

NBER WORKING PAPER SERIES

ANALYZING THE DETERMINANTS OF THE
MATCHING OF PUBLIC SCHOOL TEACHERS TO JOBS:
ESTIMATING COMPENSATING DIFFERENTIALS
IN IMPERFECT LABOR MARKETS

Donald Boyd
Hamilton Lankford
Susanna Loeb
James Wyckoff

Working Paper 9878
<http://www.nber.org/papers/w9878>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2003

The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2003 by Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets
Donald Boyd, Hamilton Lankford, Susanna Loeb, and James Wyckoff
NBER Working Paper No. 9878
July 2003
JEL No. I2, J2, J3, J4, J6, H3, H4, H7

ABSTRACT

Although there is growing recognition of the contribution of teachers to students' educational outcomes, there are large gaps in our understanding of how teacher labor markets function. Most research on teacher labor markets use models developed for the private sector. However, markets for public school teachers differ in fundamental ways from those in the private sector. Collective bargaining and public decision making processes set teacher salaries. Thus it is unlikely that wages adjust quickly to equilibrate the supply and demand for worker and job attributes. The objective of this paper is to develop and estimate a model that more accurately characterizes the institutional features of teacher labor markets. The approach is based on a game-theoretic two-sided matching model and the estimation strategy employs the method of simulated moments. With this combination, we are able to estimate how factors affect the choices of individual teachers and hiring authorities, as well as how these choices interact to determine the equilibrium allocation of teachers across jobs. Even though this paper focuses on worker-job match within teacher labor markets, many of the issues raised and the empirical framework employed are relevant in other settings where wages are set administratively or, more generally, do not clear the pertinent markets for job and worker attributes.

Donald Boyd
University at Albany
SUNY
1400 Washington Avenue
Albany, NY 12222
boydd@rockinst.org

Hamilton Lankford
University at Albany
SUNY
1400 Washington Avenue
Albany, NY 12222
hamp@albany.edu

Susanna Loeb
Stanford University
224 CERAS, 520 Galvez Mall
Stanford University
Stanford, CA 94305
and NBER
sloeb@stanford.edu

James Wyckoff
University at Albany
SUNY
1400 Washington Avenue
Albany, NY 12222
wyckoff@albany.edu

I. Introduction

The 2.8 million elementary and secondary public school teachers in the United States make up 8.5 percent of all college-educated workers 25 to 64 years old.¹ Even though there is growing recognition of the contribution of these teachers to students' educational outcomes and later economic success, large gaps exist in our understanding of how teacher labor markets function. Most research on teacher labor markets has used models developed for the private sector. However, markets for public school teachers, as well as markets for many other public employees, differ in fundamental ways from those in the private sector. The objective of this paper is to develop and estimate a model that more accurately characterizes the institutional features of teacher labor markets. The approach is based on a game-theoretic two-sided matching model and the estimation strategy employs the method of simulated moments. With this combination, we are able to estimate how factors affect the choices of individual teachers and hiring authorities, as well as how these choices interact to determine the equilibrium allocation of teachers across jobs.

Low-income, low-achieving and non-white students, particularly those in urban areas, often are taught by the least skilled teachers, a factor that likely contributes 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. Inefficient hiring and district assignment may contribute to the disparities in teacher qualifications across schools, however teacher preferences are likely to be particularly influential.² Teachers differ fundamentally from other school resources. Unlike

¹ Digest of Education Statistics 2002 and U.S Census Bureau Educational Attainment in the United States 2000 Detailed Tables.

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

textbooks, computers, and facilities, teachers have preferences about whether to teach, what to teach, and where to teach. Potential teachers prefer one type of district to another; and within districts, they prefer one school to another. 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.³ A large literature suggests that teachers respond to wages, yet research on the compensating wage differentials needed to attract teachers with particular attributes to schools with particular characteristics has not produced consistent results.⁴

The inconsistencies in the estimation of compensating differentials for teachers are not surprising given that the estimates have been based on hedonic wage models which maintain that wages adjust to equilibrate the supply and demand for worker and job attributes. This assumption is unlikely to hold for public school teaching, given that salaries are set by collective bargaining and public decision making processes, not directly as a result of market forces. In this context, some jobs may simply be “better jobs” than others, and teachers will sort into these jobs based on their ability to obtain offers from the hiring authorities.

Non-price rationing in the market for public school teachers will result in complex interdependencies in the choices made by job candidates and employers. In particular, a candidate’s willingness to accept a particular job will depend upon her own preferences as well

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.

³ 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.

⁴ 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;

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. We can analyze such an environment in a relatively straightforward manner using the standard two-sided matching model extensively studied by game theorists (Roth and Sotomayer, 1990). The contributions of this paper are to clarify issues regarding the applicability of the standard hedonic model, to note the conceptual applicability of the game-theoretic, two-side matching model and to show how the underlying preferences of job candidates and employers in such a model can be estimated using the method of simulated moments.

Our long-term goal is to identify policies that are effective for attracting and retaining teachers in low-performing or otherwise difficult-to-staff schools. As we discuss further below, such identification has many difficulties, not the least of which is the endogeneity of any district-level 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. We focus on the initial match of teachers to schools in their first job both to simplify the first implementation of the model and because, as we discuss later, the initial match appears particularly important, in comparison to transfers and quits, in determining the disparities in the qualifications of teachers across schools.

The following section of the paper briefly summarizes the data we employ and some key features of teacher labor markets. Section III contrasts the hedonic wage approach with two alternative models of job match. We outline our conceptual framework and empirical approach in section IV and present estimates of several models in section V. Estimates of hedonic wage

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.

equations are reported in section VI and simulation results are discussed that help to clarify issues pertaining to the limitations of the hedonic wage model when used to analyze teacher labor markets. Section VII concludes.

II. Data

The data we use for this analysis comes from a larger database of teachers and schools that 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 approximately 200,000 teachers identified in the PMF. We have linked these annual records through time, yielding detailed data characterizing the career history of each individual.

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, 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 for schools and

districts include enrollment, student poverty, racial composition, and district salary schedules, 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.

Our 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 a series of papers, we have used the data to document various characteristics of teacher labor markets, a number of which are pertinent here.

First, as noted above, there is a marked sorting of teachers across schools. 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 failed.⁵ In contrast, in schools in the lowest quartile of student performance, 22 percent of teachers are uncertified, 26 percent come from least competitive colleges, and 35 percent have failed a general-knowledge certification exam (Lankford, Loeb and Wyckoff, 2002). Similar patterns are found when schools are grouped based on student poverty, race/ethnicity and limited English proficiency. These differences reflect urban-suburban differences in the qualifications of teachers as well as meaningful differences across schools within urban areas.

⁵ 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.

Differences in the qualifications of teachers are the result of the decisions of individuals and school officials that determine initial job matches and subsequent decisions that affect job quits, transfers and terminations. Of these, initial job matches appear particularly important in that they account for almost all of the urban-suburban differences in teacher qualifications as well as a substantial portion of the differences between schools within urban districts. To illustrate this, we track a cohort of entering teachers and assess the spread of teacher qualifications across groups in the first year and then in each following year for that same cohort. We define groups either by urban-suburban-region-status or by quartiles of student characteristics (race/ethnicity or achievement). On the initial match of the 1995 cohort, New York City urban schools had 17.1 percentage points more teachers who had failed a teacher certification exam than did non-NYC suburban schools. This difference had increased by 5.2 points by the end of six years, implying that the initial match accounted for 77 percent of the disparity after six years, when most transfers had already taken place. Within urban areas, the contributions of initial match and exits are roughly equal in determining the overall qualifications of teachers. When we compare the proportion of teachers failing the certification exam in New York City between schools in the top and bottom quartiles of percentage of students who are non-white, the initial gap for the 1995 cohort is 11.3 percentage points. As this cohort ages to the year 2000, the gap enlarges to 19.7 percentage points, implying that quits and transfers have added 8.4 percentage points and that the initial match accounts for 58 percent of the total gap by 2000 (Boyd, Lankford, Loeb, and Wyckoff, 2002a). Given the relative importance of the initial matches of teachers to schools and the need to simplify the first implementation of our empirical model, we focus on these initial job matches and the sorting of teachers within the local labor markets.

A final characteristic of teacher labor markets worth noting is the surprisingly large number of individuals who take their first teaching job very close to where they grew up. Over 60 percent of teachers first teach within 15 miles of the district from which they graduated high school and 85 percent teach within 40 miles. Even of those who travel over 100 miles to college, most return home to teach (Boyd, Lankford, Loeb, and Wyckoff, 2002b).⁶ This proximity has two important implications for modeling the sorting of teachers across jobs. First, most teachers make job choices within a very limited geographic area. Because of this, our empirical analysis of job match presented below, focuses on the matching of teachers to jobs within relatively small geographic areas (metropolitan areas) instead of across the entire State. Second, even within each of these local labor markets, work proximity is likely to affect teachers' rankings of alternative job opportunities. Teachers will rank otherwise identical jobs differently because of differences in the relative proximity of jobs to the teachers' own locations. These ranking differences suggest that an accurate model of teacher labor markets will need to incorporate potential preference heterogeneity.

III. Common Approaches for Modeling Sorting

Before describing our empirical 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.

⁶ Information regarding home location is drawn from either College Board data for all individuals who attended a NYS high school and took the SAT since 1980 or SUNY data for all individuals who applied to a SUNY campus anytime since 1990. Thus, this analysis does not include individuals who did not apply to a SUNY school over this period and (1) attended high school out of state, (2) attended a NYS high school prior to 1980 or (3) attended a NYS high school but did not take the SAT. How the above statistics would be affected by including these individuals in the calculations is not clear as the longer distances traveled by out-of-state students could be more than offset by the many students in New York City who did not take the SAT because they planned to attend CUNY which did not require the test.

Hedonic Wage Models: Most previous studies of teacher labor markets, such as Antos and Rosen (1975), employ a hedonic wage framework, which can be illustrated using a relatively simple example. Suppose that teachers' preferences over job attributes are reflected in the utility function $U = U(S, z | q)$ where S is the salary paid and z is a non-pecuniary job attribute. q is a scalar measure of some teacher attribute which might affect his/her preferences for S and z. In a similar way, $V = V(S, q | z)$ is assumed to characterize each employer's preferences with respect to the salary paid and the teacher attribute conditional on the employer's own attribute, z. For example, suppose that z measures the innate ability of the students in the school and q measures the qualifications of teachers where $\frac{\partial U}{\partial z} > 0$ and $\frac{\partial V}{\partial q} > 0$. Furthermore, assume that U() is an increasing function of S and V() decreases in S.

In the case where salaries clear the market, let $S = S(z, q)$ represent the equilibrium wage function showing the salary a teacher having qualification q would be paid when working in a job having non-pecuniary attribute z. The above assumptions imply that S will be a decreasing function of z and an increasing function of q.

When salaries have adjusted so as to clear the market, each teacher is able to choose the school in which to teach only subject to his/her budget constraint, which would be $S = S(z, q)$.

He/she will choose to teach in a school having the value of z such that $-\frac{\partial S(z, q)}{\partial z} = MRS_{zS}^U$ where

$MRS_{zS}^U \equiv \frac{\partial U / \partial z}{\partial U / \partial S}$ is a teacher's marginal evaluation of z measured in terms of forgone salary.

Similarly, subject only to its budget constraint, each employer will choose a teacher of quality q

such that $\frac{\partial S(z, q)}{\partial q} = MRS_{qS}^V$ where $MRS_{qS}^V \equiv \frac{\partial V / \partial q}{-\partial V / \partial S}$ is the employer's marginal evaluation of q

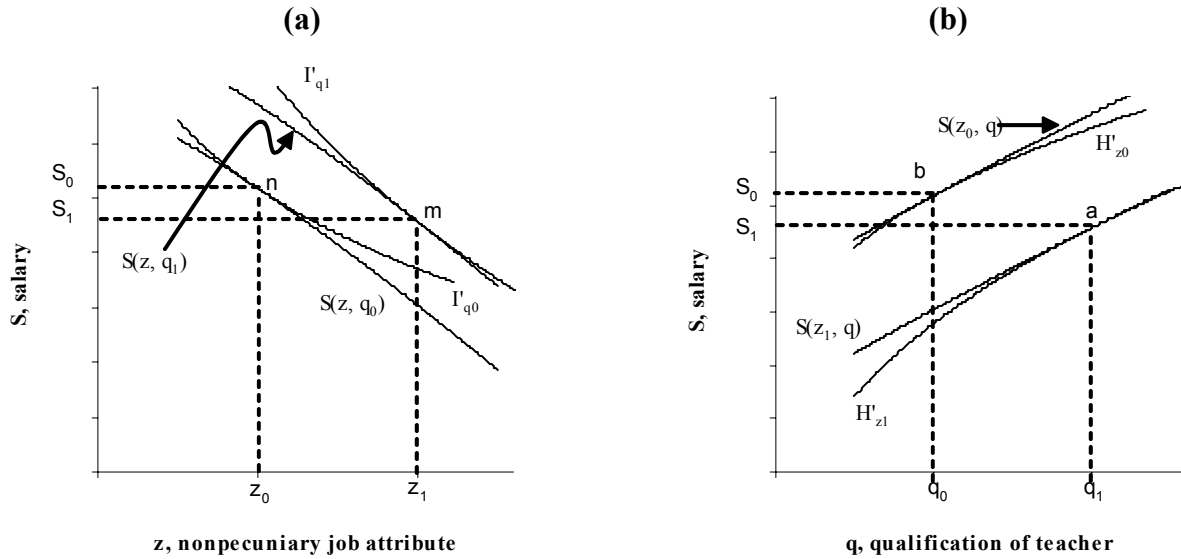
measured in terms of it's willingness to pay a higher salary. The pertinent second-order conditions are assumed to hold.

The equilibrium wage function, $S = S(z, q)$, and the equilibrium allocation of teachers to jobs, here represented by $q = q(z)$, are implied by the preferences of teachers and employers together with the distributions of z and q . For example, in addition to the above assumptions regarding preferences, suppose that more qualified teachers have relatively higher marginal evaluations of z (i.e., MRS_{zS}^U is increasing in q) and that employers having relatively larger values of z have greater willingness to pay for q (i.e., MRS_{qS}^V is increasing in q). In such a case, the equilibrium allocation of teachers to jobs, $q(z)$, will be characterized by positive assortative matching; $\frac{d q(z)}{d z} > 0$ so that schools having higher values of z will employ the teachers having higher values of q . The exact functional form for $q(z)$ will depend upon the distribution of z across jobs and the distribution of q across teachers.

Figures 1a and 1b illustrate such a model. $S(z, q_1)$ in Figure 1a reflects the salary constraint faced by a teacher having qualification q_1 . Her ability to choose any point on this S-z locus together with her preferences imply that he/she will choose to teach in a school having (S_1, z_1) corresponding to point m on indifference curve I'_{q_1} . Similarly, an individual having qualification $q_0 (< q_1)$ would face the salary constraint $S = S(z, q_0)$ and choose the salary-job attribute bundle represented by point n on indifference curve I'_{q_0} . In Figure 1b, an employer having $z = z_0$ would face the salary constraint $S = S(z_0, q)$ and choose to employ a teacher having qualification $q = q_0$ and, correspondingly, pay S_0 . As the location of $S(z_1, q)$ in the figure indicates, an employer having $z_1 (> z_0)$ would be able to pay a relatively lower salary at

each level of q and would choose (S_1, q_1) .⁷ Note that the choices of the teachers in Figure 1a are internally consistent with the choices of employers in 1b; $q_0 = q(z_0)$ and $q_1 = q(z_1)$ where $q(z)$ is the equilibrium matching function defined above.⁸

Figure 1: The Salary, Job Attribute, and Worker Attribute Relationship Used as a Basis for the Hedonic Approach to Estimating Compensating Differentials



The above model can be generalized to include vectors of job (Z) and worker (Q) attributes and, in turn, provide a theoretical basis for the empirical analysis of worker-job match. For example, the assumed equilibrium wage equation, $S = S(Z, Q)$, could be approximated using the function $S_{kj} = Z_k \theta + Q_j \beta + Z_k \Lambda Q_j + \eta_{kj}$ where θ and β are vectors of parameters, Λ is a

⁷ Even though it is not the case here, note that different preferences could have resulted in the employer with $z = z_1$ having chosen a high value of q and a corresponding salary that was higher than S_0 . This would result in a positive correlation between z and S in equilibrium, the opposite of what is shown in figures 1a and 1b.

⁸ The indifference curves in Figures 1a and 1b correspond to the case where $U(S, z | q) = S + z^{0.23} q$, $V(S, q | z) = -S + q^{0.4} z^{0.5}$, $q_0 = 10$, $q_1 = 20$, $z_0 = 10$ and $z_1 = 20$. Furthermore, in deriving the hedonic wage equation, it is assumed that there is an equal number of candidates and employers and that the distributions of q and z are identical which, with positive assortative matching, implies that $q = q(z) = z$. Following steps similar to those employed by Sattinger (1979), it can be shown that $S(z, q) = 0.4444 q^{0.90} - 0.1869 z^{1.23} + C$ where C is a constant

matrix of parameters and η_{kj} is a random disturbance. If the dependent variable were the log wage, the i^{th} element of the column vector $\theta + \Lambda Q_j'$ would be the percent change in the wage needed to attract a worker having attributes Q_j when the value of the i^{th} element in Z_k increased by one unit. Note that parameter estimates for such an equilibrium wage equation, by themselves, typically do not imply estimates of the underlying preference parameters for workers and employers (e.g., the parameters in $U(\cdot)$ and $V(\cdot)$). Inferences regarding such preference parameters typically require a second-stage estimation (Rosen, 1974, 1986).

Even though some labor market studies have carried out such two-stage estimation, there are far more studies that have focused on the first-stage estimation of empirical wage equation parameters. However, estimates of such wage equations in a wide range of settings, including teacher labor markets, have proven inconsistent. Researchers have posited a number of reasons for the inconsistencies including omitted variables (Brown, 1980; Lucas, 1977), simultaneity (McLean, 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, students and teachers and the endogenous determination of pertinent school policies have been offered as possible explanations for counterintuitive hedonic wage results. However, there are other problems with hedonic models in the context of the market for public school teachers, as well as public-sector labor markets more generally.

First, contradicting a basic assumption of the hedonic wage model, public-sector salaries are unlikely to clear their respective markets and, as a result, do not fully adjust for differences in both the attributes of workers and the non-pecuniary attributes of their jobs. Salaries in the public sector often are inflexible because they are set through a combination of political,

of integration that could be determined by specifying the reservation wage associated with alternative employment

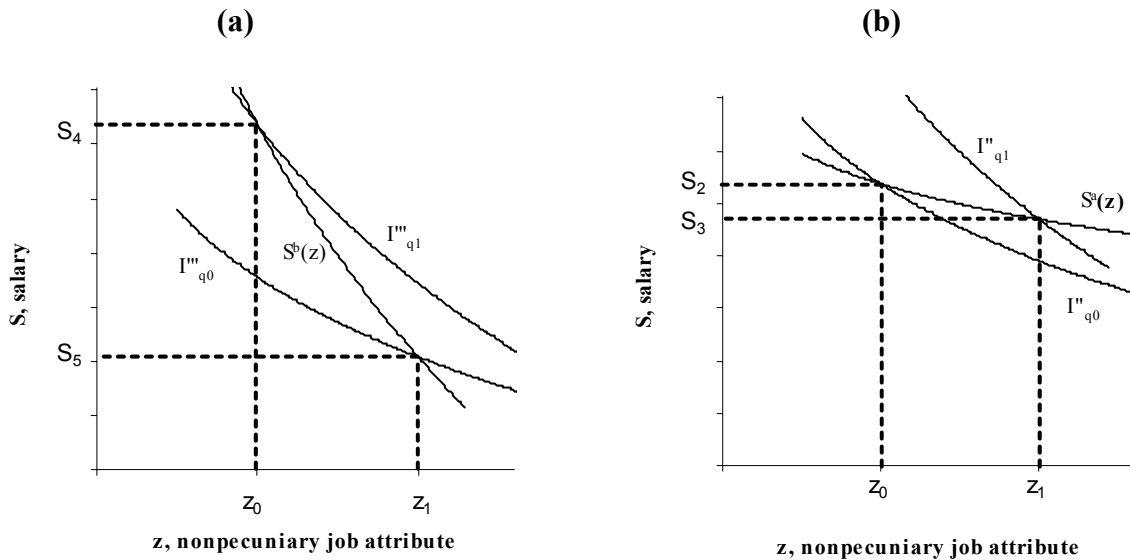
administrative and collective bargaining processes, rather than as a result of direct market forces. In the case of teacher labor markets, union contracts often set teacher salaries for three or more years and social decision-making practices limit both the variation and the flexibility of the wage. Furthermore, district wage schedules typically dictate that all teachers in the district with the same number of years of education and experience earn the same salary, regardless of their other attributes or the characteristics of the schools in which they teach. This limitation is especially restrictive in large urban districts and large countywide districts in which there is considerable within-district variation in the non-wage attributes of schools.

In a setting where teacher salaries do not clear the labor market, those salaries will not reflect compensating differentials for non-pecuniary job characteristics. Consider a modification of the above model where $S = S^*(z)$ reflect a predetermined pattern of wages. Even though $S^*(\cdot)$ also could be a function of q , we consider the simpler case since it is roughly consistent with the institutional feature that salaries do not vary with any of a range of teacher qualifications, other than educational attainment and experience. In this setting, each employer's value of $V = V(S^*(z), q | z)$ will be increasing in q over its entire range so that every employer will want to hire the highest quality teacher possible – a result which differs from that in the hedonic wage model. Whether teachers will prefer higher or lower values of z (i.e., whether $U = U(S^*(z), z | q)$ is increasing or decreasing in z) will depend upon whether MRS_{zS}^U is larger or smaller than $-\frac{dS^*}{dz}$. The general point is that with salaries being predetermined, some jobs simply will provide more utility to workers than other jobs.

in another occupation or home production. In the example, C was arbitrarily set to equal 10.

Figure 2a illustrates the case where the salary reductions teachers are willing to give up in exchange for an increase in z are everywhere smaller in magnitude than the actual salary reduction needed. In such a case, all teachers will prefer the lowest level of z possible so that there will be an excess supply for jobs having low values of z . Figure 2b illustrates the reverse case where $S^*(\cdot)$ is such that all workers prefer jobs having higher values of z , leading to those jobs having an excess supply of teachers.⁹

Figure 2: The Salary, Job Attribute, and Worker Attribute Relationship When Wages Do Not Clear the Market



With price not adjusting so as to equilibrate demand and supply at each level of q and z , a non-price rationing mechanism will be needed to determine the matching of teachers to jobs. We maintain a deferred acceptance mechanism in the game-theoretic, two-sided matching model that underlies our empirical analysis. Here we will leave the mechanism unspecified, but note that there is a fairly obvious outcome when all agents on each side of the market have the same rank

⁹ Rather than $S^*(\cdot)$ being a decreasing function of z as assumed in figure 2, the salary paid could be higher in those schools having higher values of the desirable attribute z ; $dS^*/dz > 0$. The New York City is one such example where both salaries and working are relatively more attractive in the surrounding suburbs. (Lankford, Loeb and Wyckoff (2002)).

ordering of the alternatives on the other side. In such a case, the employer having the job all teachers agree is most attractive would be able to hire its top choice. In turn, the employer having the second most attractive job will be able to hire the individual it ranks highest among the remaining individuals, and so on.

In Figure 2a, $S^*(\cdot)$ is such that all teachers prefer lower values of z . With all employers preferring higher values of q , non-price rationing would result in an allocation characterized by negative assortative matching. In Figure 2b, $S^*(\cdot)$ is such that the resulting allocation would be characterized by positive assortative matching, the same allocation as in the hedonic wage equilibrium. In neither case will the change in salary associated with a change z , $\frac{\partial S}{\partial z}$, reflect teachers' marginal evaluations of z . This results from the fact that teachers are not able to choose jobs subject only to their individual budget constraints. In particular, there is non-price rationing that results from the choices made by the most attractive job candidates and their employers affecting the available alternatives and, thereby, directly constraining the choices available to others. The implications of non-price rationing would be the same when preference heterogeneity or the properties of $S^*(\cdot)$, or a combination of the two, results in MRS_{zS}^U being greater than $-\frac{dS^*}{dz}$ for some values of z and q , with the inequality reversed for other values.

In summary, even if salaries are an important condition of work affecting the allocation of teachers across jobs, there is little reason to believe that those salaries would imply a set of implicit prices that reflect teachers' marginal evaluations of various job attributes. Insufficient price flexibility results in excess demand for the jobs that are relatively more attractive so that the job choices of workers will be constrained by more than just their individual budget constraints. The non-price rationing brought about by insufficient price flexibility will prevent

workers from choosing bundles of job attributes for which their marginal evaluations mirror the implicit prices for those attributes in the market. Thus, the applicability of the standard hedonic model is limited in settings where the choice set for each job candidate is affected directly by the choices made by others beyond the effects of those choices on the candidate's budget constraint. In such settings, there is little reason to think that the wage locus will reflect marginal valuations on either side of the market. While true for teacher labor markets, this will also be the case in other setting where wages are set administratively or for some other reason do not clear the market for job and worker attributes.

The applicability of the hedonic wage model to teacher labor markets is also affected by the apparently small geographic scope of these markets. When labor markets are small, there may not be sufficient numbers of jobs and candidates in each local market to assure that the distributions of employer and employee attributes are continuous. In this setting, discrete choice models such as random utility models are likely to be more pertinent in the analysis of job choice (Freeman 1979; Palmquist 1991).

The importance of distance to teachers making job choices implies not only that labor markets are limited geographically but also that there is likely to be heterogeneity of preferences within each labor market. Teachers may rank the same job differently because of their location relative to the school. Some hedonic models have included distance measures; rent-gradient models, for example, have assumed that individuals prefer living as close as possible to the central city. However, the importance of distance more generally implies that each job candidate has a different assessment of the value of a particular job because the assessment depends upon his/her own location. Such extreme heterogeneity is difficult to incorporate into traditional hedonic models, as the geographical distributions of jobs and job candidates typically will be

quite idiosyncratic from application to application. Yet the implicit prices of various attributes can vary by location depending upon the relative proximity of candidates and jobs having different attributes. Estimates that do not account for the heterogeneity are likely to be biased.

In light of the limitations of applying hedonic wage models to the teacher labor market, we propose to develop and estimate structural models drawing upon the game-theoretic two-sided matching literature. These models will account for pertinent features of teacher labor markets, including wage rigidities and the resulting non-price rationing of jobs and teachers, as well as for factors affecting the separate, but interdependent, choices made by job candidates and school officials. Preferences with respect to distance and other sources of preference heterogeneity enter in a straightforward manner. Furthermore, our estimation strategy allows the underlying preference parameters to be estimated in one step rather than the two-steps required when using the hedonic approach.

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.¹⁰ 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).

¹⁰ 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.¹¹ 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), other conditions of work¹², 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

¹¹ 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.

¹² In teacher labor markets, many conditions of work are inflexible, including the location of the school and student body characteristics. Job attributes such as class size and teacher preparation time, while not completely inflexible, are constrained by the political process and collective bargaining.

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. 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, in part because teachers' rankings of school alternatives are likely to differ reflecting their own proximity to those schools. For this reason, our model builds on the game-theoretic approach, with the hope of incorporating market frictions into later work.

IV. The Model

Consider an environment in which $C = \{c_1, \dots, c_J\}$ represents the set of J individuals seeking teaching jobs and $S = \{s_1, \dots, s_K\}$ 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 $u_{jk} = u(z_k^1, d_{jk} | q_j^2, \beta) + \delta_{jk}$. z_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 q_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 $v_{jk} = v(q_j^1 | z_k^2, \alpha) + \omega_{jk}$ represent the attractiveness of the j^{th} candidate from the perspective of the hiring authority for school k . The vector q_j^1 represents pertinent observed attributes of the j^{th} candidate. The vector z_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 $q_j \equiv (q_j^1, q_j^2)$ for each candidate and $z_k \equiv (z_k^1, z_k^2)$ for each job. Given the vector of parameters β and a particular set of random variable draws for the δ_{jk} , the formula $u_{jk} = u(z_k^1, d_{jk} | q_j^2, \beta) + \delta_{jk}$ implies the matrix of candidates' benefits represented in panel (A) of Figure 3. 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 $v_{jk} = v(q_j^1 | z_k^2, \alpha) + \omega_{kj}$, imply the matrix of school benefits represented in panel (B) of Figure 3 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.

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 $u_{jk} = u(z_k^1, d_{jk} | q_j^2, \beta) + \delta_{jk}$ 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

Figure 3: Utility and Rankings of Candidates and Schools

<p>(A) Candidates' benefits from alternative employment</p> $ \begin{array}{cccc} 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} $	<p>(B) Schools' benefits from alternative candidates</p> $ \begin{array}{cccc} 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} $
<p>(C) Candidates' rankings of employers</p> $ \begin{array}{cccc} 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} $	<p>(D) Schools' rankings of candidates</p> $ \begin{array}{cccc} 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} $

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.¹³

¹³ To see how one can have a model with joint decisions that avoids 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

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), 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

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 interactions. Furthermore, our model allows for unobserved heterogeneity in agents’

procedure in which candidates made offers to hiring authorities would result in an employee-optimal matching.

The equilibrium employer-optimal stable matching corresponding to the alternatives and rankings characterized in Figure 3 is represented in the left side of Figure 4. The right side of Figure 4 characterizes this matching in terms of the resulting relationship between the attributes of candidates and the schools where they are employed.

The matching of candidates to schools represented in Figure 4 corresponds to particular values of the model's random variables (δ_{jk} and ω_{kj} ; $j = 1, 2, \dots, J$ and $k = 1, 2, \dots, K$), the explanatory variables (e.g., q_j and z_k) and the parameters ($\theta = (\alpha, \beta)$) of the model. Given the

Figure 4: Resulting Matching of teachers and Jobs

School-teacher matched pairs	Joint distribution of school and teacher attributes
$\left\{ \begin{array}{l} (s_1, c_j) \\ (s_2, c_{j'}) \\ \dots \\ (s_K, c_{j''}) \end{array} \right\}$	$\begin{bmatrix} z_1 & q_j \\ z_2 & q_{j'} \\ \vdots & \vdots \\ z_K & q_{j''} \end{bmatrix} = [z \ q]$

implied rankings for candidates and employers, deriving such a stable matching is relatively easy using the Gale-Shapley matching algorithm.¹⁴ However, deriving closed-form expressions for the likelihood of observing any particular candidate-job matching or the probability distribution of any particular distribution of worker and job attributes is impossible. To compute the likelihood of a particular stable matching one would need to identify the set of all possible

rankings of alternatives whereas search models typically assume there are common rankings or only very limited heterogeneity.

¹⁴ 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.

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.¹⁵ 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, $m = 1, 2, \dots, M$, and T years, $t = 1, 2, \dots, T$. To account for this generalization, we need only add the subscripts “ m ” and “ t ” to the explanatory and random variables defined above. For example, q_{mtj} represents the attributes of the j^{th} candidate first employed in the m^{th} 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 parallels that for one-to-one matches discussed above.

Let q_{mtki} represent the attributes of the teacher newly employed during period t to fill the i^{th} vacancy of school k in labor market m where $i = 1, 2, \dots, n_{mk}$ and n_{mk} is the total number of job

¹⁵ Berry (1992) makes a similar point in a game-theoretic model of entry in the airline industry.

openings in the k^{th} school for that year. (Reflecting the two-sided match, q_{mki} from the perspective of this employer is the same as q_{mj} defined above from the perspective of the employee where j is the individual employed to fill the k^{th} firm's i^{th} position.) The structure of the two-sided matching model, values of parameters α and β and the distributions of sets of random variables δ_{jk} and ω_{kj} , together imply the joint distribution of z_{mtk} and q_{mki} , $j = 1, 2, \dots, J$ and $k = 1, 2, \dots, K$. This in turn implies the expected value of q_{mki} for the k^{th} school,

$E(q_{mki} | z_{mtk}; \theta)$. Subscript i in this expression can be dropped as a result of the assumption that all the job openings within the school are identical, which implies the expected values for all positions within the schools are identical. The above expression implies that

$E[q_{mki} - E(q_{mki} | z_{mtk}; \theta) | z_{mtk}] = 0$; for a school having attributes z_{mtk} , the difference between the attributes of the i^{th} newly hired teacher, q_{mki} , and the expected mean attributes, given z_{mtk} , is zero in expectation. In turn, this implies that $E(z_{mtk} [q_{mki} - E(q_{mki} | z_{mtk}; \theta)]) = 0$; 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 $\sum_t \sum_k \sum_i z_{mtk} [q_{mki} - E(q_{mki} | z_{mtk}; \theta)] = 0$ which can be rewritten $\sum_t \sum_k n_{mtk} z_{mtk} [\bar{q}_{mtk} - E(q_{mki} | z_{mtk}; \theta)] = 0$, where \bar{q}_{mtk} is the mean attributes of the new teachers employed by the k^{th} school. We employ this moment condition in our

estimation.¹⁶ Similarly, we use $\sum_t \sum_k n_{mtk} z_{mtk} [\bar{d}_{mtk} - E(d_{mtk} | z_{mtk}; \theta)] = 0$ which relates the average distance for those newly employed in a school, \bar{d}_{mtk} , to the corresponding expectation; and $\sum_t \sum_k n_{mtk} [\bar{d}_{mtk} - E(d_{mtk} | z_{mtk}; \theta)] = 0$, which relates to the 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 market included in the analysis. We do not employ the moment condition $\sum_t \sum_k n_{mtk} [\bar{q}_{mtk} - E(q_{mtk} | z_{mtk}; \theta)] = 0$ for the empirical estimates to follow because, 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 $E(q_{mtk} | z_{mtk}; \theta)$ and $E(d_{mtk} | z_{mtk}; \theta)$ 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. Our method of simulated moment estimation strategy is described in Appendix B. In short, the MSM estimator, $\hat{\theta}$, is the value of θ which minimizes a quadratic form defined in terms of the moment conditions. In effect, the parameter estimates minimize the distance between moments reflecting the empirical distribution of teachers across schools and the corresponding theoretical moments implied by our model.

Within the burgeoning set of papers employing the method of simulated moments, we know of three papers that have substantial overlap with our application. Epple and Sieg (2001)

¹⁶ Equivalently, we could have employed $E\left(q_{mtj} \left[z_{mtk} - E\left(z_{mtk} | q_{mtj}; \theta \right) \right] \right) = 0$ and its sample analog $\sum_t \sum_j q_{mtj} \left[z_{mtj} - E\left(z_{mtj} | q_{mtj}; \theta \right) \right] = 0$ where z_{mtj} represents the attributes of the job held by the j th candidate;

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.

V. 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 1999-2000.¹⁷ We estimate the following utility functions.

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

Thus, the j th teachers' utility associated with working in the k th job, u_{jk} , is assumed to be a function of salary, the percent of students in the school who are black or Hispanic, the percent of

$\sum_t \sum_k n_{mtk} z_{mtk} [\bar{q}_{mtk} - E(q_{mtk} | z_{mtk}; \theta)]$ will always equal $\sum_t \sum_j q_{mtj} [z_{mtj} - E(z_{mtj} | q_{mtj}; \theta)]$.

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. (While an alternative distance measure based on their location when in high school may be preferable because it is not endogenous to where teachers hope to teach, we do not have this for all teachers.) If the distance to each district in the labor market where the individual took a 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. The attractiveness of the j^{th} candidate from the perspective of the hiring authority for school k , v_{jk} , is a function of teacher qualifications measured as a scalar composite of (1) whether the teacher ever failed a liberal arts 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.¹⁸ Both equations have normal random errors that are standardized, with no loss of generality, to have standard deviations of one. We then run a number of alternative models.

Table 1 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.

¹⁷ With computational limitations necessitating that we exclude the New York City metropolitan area, our analysis includes the other large metropolitan areas in the state.

¹⁸ We used principal component analysis to determine the weights used in constructing the composite. The eigenvalue is 1.65 and the weightings are 0.6773 for test score, 0.6087 for failing, 0.3089 for college selectivity and 0.1603 for higher degree.

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.

Table 2 gives the method of simulated moments results. The results corresponding to Equation 3 are in the first column. 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). For example, the utility loss associated with teaching in a school having 30 percent more minority students (approximately one standard deviation) is 0.46, an effect that could be offset by roughly a \$3,475 increase in salary (1.3 standard deviations). 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 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.

As noted above, a potential advantage of the empirical model developed here is the ease with which preference heterogeneity can be taken into account, in particular the heterogeneity

resulting for teacher-job proximity. The large magnitude of the distance coefficient estimate underscores that this is important. To investigate the importance of accounting for distance further, the second column of Table 2 reports results with distance omitted from the model. 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 is an important explanatory variable and provides important identification in the standard model.¹⁹

The third model in Table 2 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

¹⁹ It is not surprising that the estimated coefficients for salary, percents minority and poor and the urban dummy fall when distance is dropped from the criterion function for candidates; a proportionate reduction in all these coefficients is equivalent to an increase in the standard deviation of the random error in the equation which we normalized to equal one. Dropping distance from the equation results in it being subsumed in the error term. These parameter estimates not changing proportionately and the coefficient for the quality index in the employers' criterion function also showing a marked change both likely reflect the fact that the geographical distributions of candidates and schools vary systematically with respect to their attributes.

higher percent minority schools. This change is likely a result of non-white teachers being geographically clustered near schools with higher proportions of non-white students. The final model in Table 3 adds a squared term for distance. As might be expected the linear term is negative and the squared term, positive, indicating that the effect of distance is stronger when the distance is short.

VI. Hedonic Model Estimates and Simulations

It is constructive to compare the above estimates with those from the traditional wage-equation approach. Table 3 reports parameter estimates for the following two equations:

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

We include a specification having quality as the dependent variable because some studies have used this approach as an alternative to the traditional wage equation (Loeb and Page, 2001). Furthermore, such a specification corresponds to the function $q = q(z)$ reflecting the equilibrium allocation of teacher to jobs. Note that such a function has relevance whether or not the wage clears the market. 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 wage models produce typically inconsistent results. In the wage equation, salaries 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 equations, 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.²⁰ Clearly, it would be difficult to draw policy implications from these inconsistent results.

Given the wide use of the hedonic model, it is pertinent to investigate further why a wage model and our empirical two-sided matching model yield such different results. We do this by carrying out Monte Carlo simulations. Preferences are assumed to be as follows for 150 employers each having on average 3 openings and 450 teachers seeking those positions.

$$u_{jk} = \beta_1 (Z_1) + \beta_2 (Z_2) + \beta_3 (salary) + \beta_4 (\text{distance}) + \sigma_\delta \delta_{jk}$$

$$v_{jk} = \alpha (\text{teacherquality}) + \sigma_\omega \omega_{jk}$$

The locations of teachers and schools are represented by scalar variables L_T and L_S , respectively, so that a teacher's distance to a particular school equals $|L_S - L_T|$. The values of L_S , Z_1 , Z_2 and *salary* for each school, the values of L_T and *teacherquality* for each teacher as well as the values of the errors terms δ_{jk} and ω_{jk} were obtained by making 100 sets of independent random draws from the standard normal distribution. For given values of the preference parameters and the

²⁰ Market-level hedonics produce similarly unintuitive results.

standard deviations σ_{δ} and σ_{ω} , the teacher-employer stable matching implied by the matching algorithm underlying our model was determined for each of the 100 draws. In turn, the following salary and quality equations were estimated for each draw and mean values of the parameter estimates were computed for the given values of the parameters and correlations.

$$\text{salary} = \gamma_0 + \gamma_1 (Z_1) + \gamma_2 (Z_2) + \gamma_3 (\text{teacherquality}) + \varepsilon$$

$$\text{teacherquality} = \tau_0 + \tau_1 (Z_1) + \tau_2 (Z_2) + \tau_3 (\text{salary}) + \xi$$

In this way we investigate how differing (i.) the degree of correlation among the variables and (ii.) the preferences of teachers and schools affect parameter estimates in the salary and quality equations. A number of general trends emerge which are illustrated in Table 4 where the first number in each cell is the average of the parameter estimates and the second is the proportion of the 100 estimates that are statistically significant ($p < .05$).

First, when there is no correlation among variables in the model and teachers do not have preferences over distance, the wage equation gives coefficients that qualitatively reflect preferences. In Comparison 1 β_3 , α , σ_{δ} , and σ_{ω} all equal 1 and β_4 equals zero. If β_1 and β_2 equal zero the mean estimates of γ_1 , γ_2 and γ_3 in the wage equation equal -0.0003, -0.009 and 0.699, respectively. The estimates for γ_3 are statistically significant in all of the simulations, while those for γ_1 and γ_2 are significant 24 percent and 19 percent of the time. If β_1 and β_2 increase to 0.3 and 0.6, respectively, the estimates change to -0.13, -0.27, and 0.73. If β_1 and β_2 increase again to 0.6 and 1.2, the estimates change to -0.22, -0.45 and 0.78, respectively, with the estimates of γ_2 and γ_3 statistically distinguishable from zero in all simulations and the estimate of γ_1 statistically significant in all but one draw.

Second, when salary is correlated with another school characteristics, Z_1 , the coefficient on Z_1 reflects that correlation, even if candidates do not value Z_1 . Consider the same example as above, except with β_1 and β_2 equal to zero (Comparison 2). When the correlation between Z_1 and salary equals zero, the mean estimates of γ_1 , γ_2 and γ_3 equal -0.0003, -0.009, and 0.70. The estimate of γ_1 is significant in 24 percent of the simulations. When the correlation is 0.3, the estimates are 0.16, -0.01 and 0.67; and γ_1 is significant in 97 percent of the simulations. When the correlation is 0.6, the coefficient estimates are 0.38, -0.01 and 0.54, not reflecting the underlying preferences for Z_1 at all. Furthermore, the estimate of γ_1 is significant in all of the simulations.

Third, when distance enters candidates' preferences, or similarly when a relative increase in noise raises the variance of the errors, the estimated coefficients in the wage equation drop in magnitude. This happens even when distance is not correlated with any other measure. Comparison 3 uses the reference parameter values (β_1 , β_2 , and β_3 equal 0.5, 0.5 and 1.0, respectively). If β_4 equals zero, the mean estimates of γ_1 , γ_2 and γ_3 are -0.22, -0.23 and 0.74. When β_4 equals -0.5, the mean estimates change to -0.21, -0.22 and 0.72; and when β_4 equals -1.0, the mean estimates change to -0.20, -0.21 and 0.69. When β_4 equals -1.5, the mean estimates fall further to -0.19, -0.20 and 0.67. Increases in σ_δ and σ_ω also decrease the estimates in the wage equation (results not shown in Table 4). If β_1 and β_2 equal zero and β_3 equals 1.0, the mean estimate of γ_3 is 0.70 when the standard deviations of the errors equal 1.0. When the standard deviations drop to 0.5, the mean estimate of γ_3 increases to 0.90. When they increase to 1.5, $\tilde{\gamma}_3$ drops to 0.50.

Fourth, if candidates prefer closer schools and schools that are closer to more qualified candidates systematically differ in their characteristics, then the estimated wage equation will misrepresent the value teachers place on these characteristics. Using the example in which β_1 , β_2 , β_3 and β_4 equal 0.5, 0.5, 1.0, and -1.0 respectively (Comparison 4), if the correlations both between qualifications and teacher location and between Z_1 and school location are zero then the mean estimated coefficient for Z_1 is -0.20 (statistically significant in 96 percent of simulations). When these correlations equal 0.3,²¹ the mean estimated coefficient is -0.21 . When the correlations equal 0.6, the mean estimate is -0.261 , one-third larger in magnitude compared to the case where there is no such spatial proximity.

Finally, while the quality equation more accurately reflects the underlying preferences than the wage equation in some instances, it is also subject to potential biases. For example, the increasing correlation between Z_1 and salary has little effect on the predicted relationship between salary and quality in the quality equation, while it reduces the estimated relationship substantially in the wage equation (Comparison 2). Furthermore, in Comparisons 1 and 2 the tests of statistical significance for the quality equation yield results that more accurately reflect the underlying parameter values, more so than the tests of statistical significance for the wage equation. On the other hand, increased error such as that resulting from the importance of distance has an approximately equal effect on the estimated relationship in the salary and quality equation estimates (Comparisons 3). The same is true for increasing the geographical proximity of more qualified teachers and schools having higher values of Z_1 (Comparison 4).

In summary, the simulations suggest that the standard wage model approach for estimating compensating differentials is seriously flawed. These simulations, the theoretical

²¹ As these correlations increase, teachers having greater qualifications, on average, live increasingly close to schools having higher values of Z_1 .

considerations discussed in section III, and the stark differences in the empirical results described above together cause us to conclude that the matching model employed here provides a preferable framework for analyzing teacher labor markets.

VII. Conclusion

Our descriptive analyses of teacher labor markets point to a high degree of systematic sorting of teachers across schools. Yet, hedonic wage models have not produced consistent estimates for understanding this sorting. In this paper we have used a method of simulated moments estimate of the two-sided matching model. The results suggest that this may be a useful estimation strategy to explore further. Unlike the hedonic models, our empirical 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. Our findings of the relative importance of school proximity and the interaction of geographical proximity and school racial composition is relevant to policy. The negative estimate of the effect of minority students on the utility of both minority and non-minority teachers once distance is accounted for in the model suggests that the proportion of minority students in a school may be proxying for other characteristics of the school. In future work we hope to include alternative distance measures (e.g., distance from college and distance for home measured in terms of where individuals attended high school) as well as an indicator of whether a school is in the district where one attended high school. Other important extensions will be to analyze the effects of potential policy levers, such as class size, teacher preparation time, school facilities, and other non-instructional resources and gain access to the needed additional computational power needed to analyze the New York City metropolitan area.

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 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.

Even though this paper focuses on worker-job match within the context of teacher labor markets, the issues raised and the empirical framework employed are relevant in other settings where wages are set administratively or, more generally, do not clear the pertinent markets for job and worker attributes. The theoretical points made in section III, the simulations in section VI and the differences between the estimates for the hedonic and two-sided matching models bring into question the common practice of employing hedonic models to estimate compensating differentials in such settings. Even though there is little reason to think that the wage locus in such cases will reflect marginal evaluations on either side of the market, the need to estimate compensating differentials and empirically analyze the functioning of such markets remains. The empirical framework and estimation strategy developed in this paper, thus, may prove useful in a range of other applications.

In summary, his paper is a step toward understanding the functioning of teacher labor markets and the factors that influence teachers' decisions about whether and where to teach and schools' decisions about which teachers to hire. The matching model shows promise for estimating compensating differentials and the preferences of both employers and employees in labor markets not characterized by perfect competition and the rapid adjustment of wages.

References

- Antos, J., and S. 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., and M. Podgursky (1997). *Teacher pay and teacher quality*. Kalamazoo, MI: E. Upjohn Institute for Employment Research.
- Baugh, W.H., and J.A. Stone (1982). "Mobility and wage equilibration in the educator labor market," *Economics of Education Review*, 2(3), 253-274.
- Berry, S. 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 (2003). "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, K. and M. 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., and G.H. Makepeace (1993). "Female labour force participation and the choice of occupation," *European Economic Review*, 37: 1393-1411.
- Dolton, P.J., and W. van der Klaaw (1999). "The turnover of teachers: A competing risks explanation," *Review of Economics and Statistics*, 81(3): 543-552.

- Epple, D. and H. 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. and L. Shapley (1962). "College Admissions and the Stability of Marriage," *American Mathematical Monthly* 69: 9-15.
- Hanushek, E.A., J.F. Kain, and S.G. Rivkin (1999). "Do higher salaries buy better teachers?" Working Paper.
- Hanushek, E.A., and R.R. Pace (1995). "Who chooses to teach (and why)?" *Economics of Education Review*, 14(2), 101-117.
- Hwang, H., D.T. Mortensen and W.R. Reed (1998). "Hedonic Wages and Labor Market Search," *Journal of Labor Economics*, 16(4): 815-47.
- Lang, K. and 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," *American 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, H. M. (1978). "Compensating Wage Differentials for Hazardous Work: An Empirical Analysis," *Quarterly Review of Economics and Business*, 18(3): 97-107.
- Mont, D., and D.I. Rees (1996). "The influence of classroom characteristics on high school teacher turnover," *Economic Inquiry*, 34(1): 152-167.
- Murnane, R.J., J.D. Singer, and J.B. Willett (1989). "The influences of salaries and "opportunity costs" on teachers' career choices: Evidence from North Carolina," *Harvard Educational Review*, 59(3), 325-346.

- Palmquist, R. 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. and T. Abramson (1990). "Teacher Assignment, Hiring, and Preparation: Minority Teachers in New York City," *Urban Review* 22(1): 17-31.
- Rickman, B.D., and C.D. Parker (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.
- Sattinger, M. (1979) "Differential Rents and the Distribution of Earnings," *Oxford Economic Papers*, 31(1): 60-71.
- 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., and R.M.Gritz, (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.

Table 1: 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 2: 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 (0.015)	0.156 (0.014)	0.346 (0.011)	0.346 (0.015)	0.198 (0.013)	0.334 (0.015)
Non-White			0.163 (0.007)	0.161 (0.008)	0.166 (0.009)	0.169 (0.006)
Candidates' Criterion Function						
Salary	0.342 (0.017)	0.139 (0.006)	0.343 (0.013)	0.342 (0.021)	0.0902 (0.006)	0.346 (0.014)
Percent Minority	-1.533 (0.066)	-1.150 (0.047)	-1.532 (0.047)			
% Poor, K-6	-0.187 (0.008)	-0.0633 (0.011)	-0.192 (0.010)	-0.192 (0.013)	-0.188 (0.011)	-0.187 (0.009)
Urban	-0.280 (0.013)	-0.0845 (0.003)	-0.269 (0.012)	-0.274 (0.011)	-0.025 (0.002)	-0.275 (0.011)
% Minority for Non-White				-1.629 (0.091)	0.945 (0.045)	-1.541 (0.099)
% Minority for White				-1.535 (0.067)	-1.085 (0.057)	-1.547 (0.062)
Distance	-2.455 (0.089)		-2.454 (0.089)	-2.432 (0.098)		-3.457 (0.097)
Distance Squared						0.400 (0.015)

Note: Standard errors reported in parentheses.

Table 3: Hedonic and “Quality Hedonic” Results

Variable	Salary				Quality			
	I	II	III	IV	I	II	III	IV
Salary					-0.013 (0.014)	-0.18 (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	No	Yes	Yes	Yes	No	Yes	Yes	Yes
MSA fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
R2	0.0328	0.3593	0.3595	0.3596	0.0189	0.0409	0.0495	0.0410

Note: Standard errors reported in parentheses

Table 4
Mean Estimates of Parameters in Wage and Quality Equations
 Monte Carlo Simulations Varying Either the Preference Parameters or Correlations Between Explanatory Variables

	values of preference parameters				mean estimates of parameters in wage equation			mean estimates of parameters in quality equation				
	β_1	β_2	β_4	ρ_{Z_1S}	ρ_{qL_s}	ρ_{Z_1L}	$\tilde{\gamma}_1$	$\tilde{\gamma}_2$	$\tilde{\gamma}_3$	$\tilde{\tau}_1$	$\tilde{\tau}_2$	$\tilde{\tau}_3$
Comparison 1: Variation in the preference parameters for Z_1 and Z_2	0.0	0.0	0.0	0.0	0.0	0.0	-0.0003	-0.009	0.699	0.004	-0.002	0.692
	0.3	0.0	0.0	0.0	0.0	0.0	0.24	0.19	1.00	0.08	0.08	1.00
	0.6	1.2	0.0	0.0	0.0	0.0	-0.131	-0.274	0.732	0.183	0.363	0.609
Comparison 2: Variation in the correlation between Z_1 and salary	0.0	0.0	0.0	0.0	0.0	0.0	-0.0003	-0.009	0.699	0.004	-0.002	0.692
	0.0	0.0	0.0	0.3	0.0	0.0	0.24	0.19	1.0	0.08	0.08	1.0
	0.0	0.0	0.0	0.6	0.0	0.0	0.165	-0.011	0.667	0.001	0.001	0.693
Comparison 3: Variation in teachers' distance preference parameter	0.0	0.0	0.0	0.6	0.0	0.0	0.380	-0.011	0.543	-0.006	0.002	0.696
	0.5	0.5	0.0	0.0	0.0	0.0	1.0	0.27	1.0	0.04	0.05	1.0
	0.5	0.5	-0.5	0.0	0.0	0.0	-0.219	-0.231	0.736	0.302	0.302	0.599
Comparison 4: Variation in the correlations between teachers' location and q and between school location and Z_1	0.5	0.5	-1.0	0.0	0.0	0.0	0.98	0.99	1.0	1.0	1.0	1.0
	0.5	0.5	-1.5	0.0	0.0	0.0	-0.208	-0.224	0.719	0.296	0.299	0.590
	0.5	0.5	-1.0	0.0	0.0	0.0	0.99	0.99	1.0	1.0	1.0	1.0
	0.5	0.5	-1.5	0.0	0.0	0.0	-0.195	-0.211	0.690	0.289	0.292	0.572
	0.5	0.5	-1.0	0.0	0.0	0.0	0.96	0.96	1.0	1.0	1.0	1.0
	0.5	0.5	-1.5	0.0	0.0	0.0	-0.186	-0.200	0.668	0.284	0.284	0.558
	0.5	0.5	-1.0	0.0	0.0	0.0	0.95	0.97	1.0	1.0	1.0	1.0
	0.5	0.5	-1.0	0.0	0.0	0.0	-0.195	-0.211	0.690	0.289	0.292	0.572
	0.5	0.5	-1.0	0.0	0.3	0.0	0.96	0.96	1.0	1.0	1.0	1.0
	0.5	0.5	-1.0	0.0	0.6	0.0	-0.211	-0.207	0.688	0.313	0.286	0.561
	0.5	0.5	-1.0	0.0	0.6	0.0	0.97	0.97	1.0	1.0	1.0	1.0
	0.5	0.5	-1.0	0.0	0.6	0.0	-0.261	-0.197	0.687	0.387	0.271	0.529
						1.0	0.96	1.0	1.0	1.0	1.0	

* The parameters α , β_3 , σ_ω and σ_δ are all set equal to one in each of the comparisons.

Appendix A

Workforce Database

	<i>Personnel data</i>	<i>Certification and exam data</i>	<i>SUNY student data</i>	<i>School and district data</i>
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:	<ul style="list-style-type: none"> - salary - course subject and grade - class size - experience (district and other) - years of education and degree attainment - age - gender 	<ul style="list-style-type: none"> - 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 	<ul style="list-style-type: none"> - 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 	<ul style="list-style-type: none"> - 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

As noted in Section V, it is necessary to use simulation to compute values of $E(q_{mtk} | z_{mtk}; \theta)$ in the moment condition $\sum_t \sum_k n_{mtk} z_{mtk} [\bar{q}_{mtk} - E(q_{mtk} | z_{mtk}; \theta)] = 0$ and $E(d_{mtk} | z_{mtk}; \theta)$ in the moment conditions $\sum_t \sum_k n_{mtk} z_{mtk} [\bar{d}_{mtk} - E(d_{mtk} | z_{mtk}; \theta)] = 0$ and $\sum_t \sum_k n_{mtk} [\bar{d}_{mtk} - E(d_{mtk} | z_{mtk}; \theta)] = 0$. Let $F(q_{mtk} | z_{mtk}; \theta)$ be the approximation of $E(q_{mtk} | z_{mtk}; \theta)$ obtained through simulation; and $F(d_{mtk} | z_{mtk}; \theta)$, the simulator for $E(d_{mtk} | z_{mtk}; \theta)$.

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 hth draw using the notation δ_{jk}^h , $j = 1, 2, \dots, J$ and $k = 1, 2, \dots, K$. Similarly the hth 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, $\bar{q}_{mtk}^h = \frac{1}{n_{mtk}} \sum_{i \in S_{mtk}^h} q_{mtk}^h$ and

$\bar{d}_{mtk}^h = \frac{1}{n_{mtk}} \sum_{i \in S_{mtk}^h} d_{mtk}^h$ are computed for each of the K schools hiring in the hth simulation of the outcome

in market m during period t . S_{mtk}^h 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.

$$F(q_{mtk} | z_{mtk}; \theta) = \frac{1}{H} \sum_h \bar{q}_{mtk}^h \approx E(q_{mtk} | z_{mtk}; \theta)$$

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

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

$$\begin{aligned} \psi_{mtk}^a &\equiv z_{mtk} [\bar{q}_{mtk} - F(q_{mtk} | z_{mtk}; \theta)] \\ \psi_{mtk}^b &\equiv z_{mtk} [\bar{d}_{mtk} - F(d_{mtk} | z_{mtk}; \theta)] \\ \psi_{mtk}^c &\equiv [\bar{d}_{mtk} - F(d_{mtk} | z_{mtk}; \theta)] \end{aligned} \quad (1)$$

$$\psi_m = \sum_t \sum_k n_{mtk} \begin{bmatrix} \psi_{mtk}^a \\ \psi_{mtk}^b \\ \psi_{mtk}^c \end{bmatrix} = \sum_t \sum_k n_{mtk} \psi_{mtk} = 0 \quad (2)$$

Defining $\psi(\theta)$ to be a column vector containing the stacked values of $\psi_1, \psi_2, \dots, \psi_5$ 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 $\Omega = \text{AsyVar}[\psi]$. Given our framework, Ω simplifies to the following block diagonal matrix where the m^{th} diagonal block can be approximated using the formula

$$\tilde{\Omega}_m = \frac{1}{n_m} \sum_t \sum_k \psi_{mtk}(\tilde{\theta}) \psi'_{mtk}(\tilde{\theta}) \text{ evaluated at some consistent estimate of } \theta, \tilde{\theta}.$$

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

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.

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.