*	*School and district data*
UNIVERSE:	All public school teachers, superintendents, principals, and other staff	All individuals taking certification exams	All SUNY applicants (including non-teachers)	All public schools and districts
ELEMENTS:	- salary - course subject and grade - class size - experience (district and other) - years of education and degree attainment - age - gender	- scores on NTE and NYSTCE (general knowledge, pedagogy, and content specialty) exams - college of undergraduate and graduate degrees - degrees earned - zip code of residence when certified - race	- high school attended - high school courses - high school GPA - SAT exam scores - college attended and dates - intended college major - actual college major - college GPA - degrees earned	- enrollment - student poverty (free and reduced lunch counts) - enrollment by race - limited English proficiency - student test results - dropout rates - district wealth - district salary schedule - support staff and aides
TIME PERIOD:	1969-70 to 1999-00	1984-85 to 1999-00	1989-90 to 1999-00	1969-70 to 1999-00
SOURCE:	New York State Education Department	New York State Education Department	The State University of New York	New York State Education Department

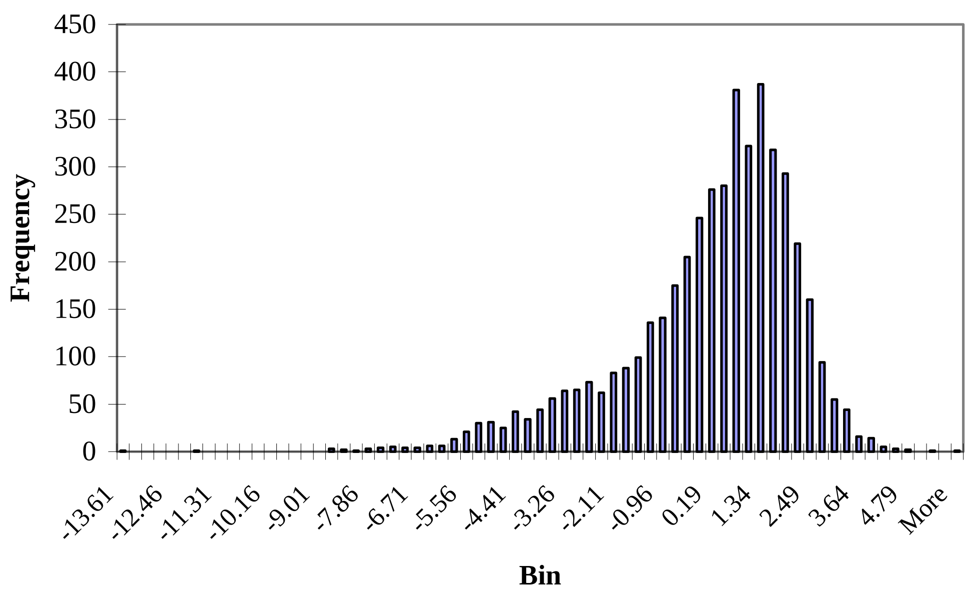
Appendix B

The Composite Measure of Teacher Quality

Components:	**Scoring Coefficients**
1. percent of teachers with less than or equal to 3 years of experience	-0.36449
2. percent of teachers with tenure	0.36032
3. percent of teachers with more than a BA degree	0.31576
4. percent of teachers certified in all courses taught	0.39435
5. percent of teachers from less-competitive or non-competitive colleges	-0.27578
6. average teacher score on the NTE communication skills exam	0.37538
7. average teacher score on the NTE general knowledge exam	0.34601
8. average teacher score on the NTE professional knowledge exam	0.38134

Eigenvalue: 4.17 (52.14% of variation)
Cronbach's alpha (reliability): 0.8641

Figure B1: Histogram of Factor



Appendix C
Additional Results

Table C1: Distance from High School to First Job, by MSA, 1997-2000

Region	0 to 15 miles	15 to 40 miles	40 to 100 miles	100+ miles
Buffalo City	77.6	6.6	4.6	11.2
Buffalo suburbs	71.8	19.3	3.9	5.0
New York City	62.4	27.3	6.5	3.8
New York City Suburbs	69.6	23.5	3.3	3.5
Rochester City	48.7	10.8	21.4	19.1
Rochester Suburbs	42.3	27.0	18.4	12.3
Syracuse City	76.4	6.0	6.0	11.5
Syracuse Suburbs	51.4	22.9	15.3	10.4
Other	48.2	23.0	14.6	14.2
All	59.0	23.7	9.3	8.0

Table C2: Conditional Logit Model of Region Selection

	coefficient	odds ratio	z statistics
Distance from home			
ln(distance)	-0.45	0.64	-5.23
ln(distance)^2	-0.071	0.93	-17.82
ln(distance)3	-0.0083	0.99	-9.32
ln(distance) * female	0.0056	1.01	0.25
* SAT	0.00031	1.00	4.51
* urban	0.027	1.03	0.58
* rural	0.0058	1.01	0.11
Region is home	1.136	3.12	14.34
Region is home * urban	-0.32	0.73	-1.22
Region is home * rural	-0.62	0.54	-4.17
Region and home same type	0.099	1.10	1.31
Region and home same type * urban	0.46	1.59	2.56
Region and home same type * rural	0.46	1.58	2.05
Region is other portion of home metro	-0.12	0.88	-0.56
Distance from college			
ln(distance)	0.030	1.03	0.85
ln(distance)^2	-0.061	0.94	-10.16
ln(distance)^3	-0.0042	0.996	-4.23
ln(distance) * female	-0.065	0.94	-2.58
* rural	0.17	1.19	5.48
Graduated from college in region	0.31	1.36	5.13
Log Likelihood = -12,696.0	*Sample Size = 11,484*		

Table C3: Effects of Location Similarity Net of Distance Effects
Odds ratio of First Job Being in Home Region vs. Various Alternatives

	Alternative Region			
Individual having:	Other part of same metro area	Another metro area: urban portion	Another metro area: suburban portion	Another rural area
Urban Home	5.16	2.26	5.69	5.69
Suburban Home	4.49	4.96	1.97	1.97
Rural Home	n/a	2.35	2.35	1.67

Figure C1: Effect of Distance from Hometown on Employment Location in Region 1 Relative to Region 2 for Teachers with Various Attributes (Region 1 Distance Equals 5 Miles)