INTERNATIONAL ECONOMIC REVIEW
Vol. 42, No. 3, August 2001

COMPENSATION POLICIES AND TEACHER DECISIONS*

BY TODD R. STINEBRICKNER[1]

The University of Western Ontario, Canada

A dynamic, discrete choice framework with a flexible structure for unobserved heterogeneity is used to model the occupational decisions of individuals who are certified to teach in elementary and secondary schools. The model is estimated using data from the National Longitudinal Study of the High School Class of 1972 and is used to examine the effects of possible changes in the compensation policy for teachers.

1. INTRODUCTION

Understanding the labor supply decisions of individuals who are certified to teach in elementary and secondary schools is of current interest. One reason for this is that demographic trends in the United States suggest that additional teachers will be needed in the near future to accommodate increases in student enrollments.[2] However, ensuring that schools have an ample supply of teachers is not the only reason that this issue is of interest. If teacher quality is related to academic ability, it is also important to understand why previous research has found that academically gifted teachers are more likely to leave teaching than other teachers. In particular, from a policy standpoint it is important to understand the extent to which this finding can be attributed to the current rigid wage structure for teachers which does not allow wage premiums for academically gifted teachers who may tend to have better non-teaching alternatives than other teachers. An alternative explanation of this finding is that academically gifted teachers simply tend to find teaching less challenging or enjoyable than other teachers.

Previous research, which has primarily involved the use of reduced-form duration models, has also established that a strong, positive relationship exists between teacher pay and the length of time that a person remains in his/her first teaching job.[3]

* Manuscript received November 1998; revised August 1999.

[1] The author thanks Steve Stern, Michael Brien, John Rust, two anonymous referees, and numerous seminar participants.

[2] See "Projections of Education Statistics to 2007," U.S. Department of Education, NCES 97-382, and the associated special report "A Back to School Special Report on the Baby Boom Echo: Here Come the Teenagers."

[3] Previous work in the area of teacher decisions that primarily involves the use of reduced-form models includes Eberts (1987), Murnane and Olsen (1989, 1990), Murnane et al. (1989), Rickman and Parker (1990), Brewer (1996), Mont and Rees (1996), Gritz and Theobald (1996), Theobald and Gritz (1996), Stinebrickner (1998, 1999, forthcoming), and Dolton and van der Klaauw (1995, 1999). van der Klaauw (1996a) and Stinebrickner (2000, 2001) use structural models to examine the issue.

This suggests that wage increases may represent an effective policy instrument from the standpoint of influencing teacher labor supply. The ultimate goal of this article is to specify and estimate a model of individual decision-making which is conducive for simulating the effects that potential wage changes would have on both overall teacher labor supply and the labor supply of different types of teachers. For example, one might be interested in the effects of a wage increase in which all teachers receive an identical pay raise. This type of uniform wage increase conforms to the current, rigid wage structure and is typically supported by teacher unions. However, one might also be interested in analyzing whether alternative policies which deviate from the traditional wage structure might have more desirable effects on either overall teacher labor supply or the labor supply of particular groups. For example, one might be interested in comparing a uniform wage increase to a wage increase which involves the same amount of total spending but allows the amount of the increase that a particular teacher receives to depend on the person's academic ability.

A dynamic, discrete choice, utility maximizing model is used to analyze the effects of the types of wage policies described above. Flexible forms of unobserved heterogeneity are accommodated by specifying person-specific, permanent determinants of wages and nonpecuniary utility in a nonparametric fashion which was shown to be useful in these types of applications by Keane and Wolpin (1997) and is analogous to the method which was proposed by Heckman and Singer (1984). The choice set considered by agents in this article differentiates between individuals who choose to work in nonteaching jobs and individuals who choose to leave the workforce altogether. This distinction implies that previously utilized "teacher-specific" data, which follow a group of teachers only as long as they remain in teaching, are not suitable for the estimation in this article.[4] Instead, this article utilizes data from a general longitudinal survey which follows each individual regardless of whether he/she is currently teaching. At first glance, one might think that the reason that a person chooses not to teach is irrelevant from an educational standpoint because losing a teacher has the same effect on a particular school district regardless of the reason that the teacher chooses not to teach. However, it is important to realize that the reason that a person is not teaching may contain information which is important to understanding the teacher decision process. For example, in Section 3, a simple descriptive look at the data used in this article will suggest that a large proportion of the decrease in the teaching participation rate which occurs over time after certification arises because women leave the workforce altogether. Further, estimates of the model indicate that children play an important role in the decision to leave the workforce. This suggests that child-care subsidies may represent a cost-effective way to increase teacher labor supply. This possibility will be discussed in addition to the types of wage policies mentioned earlier.

The remainder of the article proceeds as follows. Section 2 describes the model. Section 3 describes the data and the results. Section 4 presents policy simulations and conclusions.

[4] Data on North Carolina and Michigan teachers are used by Murnane and Olsen (1989, 1990) and Murnane et al. (1989). Mont and Rees (1996) use information on school districts in New York State. Gritz and Theobald (1996) and Theobald and Gritz (1996) use data from the state of Washington.

2. MODEL AND ESTIMATION

2.1. *Choice Set.* Each individual has a finite decision horizon beginning at the time he/she becomes certified to teach and ending at age T. At each time t, an individual chooses among three mutually exclusive and exhaustive alternatives: work in a teaching job, work in a nonteaching job, or be out of the workforce. Let $d_m(t) = 1$ if alternative m is chosen ($m = 1, 2, 3$) at age t and zero otherwise. There are two things to note. First, because the decision to become certified is not modelled, the empirical analysis will involve a sample of individuals who have chosen to become certified. It is important that the policy discussion in Section 4 takes this into account. Second, the choice set does not include marital and fertility options. The assumptions which are made about family variables are discussed in Section 2.3.

2.2. *Current Period Rewards.* The current period reward in any year t is given by

$$(1) \qquad R(t) = \sum_{m=1}^{3} R_m(t) d_m(t)$$

where $R_m(t)$ is the current period reward associated with the mth alternative at time t. These rewards contain all the benefits and costs associated with each alternative.

$R_m(t)$ is the sum of the wage, $w_m(t)$, and the nonpecuniary utility, $q_m(t)$, that the person receives from option m at time t:

$$(2) \qquad R_m(t) = w_m(t) + q_m(t) \qquad m = 1, 2, 3$$

2.2.1. *Wages.* The individual is assumed to have a wage of zero if she is out of the workforce, $w_3(t) = 0$. The individual's log wage in each of the two work options is assumed to be a linear function of a vector of observable characteristics of the individual at time t, $X(t)$, a person-specific permanent heterogeneity component, γ_m^w, and a random shock, $\nu_m(t)$:

$$(3) \qquad w_m(t) = \alpha_m^w X(t) + \gamma_m^w + \nu_m(t) \qquad m = 1, 2$$

Note that $X(t)$ contains all observable characteristics in the model. Thus, it contains not only variables such as a person's sex but also variables such as a person's accumulated years of teaching and nonteaching experience which depend on the choices which a person makes before time t.

α_m^w is a vector of coefficients which indicates the effect that $X(t)$ has on the average log wage in option m. Given the rigidity of the teaching wage structure in most schools, a person's post-bachelor education level and accumulated teaching experience are expected to be the best predictors of her teaching wages. Nonetheless, the elements of α_1^w associated with other elements of $X(t)$, such as a measure of a person's academic ability, are not constrained to be zero because it is at least conceivable that individuals with more desirable characteristics may, on average, be able to obtain jobs in higher paying school districts. By comparing the importance of this effect to the wage premium that academically gifted individuals receive in nonteaching occupations, it is possible to determine the extent to which the rigidity of the teaching wage structure is responsible for the reality (which will be shown in Section 3.4) that academically gifted individuals are less likely to teach in a particular year than other individuals.

The person-specific, permanent heterogeneity term in the teaching wage equation, γ_1^w, allows for the reality that some individuals will tend to receive more lucrative teaching offers throughout their teaching careers than other individuals. Given that wages have not been adjusted for geographic differences in price levels, this term will to some extent capture the reality that teaching wages will tend to be higher for a person if she teaches in an area where the cost of living is higher.[5] However, given that teaching wages may also vary across school districts within a particular geographic region, γ_1^w will also tend to represent any unobservable individual attributes which make a person more likely to obtain a higher paying job within a particular geographic region. Thus, if schools have a preference for high quality teachers and are able to determine teacher quality during the hiring process, it is reasonable to think that γ_1^w measures a portion of the person's teaching ability that is not captured by observable characteristics such as a person's academic ability.[6] Given these interpretations of γ_1^w, it seems reasonable to make the modelling assumption that each individual knows her specific value of γ_1^w (and all other person-specific heterogeneity terms).[7]

The interpretation of the person-specific heterogeneity term in the nonteaching wage equation, γ_2^w, is similar in spirit to the interpretation of γ_1^w.

2.2.2. *Nonpecuniary utility.* The equations which represents the nonpecuniary utility, $q_m(t)$, associated with each option m are similar in structure to the wage equations:

$$q_m(t) = \alpha_m^q X(t) + \gamma_m^q + \epsilon_m(t) \qquad m = 1, 2, 3 \tag{4}$$

The first term suggests that the nonpecuniary utility derived from a particular option may depend on a person's observable characteristics. For example, an academically gifted individual may find being out of the workforce particularly undesirable. α_3^q is normalized to zero so that α_1^q and α_2^q represent the effect that $X(t)$ has on the nonpecuniary utility derived from teaching and nonteaching jobs relative to the nonpecuniary utility derived from not working.[8]

The descriptive analysis in Section 3 will indicate that the behavior of male and female teachers is quite similar in the early years after certification but becomes

[5] Wages have been adjusted for inflation.

[6] However, if as suggested by Ballou (1996), schools do not necessarily hire the best applicants, this unobserved component could simply represent a person's "connections" or other randomness which determines whether an individual becomes associated with a particular high paying school district.

[7] The assumption that heterogeneity is permanent in nature would not be particularly appealing if teachers tend to change geographic areas or school districts frequently and γ_1^w tends to capture characteristics which are specific to the geographic area or school district rather than characteristics which are specific to the individual. It seems likely that there would not be a large degree of geographic mobility for these teachers. Unfortunately, the extent to which individuals change school districts cannot be examined directly in the data because questions about job changes do not allow one to differentiate between a job change which takes place within a particular school district and a job change which takes place between school districts.

[8] This type of normalization is standard in discrete choice models.

quite different in later years. One intuitively appealing explanation of this finding is that family changes, which are likely to become more prominent after the early years of teaching, tend to influence the nonpecuniary utility which is derived by men and women differently.[9] This can be explored formally because, unlike the teacher-specific data which have been used in the past to study teacher attrition, the data used here contain detailed information about a person's marital and fertility histories. For example, by including a person's marital status at time t in $X(t)$, the model can determine whether changes in marital status influence the nonpecuniary utility which is derived from the various options (and whether effects are different for males and females). Two versions of the model are estimated under different assumptions about the manner in which children influence nonpecuniary utility. In the first specification, it is assumed that a person's children situation at time t can be characterized by the total number of children which are present. This approach will be desirable if the opportunity cost of working varies substantially with the number of children.[10] However, this approach will not be altogether satisfactory if the effect that a particular child has on a person's labor supply varies substantially with the child's age. For example, it might be expected that the nonpecuniary benefits that a person receives from remaining at home would be highest when a child is very young and would diminish significantly as the child approaches school age. To allow for the possibility of this type of effect, the second specification assumes that a person's children situation can be characterized by the age of the person's youngest child. It should be noted that a specification which combines the previous assumptions and characterizes a person's children situation by describing both the number of children and the age of each child would be most desirable. Unfortunately, this type of specification would require a large amount of additional computational time and, given the relatively small sample size described in Section 3, would also potentially lead to difficulties of identification.

The person-specific, permanent heterogeneity terms, γ_m^q, in the nonpecuniary utility equations allow for the possibility that some individuals may have an unobserved "love for teaching" while others may consistently derive more utility from a nonteaching job or from the option of not working. The model is estimated with γ_3^q normalized to zero so that γ_1^q and γ_2^q represent the person-specific, nonpecuniary benefits that a person derives from teaching and nonteaching jobs relative to the nonpecuniary benefits which the person derives from being out of the workforce.

2.2.3. *Summary of $R_m(t)$ and specification of stochastic elements unknown to individual.* To summarize,

$$
\begin{aligned}
(5)\qquad R_m(t) &= w_m(t) + q_m(t) = \alpha_m^w X(t) + \gamma_m^w + \nu_m(t)\\
&\quad + \alpha_m^q X(t) + \gamma_m^q + \epsilon_m(t) \qquad m = 1, 2\\
R_3(t) &= q_3(t) = \epsilon_3(t)
\end{aligned}
$$

Random effects,

[9] Using a competing risks duration models, Stinebrickner (forthcoming) finds a strong correlation between family variables and the length of a person's first spell in teaching.

[10] This could be the case because total day-care costs rise with the number of children or because the nonpecuniary benefits of being at home vary with the number of children.

$\epsilon_m(t)$, $m = 1, 2, 3$, are assumed to be i.i.d. extreme value and serially uncorrelated. This specification has computational benefits which will be discussed in Sections 2.3 and 2.4. The random components of wages are also assumed to be contemporaneously and serially independent with $\nu_m(t) \sim N(0, \sigma_m^2)$, $m = 1, 2$.

2.3. *The Individual's Decision.* In any year t, the individual's objective is to choose the option which maximizes the expected present value of remaining lifetime rewards. The individual cannot determine her exact future utility with certainty because in each future period $t + r$ there exists a set of stochastic components whose realizations are not known at time t. The person-specific heterogeneity terms (the γ's) are assumed to be known to the individual. Thus, from the standpoint of the individual at time t, the set of stochastic variables at time $t + r$ includes the set of wage errors, $\nu(t + r) = \{\nu_1(t + r), \nu_2(t + r)\}$, the set of nonpecuniary errors, $\epsilon(t + r) = \{\epsilon_1(t + r), \epsilon_2(t + r), \epsilon_3(t + r)\}$, and any stochastic elements of $X(t + r)$.

Some elements of X, such as a person's sex and a person's college entrance exam score, are constants which are clearly predetermined at the time the person becomes certified and are known to the individual for all periods. In reality, the remainder of the variables which appear in the two models, teaching and nonteaching experience at time t, the number of children that the person has at time t, the age of the person's youngest child at time t, the person's marital status at time t, and the person's post-bachelor education level at time t, are determined (either deterministically or stochastically) by the choices that a person makes. Unfortunately, computational time constraints and certain modelling difficulties, make it difficult to specify and estimate a model which allows all of these variables to be endogenously determined.[11] Instead, as is evident from the choice set which is specified for the individual in Section 2.1, only years of teaching experience and years of nonteaching experience are assumed to be endogenously determined by choices in the model. One option for including the other variables is to simply treat them as deterministic variables by making the assumption that their values are known by the individual for all future periods. Clearly removing all future uncertainty associated with these variables is less than ideal. Thus, although, the assumption is made that a person knows her future marital status values and post-bachelor education levels, this assumption is relaxed for a person's future child outcomes which were shown to be very important determinants of teacher decisions in Stinebrickner (forthcoming). In particular, it is assumed

[11] As is well known, the addition of choices or state variables can lead to substantial increases in computational time. Allowing the post-bachelor education levels for this group to be determined endogenously would create some difficult modeling issues. In some states, teachers are required to obtain a master's degree within a relatively small number of years after becoming certified. For teachers in these states, the accumulation of post-bachelor education levels essentially follows a deterministic path given the decision to remain in teaching. That is, for these individuals, obtaining post-bachelor education could be considered part of the job requirement. However, in other school districts, where these requirements do not exist, or in cases in which individuals are obtaining post-bachelor education for the purpose of working in a nonteaching occupation, whether to obtain post-bachelor education does require an additional decision on the individual's part. One option would be to ignore information about the post-bachelor education level. However, given an interest in separating the effect of education on wages from the effect of experience on wages, it was decided to include it in a predetermined fashion.

Θ (xper. B_t $B' = 1$ if

Counter — state contr

$\Phi(\ )$

that the birth of a child in year t is a stochastic event whose probability is known by the individual. Letting $B(t+1) = 1$ if a new child is present in year $t+1$, and $B(t+1) = 0$ otherwise (i.e., $B(t+1) = 1$ if a birth takes place sometime in year t),

$$B(t+1) = 1 \qquad \text{iff } B^*(t+1) = \alpha^B X(t) + \gamma^B + \eta(t+1) > 0 \tag{6}$$

The first term allows a person's probability of having a birth to depend on observable characteristics $X(t)$. For example, the first term allows for the possibility that women, who tend to get married at younger ages than men, are likely to have children at younger ages. Note that the observable characteristics which are assumed to influence births, the observable characteristics which are assumed to influence wages, and the observable characteristics which are assumed to influence nonpecuniary utility are not identical (i.e., as will be discussed in more detail in Section 3.2, some elements of α_1^w, α_2^w, α_1^q, α_2^q, and α^B are constrained to be zero during estimation).

The second term in Equation (6) is a person-specific heterogeneity term which allows for the possibility that some individuals may have a "love for children" which makes them more likely to have children in each period.[12] The final term, $\eta(t+1)$, captures random determinants of births. Assuming that $\eta(t+1) \sim N(0,1)$ implies a probit form for the birth probability associated with Equation (6).

While this approach to dealing with births is less ideal than explicitly modelling the decision to have a birth (see, e.g., Wolpin, 1984), it does allow the model to deal with some of the problems which would potentially arise if a person's fertility history was treated as predetermined. For example, if individuals who are unhappy being employed (because of unhappiness with wages or the nonpecuniary aspects of jobs) are more likely to have children, treating children as predetermined would potentially lead to an overstatement of the effect that children have on labor supply. Intuitively, this would occur under this scenario because the model could not take into account that, on average, individuals with more children would be less likely to work (than individuals with less children) even if they had the same number of children. In this model, allowing the person-specific unobserved propensity for children, γ^B, to be correlated with the person-specific wage and nonpecuniary heterogeneity terms associated with the working options represents one way to attempt to address this problem.[13]

Define $\gamma = \{\gamma_1^w, \gamma_2^w, \gamma_1^q, \gamma_2^q, \gamma^B)$. The state space at time t, which represents the information which is relevant for a person's decision, is $S(t) = \{\gamma, \nu(t), \epsilon(t), X(t)\}$. Depending on the particular model specification, $X(t)$ contains either the person's total number of children or the age of the person's youngest child. These variables summarize the information in $\eta_{(t)}, \eta_{(t-1)}, \eta_{(t-2)}, \ldots$ which is relevant for the person's situation at time t. The set of time $t+1$ stochastic components whose realizations are not known when the person makes her time t decision is given by $\zeta(t+1) = \{\nu(t+1), \epsilon(t+1), \eta(t+1)\}$.

[12] The assumption that individuals know their person-specific values of unobserved heterogeneity implies that some individuals realize that they are more likely to have children than other individuals.

[13] This would not address a scenario in which the decision to have a child depends on yearly random shocks to wages or nonpecuniary utility.

The expected present value of lifetime rewards associated with option m can be represented by the Bellman equation (Bellman, 1957),

$$(7)\quad V_m(S(t),t) = R_m(S(t),t) + \mathrm{BE}[V(S(t),t) \mid S(t), d_m(t) = 1] \qquad m = 1, 2, 3$$

where $V(S(t),t) = \max[V_1(S(t+1),t+1), V_2(S(t+1),t+1), V_3(S(t+1),t+1)]$.

The value functions in (7) do not have closed form solutions. However, numerical solution of (7) can be carried out by backward recursion. For each value of $S(t)$, the computation of $E[V(S(t),t) \mid S(t), d_m(t) = 1]$ for a particular option m involves the evaluation of a multidimensional integral over the distribution of $\zeta(t+1)$. Taking into account that the person uses the distribution of $\eta(t+1)$ to compute the probability that a new child will be born before the next period, the expected maximum can be written as[14]

$$\begin{aligned}(8)\qquad & E[V(S(t),t) \mid S(t), d_m(t) = 1] \\ &= \mathrm{PR}(B(t+1) = 1) \\ &\quad \cdot \iint [V(S(t),t) \mid S(t), d_m(t) = 1, B(t+1) = 1, \nu(t+1), \epsilon(t+1)] \\ &\quad \cdot dG(\epsilon(t+1))dF(\nu(t+1)) + \mathrm{PR}(B(t+1) = 0) \\ &\quad \cdot \iint [V(S(t),t) \mid S(t), d_m(t) = 1, B(t+1) = 0, \nu(t+1), \epsilon(t+1)] \\ &\quad \cdot dG(\epsilon(t+1))dF(\nu(t+1))\end{aligned}$$

where G and F are the distribution functions of ϵ and ν, respectively.

The assumption that the ϵs are i.i.d. extreme values is computationally desirable because it implies that the inner (multidimensional) integral in each of the two terms in Equation (8) has a closed-form solution conditional on the value of $\nu(t+1)$.[15] The outer integral is approximated using Gaussian quadrature integral approximation methods. Stinebrickner (2000) discusses the operational details of the integral approximation methods in more detail.

2.4. *Estimation*. The model is estimated by maximum likelihood. The likelihood contribution for person i at time t is the joint probability of the choice that the person makes, the birth outcome at time t, and the person's wage if the person works and a wage is observed. Conditional on knowing the permanent heterogeneity values γ, the likelihood contribution for person i at time t, $L(i, t \mid \gamma)$, is straightforward to

[14] More specifically, under the first specification, the individual is interested in the probability that she will have an additional child in the next period; i.e., number-of-children $(t+1)$ =number-of-children(t) if $B(t+1) = 0$ and number-of-children $(t+1)$ =number-of-children$(t) + 1$ if $B(t+1) = 1$. Under the second specification, the individual is interested in the probabilities associated with each possible age that her youngest child could be at time $t+1$; i.e., age-of-youngest$(t+1)$ =age-of-youngest$(t) + 1$ if $B(t+1) = 0$ and $0 <$ age-of-youngest$(t+1) < 1$ if $B(t+1) = 1$.

[15] See, e.g., Rust (1987) or Berkovec and Stern (1991).

calculate. The computation of the birth probability and the wage probability (if a wage is observed) are straightforward given the wage Equation (3) and the probit birth probability from Equation (6). The computation of the choice probability at time t conditional on γ is also reasonably straightforward once the value functions $V_m(S^*(t), t)$ have been solved, where $S^*(t)$ is the person's observed state at time t excluding the values of $\epsilon(t)$. This is the case because the extreme value assumption for $\epsilon(t)$ implies that, conditional on $\nu(t)$ and γ, the choice probabilities take on a multinomial-logit, closed-form solution. In practice, at least one of the components of $\nu(t)$ will not be observed. Thus, in order to obtain the choice probability conditional on only γ, it is necessary to integrate the closed-form, multinomial-logit solution over the distribution of the values of $\nu(t)$ which are not observed.

Given the assumption that the stochastic components in the model are serially uncorrelated, the likelihood contribution for the person conditional on γ is the product of the year-specific conditional likelihood contributions over all of the years for which a person is observed,

$$L(i|\gamma) = \prod_t L(i, t|\gamma) \tag{9}$$

Ran Eff (4)

The person-specific heterogeneity values, although assumed to be known by the individual, are not observed by the econometrician. The method to deal with this follows the nonparametric approach taken by Keane and Wolpin (1997). In a fashion analogous to the method introduced by Heckman and Singer (1984) in the context of duration models, it is assumed that there are a discrete number of types of individuals, each with a different value γ_k of $\gamma = \{\gamma_1^w, \gamma_2^w, \gamma_1^q, \gamma_2^q, \gamma^B\}$.[16] The proportion of individuals of each type k is given by π_k. Here it is assumed that there are four types. In this case, the unconditional likelihood contribution of person i is given by

$$L(i) = \sum_{k=1}^{4} \pi_k L(i|\gamma_k) \tag{10}$$

4

and the likelihood contribution for the sample is given by

$$L = \prod_{i=1}^{I} L(i) \tag{11}$$

where I is the total number of individuals in the sample.

[16] It would be desirable to estimate types separately for different combinations of observable characteristics that are present at the time individuals become certified. This is especially true for the individuals in the sample that have children at the time they become certified. However, given the additional estimation difficulties that would potentially be created (e.g., increases in the number of nonparametric heterogeneity parameters being estimated and potential difficulties in practice with identification) and given that only a small proportion of individuals in the sample fit this description, it was decided to (somewhat incorrectly) estimate one set of type probabilities for the sample as a whole.

4 x 551 ≈ 2200
"30" = 120

3. DATA AND RESULTS

3.1. *Data*. Data from the National Longitudinal Study of the High School Class of 1972 (NLS-72) are used to estimate the model. The first wave of this survey, which was completed in 1972, includes interviews with 22,652 students who were expected to graduate from high school in that year. Included in the first wave is information on aptitude tests such as the Scholastic Aptitude Test (SAT). Follow up surveys were taken in 1973, 1974, 1976, 1979, and 1986. Thus, for each person, the survey contains detailed information about work experience, education, marriage, and fertility for approximately 14 years after the person graduated from high school.[17]

Further, 832 individuals, who were certified to teach in elementary or secondary schools, were sent supplemental questionnaires which asked questions about their teaching experiences. The final sample used in this paper consists of 551 of these individuals who became certified to teach at some point between 1975 and 1985.[18] The majority of the people who were sent the teaching supplement but do not appear in the final sample had missing observable characteristics. The others had crucial missing information which made the construction of job or personal histories impossible. Since this article involves the career choices of individuals only after they become certified, which usually requires a minimum of 4 years of training, the data contain between 1 and 11 years of work histories and personal information for each person. On average, 8.9 years of data are observed for a person in the sample.[19]

3.2. *Model Specification*. The variables which are included in $X(t)$ are summarized in Table 1. There are several things to note about the model specification. First, as mentioned earlier, some elements of α_1^w, α_2^w, α_1^q, α_2^q, and α^B are constrained to be zero so that the sets of variables which enter the various equations are not identical. For example, it is assumed that the family variables do not affect the person's wages and the average nonpecuniary benefits from teaching do not depend on the year. Second, although it would be desirable to estimate the model separately for males and females, this is not done because of the relatively small sample. However, because males and females are likely to react differently to family changes which take place in the years after certification, the specifications include interaction terms which allow marital and children variables to have different effects for males and females. Finally, although in the second specification (in which a person's children situation is characterized by the age of the person's youngest child) it would be desirable to estimate a separate effect for each possible age that a person's youngest child could have, identification of this specification turned out to be difficult in practice given the

[17] Since survey waves did not occur in every year, some of the survey waves ask the individual retrospective questions which cover several years of the individual's life. One consequence of this is that the estimation algorithm must take into account that the individual is not asked about wages for every year in which she works.

[18] It should be noted that because teacher certification is used as a selection criteria for the sample, the estimates in the model do not represent the preferences of the entire population of 1972 high school graduates or even the subset of 1972 high school graduates who also graduated from college.

[19] The variance of the number of observed years is approximately 4. A small number of people finished college in 3 years. For these people 11 years of data are observed. For all other people 10 or less years are observed.

TABLE 1
DESCRIPTIVE STATISTICS FOR VARIABLES IN $X(t)$

Variable	Mean	Standard Deviation
Math SAT	474	91.7
Percent female	72.3	
Number of children (in first year of certification)	0.19	0.5
Number of children (in 1986)	1.1	1.1
Percent with at least one child (in first year of certification)	12.7	
Percent with at least one child (in 1986)	56.1	
Percent married (in first year of certification)	37.3	
Percent married (in 1986)	77.4	
Percent married in at least one period	81.4	
Number of years of post-bachelor education (as of 1986)	1.2	1.1
Years of teaching experience (as of 1986)	4	3.2
Years of nonteaching experience (as of 1986)	2.3	2.9

sample size. As an alternative, the age of a person's youngest child is described by two indicator variables: whether the child is younger than two years of age (young-child) and whether the child is older than two years of age (older-child).[20] The effects that these variables have on a person's nonpecuniary utility are identified relative to the case of a person with no children. In the first specification (in which a person's children situation is characterized by the total number of children that the person has), the number of children variable (number-of-children) is assumed to enter the model in a linear fashion.

The set of variables which are allowed to potentially influence the pecuniary and nonpecuniary utility equations and the birth equation can be seen clearly in Table 2 which presents the model estimates under the two different specifications/assumptions about children. The model is estimated assuming a discount rate $\beta = 0.95$.[21]

The next several sections examine how well the model fits the data for the overall sample, individuals with "high" academic ability, individuals with "low" academic ability, male teachers, and female teachers. The estimates of the model in Table 2 will be discussed in the context of what insight they provide about the differences which are observed between certain groups in the data.

3.3. *A Look at the Overall Sample.* For each year that each person is observed, the data reveal whether the person chooses a teaching job, a nonteaching job, or not to work. The first entry in Figure 1 (labelled "data") shows that, of the aggregated 4874 person years of data, 0.50 of the years are spent teaching, 0.30 of the

[20] The estimates of the model were quite insensitive to changes in the age which determines whether a particular child is a "young" child or an "older" child.

[21] It is assumed that the end of an individual's decision-making horizon ends at the age of 45 (i.e., $T = 45$). It is then assumed that individuals remain at their age 45 job for the remainder of their working lives. Because individuals discount future utility, estimates were found to change very little when the model was estimated with different values of T.

$T = 65 - 22 = 43$

TABLE 2

STRUCTURAL MODEL ESTIMATES

Variable $X(t)$	Specification 1	Specification 2
Teaching Wage	α_1^w	α_1^w
Dummy for male	−0.015	−0.024
	(0.021)	(0.021)
Math SAT/100	−0.004	−0.005
	(0.010)	(0.011)
Years of post-bachelor education	0.040*	0.037*
	(0.010)	(0.010)
Years of teaching experience	0.051*	0.053*
	(0.006)	(0.006)
Years of nonteaching experience	−0.008	−0.007
	(0.013)	(0.014)
Year (1975=1, 1976=2,...)	−0.101*	−0.103*
	(0.017)	(0.018)
Year*year	0.006*	0.003
	(0.001)	(0.002)
Constant	4.917*	4.962*
	(0.071)	(0.071)
γ_1^w Type 1	•	•
γ_1^w Type 2	0.845*	0.796*
	(0.048)	(0.047)
γ_1^w Type 3	0.329*	0.311*
	(0.031)	(0.029)
γ_1^w Type 4	0.398*	0.366*
	(0.027)	(0.025)
σ_1	0.319*	0.320*
	(0.004)	(0.004)
Nonteaching Wage	α_2^w	α_2^w
Dummy for male	0.070*	0.069*
	(0.031)	(0.032)
Math SAT/100	0.024*	0.025*
	(0.013)	(0.013)
Years of post-bachelor education	0.046*	0.045*
	(0.009)	(0.009)
Years of teaching experience	−0.037*	−0.035*
	(0.010)	(0.010)
Years of nonteaching experience	0.027*	0.021*
	(0.010)	(0.009)
Year (1975=1,1976=2,...)	0.004	0.011
	(0.026)	(0.025)
Year*year	0.004*	0.003
	(0.002)	(0.002)

(continued)

TABLE 2
CONTINUED

Variable $X(t)$	Specification 1	Specification 2
Constant	4.141*	4.117*
	(0.109)	(0.107)
γ_2^w Type 1	•	•
γ_2^w Type 2	0.968*	0.968*
	(0.056)	(0.054)
γ_2^w Type 3	0.330*	0.377*
	(0.052)	(0.052)
γ_2^w Type 4	0.538*	0.530*
	(0.040)	(0.038)
σ_2	0.383*	0.377*
	(0.007)	(0.007)
Teaching Nonpecuniary Utility	α_1^q	α_1^q
Dummy for male	−0.093	−0.120
	(0.095)	(0.098)
Math SAT	−0.075*	−0.037
	(0.031)	(0.031)
Years of teaching experience	−0.071	−0.072*
	(0.011)	(0.010)
Years of nonteaching experience	−0.083*	−0.092*
	(0.019)	(0.020)
Number-of-children	−0.488*	•
	(0.056)	
Number-of-children × male	0.421*	•
	(0.119)	
Young-child (less than two years of age)	•	−1.380*
		(0.161)
Young-child × male	•	1.366*
		(0.273)
Older-child	•	−1.020*
		(0.141)
Older-child × male	•	0.665*
		(0.292)
Dummy for marriage	−0.413*	−0.286*
	(0.082)	(0.082)
Marriage × male	0.301*	0.200
	(0.080)	(0.122)
Constant	−2.365*	−2.658*
	(0.292)	(0.272)
γ_1^q Type 1	•	•
γ_1^q Type 2	0.383	0.298
	(0.258)	(0.244)

(continued)

TABLE 2
CONTINUED

Variable $X(t)$	Specification 1	Specification 2
γ_1^q Type 3	−1.720*	−1.600*
	(0.164)	(0.151)
γ_1^q Type 4	0.089	0.321*
	(0.106)	(0.121)
τ	0.554*	0.573*
	(0.065)	(0.068)
Nonteaching Nonpecuniary Utility	α_2^q	α_2^q
Dummy for male	−0.204*	−0.232*
	(0.099)	(0.010)
Math SAT	−0.080*	−0.045
	(0.030)	(0.031)
Years of teaching experience	−0.062*	−0.066*
	(0.014)	(0.014)
Years of nonteaching experience	0.019	0.012
	(0.013)	(0.013)
Number of children	−0.473*	•
	(0.056)	
Number children × male	0.524*	•
	(0.124)	
Young-child (less than two years of age)	•	−1.434*
		(0.169)
Young-child × male	•	1.445*
		(0.285)
Older-child	•	−0.907*
		(0.131)
Older-child × male	•	0.600*
		(0.280)
Dummy for marriage	−0.505*	−0.448*
	(0.090)	(0.094)
Marriage × male	0.281*	0.352*
	(0.084)	(0.127)
Constant	−2.971*	−3.159*
	(0.222)	(0.217)
γ_2^q Type 1	•	•
γ_2^q Type 2	0.285	0.186
	(0.264)	(0.247)
γ_2^q Type 3	−1.796*	−1.748*
	(0.184)	(0.172)
γ_2^q Type 4	−0.073	0.142
	(0.105)	(0.121)
τ	0.554*	0.573*
	(0.065)	(0.068)

(continued)

TABLE 2
CONTINUED

Variable $X(t)$	Specification 1	Specification 2
Birth Equation	α^B	α^B
Dummy for male	0.011	−0.346
	(0.567)	(0.681)
SAT/100	0.028	0.046
	(0.032)	(0.036)
Years of teaching experience	0.029	0.018
	(0.022)	(0.022)
Years of nonteaching experience	0.016	−0.013
	(0.023)	(0.024)
Years of post-bachelor education	0.018	0.015
	(0.045)	(0.049)
Year (age) − 1975 = 1, 1976 = 2, ...	0.207*	0.175
	(0.076)	(0.091)
Year*year (age*age)	−0.014*	−0.013
	(0.006)	(0.007)
Male × year	−0.027	0.096
	(0.174)	(0.217)
Male × year*year	0.004	−0.005
	(0.013)	(0.016)
Number-of-children	−0.124*	•
	(0.046)	
Young-child	•	0.404*
		(0.105)
Older-child	•	−0.178*
		(0.081)
Dummy for marriage	1.650*	1.628*
	(0.137)	(0.142)
Constant	−3.565*	−3.436*
	(0.314)	(0.365)
γ^B Type 1	•	•
γ^B Type 2	0.277	0.209
	(0.188)	(0.189)
γ^B Type 3	0.356*	0.230
	(0.124)	(0.137)
γ^B Type 4	0.125	0.077
	(0.117)	(0.125)
Log Likelihood Function	−5642.47	−5554.76

The numbers are estimates from a specification in which the discount factor, β, is set to 0.95. The numbers in parentheses are asymptotic standard errors. *denotes an asymptotic t ratio greater than two. Specification 1 is the model which includes a person's number of children. Specification two is the model which includes the age of a person's youngest child.

years are spent working in nonteaching jobs, and the other 0.20 of the years are spent not working. The second and third entries in Figure 1 (labelled "Model 1" and "Model 2") show the simulated choice proportions for the entire sample under the two model specifications/assumption about children. Both models appear to fit the overall choices of the individuals in the sample well. This is formally confirmed

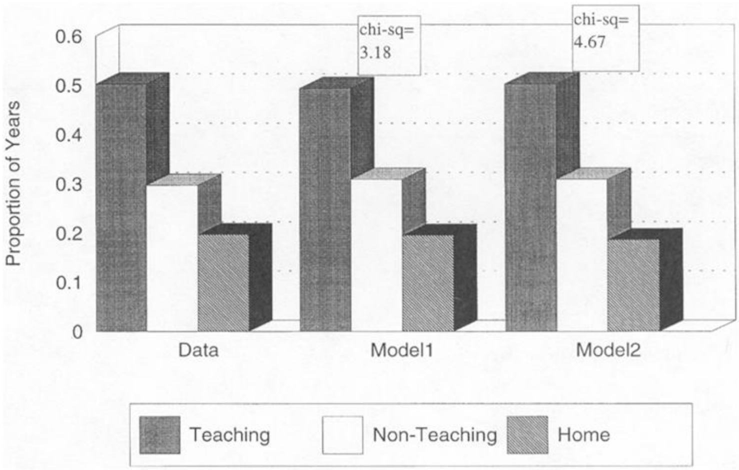

FIGURE 1

PROPORTION AGGREGATE YEARS EACH OPTION, DATA AND SIMULATED MODEL PREDICTIONS, FULL SAMPLE

by chi-square goodness of fit test statistics (shown above the Model 1 and Model 2 entries) which, at a 0.05 level of significance, do not lead to a rejection of the null hypothesis that the choice proportions generated by the model are the same as the choice proportions observed in the data.[22]

The proportions in Figure 1 do not reveal much information about the trends of particular individuals. For example, the aggregate teaching proportion of 0.50 is consistent with every teacher choosing to teach in half of the years. However, it is also consistent with 50 percent of all teachers never choosing to teach and the other 50 percent choosing to teach in every year. In the data, 0.79 of individuals choose to teach in at least one year and the probability that a person leaves her first teaching job before the start of the fifth year is approximately 0.48.[23] It was found that the models tend to somewhat overstate the participation rate and understate the length of individuals' first teaching spell; the simulations indicate that both models predict that 0.86 of all individuals will teach in at least one year and that 0.58 of all individuals will leave their first teaching spell before the start of the fifth year.

3.4. *Differences by Academic Ability.* Using individual-level data from Texas, Rivkin et al. (1998) provide compelling evidence that teacher quality is a very important determinant of student learning.[24] However, while this research highlights the importance of identifying high quality teachers, there remains a large degree

[22] 22. The chi-square critical value for the test at a 0.05 level of significance is $\chi_2^2 = 5.99$.

[23] Some care must be taken in interpreting the former number (0.79) because some individuals became certified near the end of the sample period. The latter number (0.48) is derived using a Kaplan–Meier survivor function which takes into account the presence of right-censored observations.

[24] They find that variations in teacher quality account for at least 7.5 percent of the total variation in student achievement across schools. The effects of teachers are found to be much larger than the effects of overall school organization, leadership, or financial conditions.

of uncertainty about the extent to which this can be achieved solely on the basis of observable characteristics. It is quite clear that some observable characteristics of teachers are not useful from the standpoint of identifying high quality teachers. For example, Hanushek's (1986) examination of 106 previous studies provides very strong evidence that teachers with master's degrees are no more effective than teachers with only bachelor degrees.[25] However, for other observable characteristics, such as scores on standardized academic tests, previous research does not always provide as conclusive a picture. For example, an examination of the predictive ability of scores on the National Teacher Examination (NTE) led Haney et al. (1987) to conclude that, "the available evidence is none too good, but it indicates that teacher tests have little, if any, power to predict how well people perform as teachers...." However, a more recent overview by Bishop (1996) reaches a more positive conclusion about the potential usefulness of standardized test scores: "The teacher characteristic that most consistently predicts student learning are tests assessing the teacher's general academic ability and subject knowledge."[26] Certainly, it is intuitively appealing to believe that a solid foundation in the types of reading and math skills that are tested on general, standardized tests will to some degree contribute to a person's teaching performance. However, it is clear from the previous literature that, due to a variety of methodological and measurement difficulties that past researchers have faced, the exact strength of the relationship between academic ability (as measured by standardized test scores) and teaching ability is far from being well established. Thus, although the remainder of this section will indicate that significant behavioral differences exist between individuals of different academic ability (as measured by test scores), the policy conclusions in Section 4 will take into account that the benefits of attempting to address these differences depend on the true strength of the relationship between teachers' academic ability as measured by tests scores and student learning. In this respect, more research on the nature of this relationship is important from a policy standpoint.

The measure of ability that is used in this analysis is the individual's score on the math SAT score.[27] The first three entries in Figure 2 show the proportion of aggregate person years spent in teaching jobs, nonteaching jobs, and the nonwork

[25] Rivkin et al. (1998) find evidence to support these conclusions.

[26] This conclusion appears to be drawn largely on the basis of the work by Hanushek (1971), Strauss and Sawyer (1986), Ferguson (1990), Ehrenberg and Brewer (1993), and Monk (1994).

[27] The behavior of individuals with higher verbal SAT scores is certainly also of interest. However, it was decided to concentrate on math SAT scores because it is likely that retaining and recruiting teachers with high math SAT scores will be an important priority given that students in the United States are more deficient in math and science than reading (Bishop, 1996). Retaining those high math SAT teachers who do enter teaching may be especially important because individuals who enter teaching score further below national averages on the math SAT than they do on the verbal SAT (see, e.g., Bishop, 1996). In general, the retention of teachers with high math SAT scores may be more difficult than the retention of teachers with high verbal SAT scores because the math SAT score is a strong predictor of opportunity costs. For example, Murnane et al. (1995) find that males graduating from high school in 1972 with strong basic math skills have significantly higher hourly earnings at age 24 than males graduating in that year with average math skills. Using the data of this article, simple wage regressions indicate that teachers with high math SAT scores do receive higher nonteaching wages than other teachers. However, this nonteaching wage premium is not present

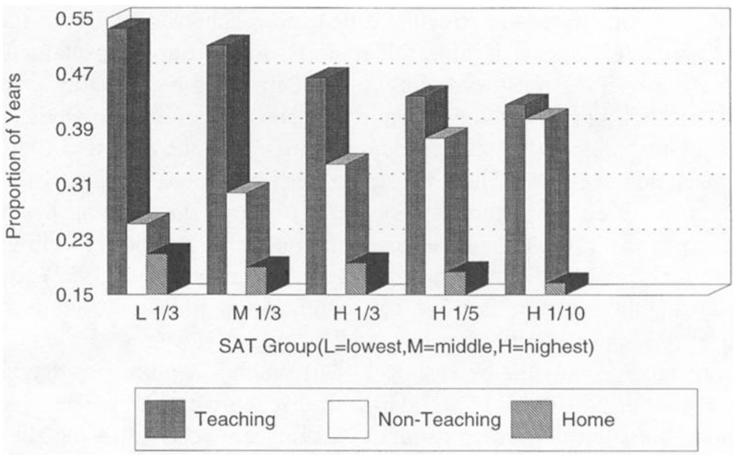

FIGURE 2

PROPORTION AGGREGATE YEARS EACH OPTION, DATA BY SAT GROUP

(home) option for three mutually exclusive groups of teachers: teachers in the lowest third of the sample in terms of SAT scores, teachers in the middle third, and teachers in the top third. Two additional groups, teachers in the highest fifth and teachers in the highest tenth, are added to give more detail about teachers at the top of the ability distribution. Notice that the proportion of years spent in teaching decreases as the academic ability of a group increases. This is accompanied by an increase of similar size in the proportion of years spent in nonteaching jobs. Thus, Figure 2 provides evidence that academically gifted teachers do choose teaching jobs less frequently and nonteaching jobs more frequently than other teachers under the current rigid wage structure.[28]

To simplify discussion, the remainder of the article will concentrate on differences between individuals in the top third of the SAT distribution and individuals in the bottom two-thirds of the SAT distribution. Figure 3 indicates that the models are able to capture the differences which exist between the behavior of "high" ability teachers and the behavior of "low" ability teachers. From a policy standpoint, it is of importance to understand why the model predicts differences between the groups. Both specifications in Table 2 show that academically gifted individuals receive a statistically significant wage premium in nonteaching jobs but not in teaching jobs.[29] Further, the estimates reveal no evidence that academically gifted individuals find

for individuals with high verbal SAT scores. Using combined math and verbal scores would tend to obscure some of the problems/solutions associated with retaining high math SAT teachers.

[28] Similar differences exist when Figure 2 is constructed separately for females. For example, the proportion of time spent teaching for the five SAT groups (from lowest to highest) is 0.55, 0.49, 0.45, 0.42, and 0.41. For males, the numbers are 0.53, 0.55, 0.50, 0.48, and 0.45.

[29] The coefficient on SAT is approximately 0.024 in the nonteaching wage equations and is approximately −0.004 in the teaching wage equations. Recall that wages are specified in logs. Thus, each additional 100 points on the SAT test leads to a 2 percent increase in a person's wage. That this type of wage difference can create the type of labor supply differences between SAT groups which are

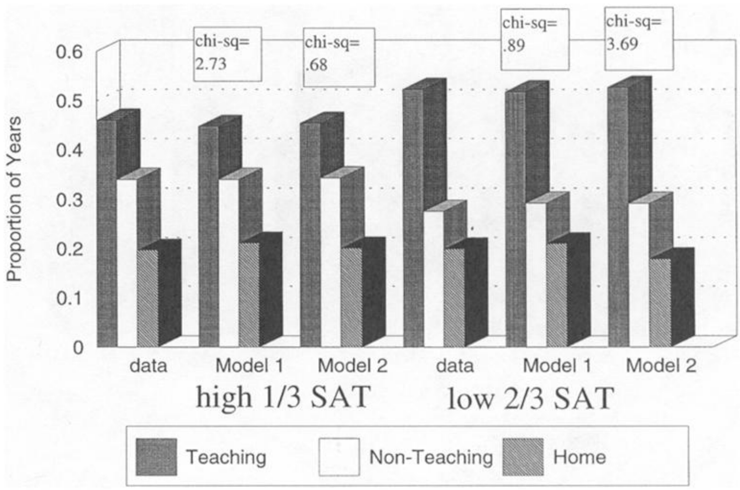

FIGURE 3

PROPORTION AGGREGATE YEARS EACH OPTION, DATA AND SIMULATED MODEL PREDICTIONS, HIGH AND LOW SAT

nonteaching jobs more enjoyable than teaching jobs (i.e., the effect of SAT is similar in the nonpecuniary teaching equation and the nonpecuniary, nonteaching equation). Thus, the models suggest that the rigidity of the teaching wage structure is largely responsible for the reality that academically gifted individuals are less likely to choose teaching jobs and more likely to choose nonteaching jobs than other teachers.

3.5. *Differences by Sex.* The first and fourth entries in Figure 4 (labelled "data") indicate that the choices of males and females differ substantially in the data. For females in the data, the proportion of aggregate person years spent in teaching jobs, nonteaching jobs, and out of the workforce is 0.49, 0.27, and 0.24, respectively. For males, the proportions are 0.52, 0.38, and 0.10. The remaining entries in Figure 4 show that the models correctly predict that, when compared to men, women spend a higher proportion of time out of the workforce and a lower proportion of time in nonteaching jobs. However, because the models tend to somewhat understate the magnitude of the differences between males and females, the chi-square test statistic values in Figure 4 indicate that the two models do not pass the goodness of fit tests at reasonable levels of significance.

The estimates in Table 2 reveal that, although males and females receive similar wages in teaching jobs, males receive a wage premium of approximately 7 percent in the nonteaching sector. Thus, the rigidity of the teaching wage structure provides an explanation for the finding that men are more likely to choose nonteaching jobs than women. However, this wage differential is unlikely to explain the large difference in the proportion of men and women who are out of the workforce altogether. The

observed in Figure 3 is consistent with the sensitivity of labor supply to wages that is found in the policy simulations in Section 4.

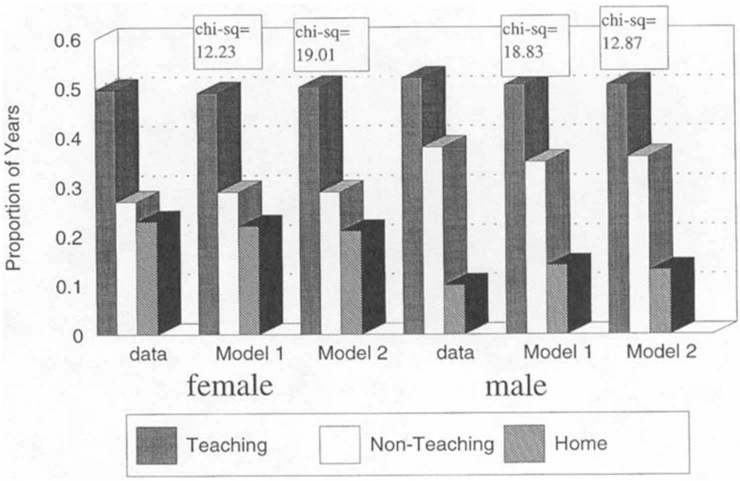

FIGURE 4

PROPORTION AGGREGATE YEARS EACH OPTION, DATA AND SIMULATED MODEL PREDICTIONS, MALES AND FEMALES

fact that the male coefficient is not positive in either of the nonpecuniary utility equations (this is true for both Model 1 and Model 2) suggests that this difference is also not driven by permanent preference differences between the sexes. Instead, Table 2 suggests that previously unavailable marital and fertility information is very important from the standpoint of explaining male/female differences. The effect of the number-of-children variable in the first specification, which represents the effect of children for females, indicates that a woman with a particular number of children is significantly less likely to be working in either a teaching job or nonteaching job than a woman with less children (i.e., the presence of children implies negative nonpecuniary utility from either type of job relative to the alternative of not working). This is consistent with the second specification which shows that the presence of a young-child leads a person to enjoy working significantly less relative to the alternative of being at out of the workforce. Further, the fact that the estimated effect of an older-child is also large (although somewhat smaller than the estimated effect of a young-child) implies that the birth of a child has an effect on a woman's labor supply which lasts beyond two years.[30] Unfortunately, because the sample ends at the time that individuals are approximately 31 years old, only a very small number of decisions are observed for women whose youngest child is old enough to attend school. Thus, it is not possible to determine the extent to which a child influences a person's labor supply when the child becomes old enough to enter school.[31]

[30] This is consistent with the simple descriptive analysis in Stinebrickner (1996b) which showed that most women in this sample who leave the workforce altogether do not return to teaching within the first several years.

[31] Although it is not possible to estimate the effect that a school-age child will have on labor supply, this information is necessary to solve the value functions of forward-looking individuals who take into account that their young children will eventually become older. However, it was found

The number-of-children × male variable in the first specification is an interaction term which represents the additional effect of children for males. Not surprisingly, the effect of this interaction variable is large and positive in both the teaching and nonteaching nonpecuniary utility equations. Therefore, the total effect of children for males (number-of-children + number-of-children × male) is small and insignificant in both the teaching and nonteaching nonpecuniary equations. This is also true for the second specification because, for both nonpecuniary equations, the effect of the young-child × male and older-child × male variables are of similar size but of different sign than the young-child and older-child variables in both nonpecuniary equations. Similarly, being married makes a woman much less likely to work but has little effect on males.[32]

3.6. *Differences Over Time After Certification.* If teacher attrition is indeed important, the proportion of individuals in the sample who choose to teach may not be constant over time. Figure 5 shows that this is true in the data for females and figure 6 shows that this is true in the data for males. In the second year after certification, approximately 0.60 of all females and 0.65 of all males are teaching. By the ninth year after certification, both of these proportions have fallen to 0.43. However, while the decreases in teaching proportions are similar for males and females, the "reasons" for these decreases are quite different. Figure 5 shows that for females, the decrease in the teaching proportion is accompanied by an increase of almost identical size in the proportion of women who are out of the workforce. Thus, the proportion of women who are in nonteaching occupations (which is not included in Figure 5 in order to preserve readability but can be easily calculated because the choice proportions must sum to one) changes very little between the second year after certification and the ninth year after certification. On the other hand, from Figure 6 it can be seen that the proportion of men who are out of the workforce changes very little over time after certification. Instead, the decrease in the teaching proportion over time is accompanied by an increase of similar size in the proportion of the sample who choose to work in nonteaching occupations (again the nonteaching proportions are omitted to maintain readability).

The estimated effects of the marital and fertility variables suggest that the models can potentially explain these findings. As shown in Table 1, the majority of individuals in the sample are not married and do not have children when they initially become certified to teach. Thus, the models predict very little difference in the behavior of men and women at the early stages of their careers. However, as families are created or enlarged in the years after certification, the models predict that women become

that model estimates and policy simulations were quite insensitive to two very different assumptions about the influence of school-age children. Under the first assumption, school-age children were assumed to have the same effect on nonpecuniary utility as children between the ages of two and five years. Under the second assumption, children of school age were assumed to have no effect on nonpecuniary utility (i.e., the same effect as no children). The estimates in the article are obtained using the first assumption.

[32] As discussed earlier, the model (incorrectly) assumes that marital status is exogenous and predetermined. van der Klaauw (1996b) estimates a female labor supply model with endogenously determined marital status.

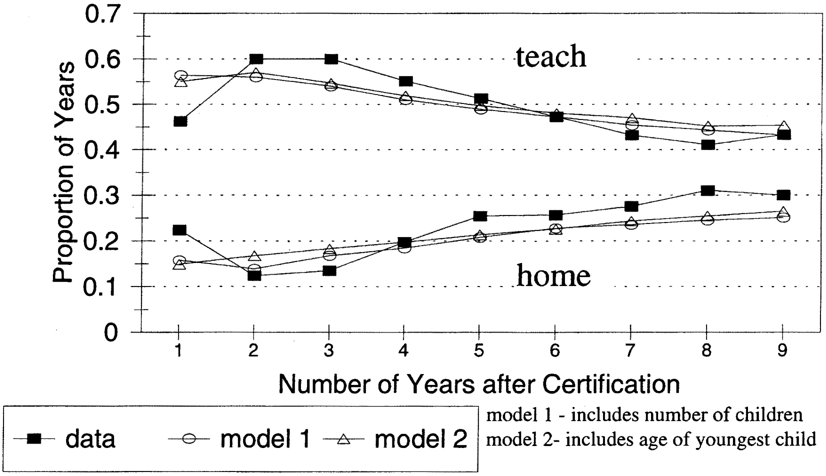

FIGURE 5

PROPORTION AGGREGATE YEARS EACH OPTION, DATA AND SIMULATED MODELS OVER TIME AFTER CERTIFICATION—FEMALES

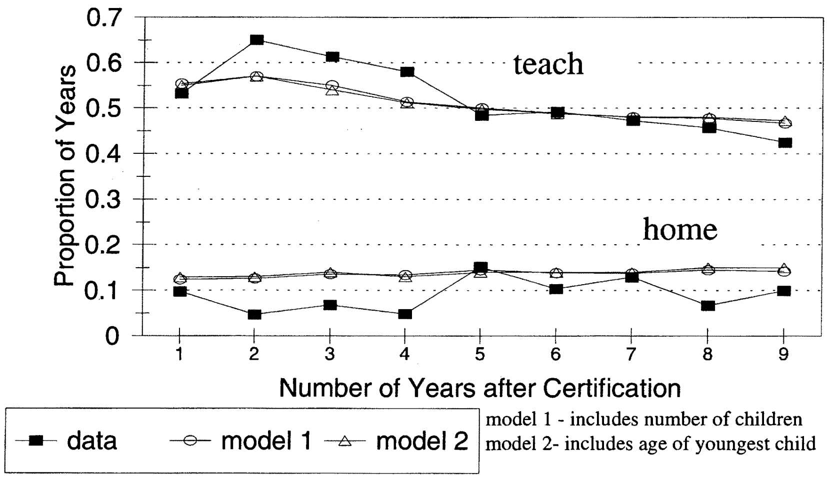

FIGURE 6

PROPORTION AGGREGATE YEARS EACH OPTION, DATA AND SIMULATED MODELS OVER TIME AFTER CERTIFICATION—MALES relatively less likely to work in either the teaching occupation or a nonteaching occupation and become more likely to be out of the workforce altogether. The extent to which the model is able to capture the observed changes in the behavior of males and females which take place after certification can be examined by comparing the simulations from the two models with the data. Figure 5 shows that the models predict a decrease in the proportion of females who choose to teach over time and correctly attribute this decrease to an increase in the proportion of individuals who are out of the workforce. Figure 6 shows that the models also predict a decrease in the proportion of males who choose to teach over time. The fact that the models correctly attribute this decrease to an increase in the proportion of males who choose to work in a nonteaching job can be seen by observing that the simulated proportion of males who are out of the workforce is essentially constant over time. For both males and females, the models somewhat understate the decrease in the teaching proportion which occurs over time. To some extent, this is probably related to the fact that the model is not able to predict the large changes which take place between the first year after certification and the second year.[33]

3.7. *Unobserved Differences.* The estimates of γ_k and π_{10}, $k = 1, 2, 3, 4$, and can be used to compute both the variance of the permanent heterogeneity term in each wage and nonpecuniary equation and the correlations between the various heterogeneity terms. These calculations from the two models indicate that permanent heterogeneity is important.[34] For example, under the first specification in Table 2 (Model 1), the variance of the heterogeneity term in the teaching wage equation, γ_1^w, is 0.037. Thus, γ_1^w accounts for approximately 0.27 of the total unexplained variation in teaching wages (i.e., 0.20 of the variation in $\gamma_1^w + \nu_1(t)$). Similarly, the variance of the heterogeneity term in the nonteaching wage equation, γ_2^w, is 0.055 which accounts for 0.27 of the total unexplained variation in nonteaching wages (i.e., 0.27 of the variation in $\gamma_2^w + \nu_2(t)$). Permanent heterogeneity is found to be even more important in the nonpecuniary equations. Under Model 1, the variance of the nonpecuniary heterogeneity terms (γ_1^q and γ_2^q) is 0.67 and 0.64 in the teaching and nonteaching equations, respectively. Thus, permanent heterogeneity accounts for approximately 0.57 of the total unexplained variation in teaching, nonpecuniary utility (i.e, 0.57 of the variation in $\gamma_1^q + \epsilon_1(t)$) and 0.56 of the total unexplained variation in nonteaching, nonpecuniary utility (i.e, 0.57 of the variation in $\gamma_2^q + \epsilon_2(t)$).[35] Permanent heterogeneity is not found to be important in the birth equation; the calculated variance of γ^B is only 0.014 which is quantitatively unimportant given the assumption that the random shock in the birth equation η is assumed to have a variance of one.

High correlations are found between the various elements of γ. For example, the correlation between the heterogeneity components in the two wage equations

[33] The first year is likely to be a transitional period. For example, some individuals may not actively search for teaching jobs during their final year of school.

[34] For ease of exposition, the discussion here is limited to the results from Model 1. However, as can be seen from Table 2, the results from Model 2 are very similar.

[35] The variances of the extreme value nonpecuniary errors $\epsilon_1(t)$ and, $\epsilon_2(t)$ are given by $\tau^2 pi^2/6$ where the model 1 estimate of τ is 0.554 and pi= 3.141.... Thus, the estimated variances of $\epsilon_1(t)$, $\epsilon_2(t)$, and $\epsilon_3(t)$ are 0.502.

TABLE 3

CHOICES, WAGES, AND BIRTHS OF FOUR TYPES

	Teaching Proportion	Nonteach Proportion	Home Prop.	Wage Teach	Wage Nonteach	Birth Proportion
Type 1 prob = 0.13	0.49	0.37	0.14	123	112	0.08
Type 2 prob = 0.08	0.45	0.52	0.02	232	290	0.11
Type 3 prob = 0.27	0.35	0.1	0.54	169	158	0.16
Type 4 prob = 0.52	0.58	0.34	0.07	179	183	0.11

(γ_1^w and γ_2^w) is 0.97. This could occur because certain unmeasured individual attributes may tend to be desirable to both teaching and nonteaching employers but could also occur because both teaching and nonteaching jobs are likely to pay higher wages in geographic areas where the cost of living is high. The correlation between the heterogeneity components in the nonpecuniary teaching and nonteaching equations (γ_1^q and γ_2^q) is also extremely high, 0.99. Thus, individuals who have a propensity to find teaching to be more appealing than not working also tend to find nonteaching jobs more appealing than not working.

As recognized by Keane and Wolpin (1997), although the exact type of a particular person cannot be observed, Bayes' theorem can be used to compute the posterior probabilities that a particular person is of each of the four types conditional on choices and observed wages. In order to get a better feel for the importance of unobserved heterogeneity in the model, these posterior probabilities are used in Table 3 to examine how the choices, wages, and births differ across types. Not surprisingly there is a large amount of variation in the wages across types. However, as expected, the wages in teaching jobs and nonteaching jobs tend to move together. There are also large differences in the choices that the different types make. Most striking is the Type 3 person who receives very large amounts of nonpecuniary disutility from working in either option relative to staying at home and also has the highest permanent propensity to have children. These factors cause the Type 3 people to stay at home in 0.54 of all person years.

4. POLICY SIMULATIONS AND CONCLUSIONS

From a policy standpoint it is important to understand the effects that particular types of policy changes would have on teacher labor supply. We begin by analyzing the effect of a policy which raises the salary of all teachers by 20 percent. This uniform wage increase, which will be referred to as "policy one," conforms to the rigid wage structure which is currently in place in public schools. Then, to illustrate the potential benefits of deviating from the rigid wage structure, we analyze the effects of a policy which is designed to cost the same amount of money as the uniform increase in policy one, but which allows a person's wage increase to depend on her academic ability. This policy will be referred to as "policy two."

Figure 7 indicates that teachers' labor supply decisions are very sensitive to wages. Policy one and policy two both cause the proportion of aggregate person years that are spent in teaching to increase from 0.50 to 0.80. However, although policy one and

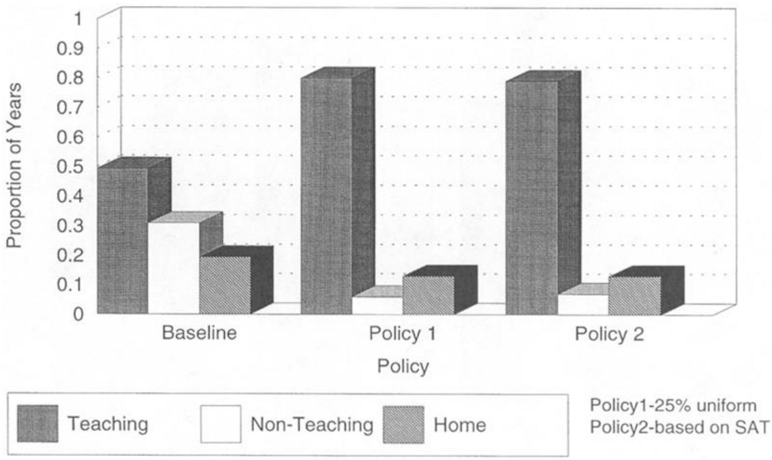

FIGURE 7

PROPORTION AGGREGATE YEARS EACH OPTION, BASELINE, POLICY 1, POLICY 2, FULL SAMPLE USING MODEL

policy two have very similar effects on overall labor supply, Figure 8 indicates that the two policies have somewhat different effects on the types of teachers who choose to teach. In the data, individuals with SAT scores in the top one-third of the sample choose teaching 0.88 as often as individuals with SAT scores in the bottom two-thirds of the sample. Under policy one, this number increases to 0.96. However, the second policy has an even larger positive effect on the academic ability composition of the teaching workforce. Under policy two, individuals with SAT scores in the top third of the sample choose teaching 1.10 as frequently as individuals with SAT scores in the bottom two-thirds of the sample.

Thus, these policy simulations raise the possibility that there may be benefits from deviating from the current, rigid wage structure in some manner. What is less clear is whether policy two should be thought of purely as an illustrative instrument or whether it is the type of policy which deserves serious consideration from the standpoint of implementation. A general caution of the difficulties which might be encountered if a school's attempt to move away from the traditional wage structure is issued by Murnane and Cohen (1986), Murnane et al. (1991), and Cohn (1996) who indicate that past attempts to implement "merit pay" policies have not been particularly successful in practice. The difficulties with these policies seem to arise largely because many teachers perceive that the subjective evaluations which are often used to determine a particular person's merit pay are unfair. However, policy two, which does not rely on subjective evaluations, is not really a merit pay program of the sort discussed by the previous authors. Instead, paying academically gifted teachers more than other teachers would seem to be similar in spirit to the more widely accepted proposal to pay higher wages to individuals with academic backgrounds in certain "shortage"

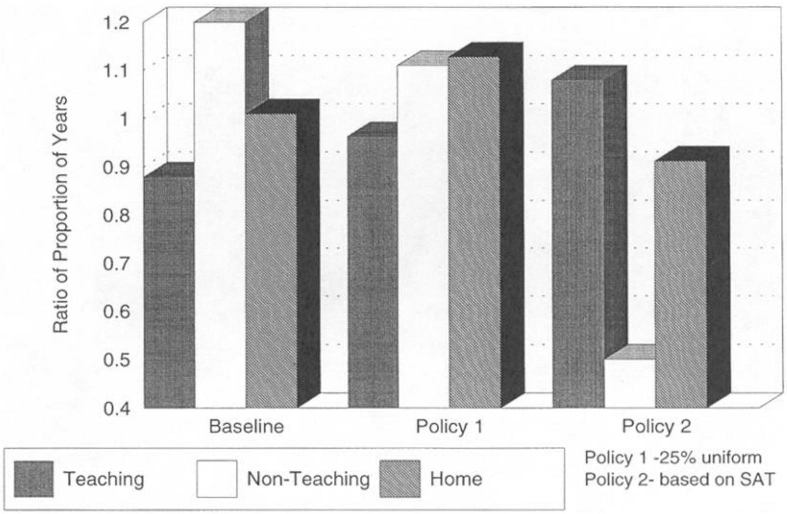

FIGURE 8

RELATIVE LABOR SUPPLY HIGH SAT GROUP, BASELINE, MODEL 1, AND MODEL 2. FIGURE SHOWS PROPORTION OF AGGREGATE YEARS THAT INDIVIDUALS IN HIGH SAT GROUP CHOOSE A PARTICULAR OPTION DIVIDED BY THE PROPORTION OF AGGREGATE YEARS THAT INDIVIDUALS IN LOW SAT GROUP CHOOSE THE OPTION.

subject areas such as science.[36] Clearly, paying teachers with science backgrounds more than other teachers is only justifiable if, on average, individuals with science backgrounds are actually better science teachers than individuals who are certified in other areas. Similarly, the benefit of paying higher wages to academically gifted teachers depends critically on the strength of the relationship between academic ability and teaching ability.[37] The literature review in Section 3.4 suggests that to a large extent the exact nature of this relationship is still an open question.

Most previous research in the area of teacher attrition has utilized teacher-specific data which do not contain information about the family situations of teachers. Thus, little has been known about the effect that fertility outcomes have on teacher decisions. Section 3 indicates that fertility outcomes play a very important role in determining whether a female decides to teach in a particular year. On the surface, this would seem to suggest that giving subsidies to women with newborn children to

[36] Early support for this idea can be found in Kershaw and McKean (1962). It could be argued that academically gifted teachers are a "shortage" areas in many schools.

[37] Clearly academic ability and teacher quality are not perfectly correlated. Thus, this policy would lead to some cases where less gifted teachers are paid more than more gifted teachers. However, this is also true under the policy of paying individuals with academic backgrounds in science fields more than other teachers because, in practice, an individual without a background in science will sometimes turn out to be a better science teacher than a particular individual with an academic background in science. Clearly policy two would not have the on-the-job incentives that a successful merit pay system (with, e.g., yearly evaluations) would have. Therefore, this policy would be the most beneficial if teacher quality tends to be reasonably "permanent" in nature (i.e., good teachers are essentially always good and bad teachers are essentially always bad) and the relationship between academic ability and this "permanent" teacher quality is strong.

help offset the costs of child care might be a cost-effective way to reduce teacher attrition. However, a closer look at the estimates of Table 2 suggests that this is not likely to be the case. For example, the estimated effect that the number-of-children has on the nonpecuniary utility which is derived from teaching suggests that a woman with a single child would have to receive a pay raise of approximately 60 percent to keep her utility from teaching (relative to not working), the same as it would be if she had no children.[38] The young-child estimate from the second column of Table 2 indicates that an even larger subsidy would be needed for a woman who has a child who is younger than the age of two. Further, the large estimated effect of the older-child variable in column 2 suggests that substantial payments would have to be made for at least several years.

The conclusion that the presence of young children has an effect on utility for women which is much larger than the cost of child care can be seen in Figure 7 where the 20 percent wage increases are shown to lead to a very large decrease in the proportion of individuals who choose nonteaching jobs but have only small impacts on the proportion of individuals who are out of the workforce. Nonetheless, although child-care subsidies do not appear to be a particularly promising way to deal with teacher retention problems, other child-care policies may turn out to be very useful. For example, it seems very possible that schools, which already have the infrastructure in place to take care of young children, might find it cost effective to provide inexpensive on-site child care. In effect, this would give parents the new, potentially appealing option of being able to both work and be close to their young children.

Although this article does not examine an individual's certification decision, influencing the set of individuals who choose to become certified is likely to be a primary motivation for raising the wages of teachers. It would seem to be quite likely that wage increases would be successful from this standpoint if the goal is simply to increase the pool of applicants. From the standpoint of increasing the number of academically gifted individuals who are teaching, a particular wage increase will tend to be successful if it has beneficial consequences on the ability composition of those seeking certification, or if it increases the number of certified teachers who are looking for jobs *and* schools hire the best applicants.[39] However, if encouraging academically gifted teachers to seek certification and enter teaching is the primary purpose of a wage increase, it is important to note that the decisions of individuals who are currently teaching will be important in determining the amount of time which will be needed before turnover of current teachers occurs and the new teachers can be hired into the schools. The simulations in this article suggest that a substantial increase in wages would significantly slow the rate at which current teachers leave.

[38] Recall that nonpecuniary utility is in log wage units.

[39] Ballou (1996) finds that schools may not necessarily hire the most qualified applicants. If teachers are selected randomly from the pool of applicants, simply having a larger pool of applicants will not increase quality.

REFERENCES

BALLOU, D., "Do Public Schools Hire the Best Applicants?," *Quarterly Journal of Economics* 111 (1996), 97–133.

BELLMAN, R., *Dynamic Programming* (Princeton, NJ: Princeton Univ. Press, 1957).

BERKOVEC, J., AND S. STERN, "Job Exit Behavior of Older Men," *Econometrica* 59 (1991), 189–210.

BISHOP, J. H., "Incentives to Study and the Organization of Secondary Instruction," in W. Becker and W. Baumol, eds., *Assessing Educational Practices* (Cambridge, MA: MIT Press, 1996), Chap. 5.

BREWER, D., "Career Paths and Quit Decisions: Evidence from Teaching," *Journal of Labor Economics* 14 (1996), 313–39.

COHN, E., "Methods of Teacher Remuneration: Merit Pay and Career Ladders," in W. Becker and W. Baumol, eds., *Assessing Educational Practices* (Cambridge, MA: MIT press, 1996), Chap. 8.

DOLTON, P., AND W. VAN DER KLAAUW, "The Turnover of Teachers: A Competing Risks Analysis," *The Review of Economics and Statistics* 81 (1999), 543–550.

——— AND ———, "Leaving Teaching in the UK, A Duration Analysis," *The Economic Journal* 105 (1995), 431–44.

EBERTS, R., "Union-Negotiated Employment Rules and Teachers Quits," *Economics of Education Review* 6, (1987), 15–25.

EHRENBERG, R., AND D. BREWER, "Did Teacher's Race and Verbal Ability Matter in the 1960's," (Ithaca, NY: Cornell University, School of Industrial and Labor Relations, 1993), pp. 1–57.

FERGUSON, R., "Racial Patterns in how School and Teacher Quality Affect Achievement and Earnings," report, Kennedy School of Government, Harvard University, Cambridge, MA, 1990.

GRITZ, M. R., AND N. D. THEOBALD, "The Effects of School District Spending Priorities on Length of Stay in Teaching," *Journal of Human Resources* 31 (1996), 477–512.

HANEY, W., R. MADAUS, AND A. KREITZER, "Charms Talismanic: Testing Teachers for the Improvement of American Education," in E. Rothkopf, ed., *Review of Research on Education* (Washington, D.C.: American Educational Research Association, 1987).

HANUSHEK, E., "The Economics of Schooling," *Journal of Economic Literature* 24 (1986), 1141–77.

———, "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-Data," *American Economic Review*, 61 (1971), 280–88.

HECKMAN, J., AND B. SINGER, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52 (1984), 271–320.

KEANE, M., AND K. WOLPIN, "The Career Decisions of Young Men," *Journal of Political Economy* 105 (1997), 473–522.

KERSHAW, J., AND R. MCKEAN, *Teacher Shortages and Salary Schedules* (New York: McGraw–Hill, 1962).

MONK, D., "Subject Area Preparation of Secondary Mathematics and Science Teachers and Students Achievement," *Economics of Education Review* 13 (1994), 125–145.

MONT, D., AND D. REES, "The Influence of Classroom Characteristics on High School Teacher Turnover," *Economic Inquiry* 34 (1996), 152–67.

MURNANE, R., AND D. COHEN, "Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and a Few Survive," *Harvard Education Review* 56 (1986), 1–17.

——— AND R. OLSEN, "The Effects of Salaries and Opportunity Costs on Duration in Teaching: Evidence from Michigan," *Review of Economics and Statistics* 71 (1989), 347–52.

——— AND ———, "The Effects of Salaries and Opportunity Costs on Length of Stay in Teaching, Evidence from North Carolina," *The Journal of Human Resources* 25 (1990), 106–24.

———, J. SINGER, AND J. WILLET, "The Influences of Salaries and Opportunity Costs on Teachers' Career Choices: Evidence from North Carolina," *Harvard Educational Review* 59 (1989), 345–46.

———, ———, ———, J. JAMES, AND J. RANDALL, *Who Will Teach? Policies That Matter* (London: Harvard Univ. Press, 1991).

———, J. WILLET, AND F. LEVY, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics* 77 (1995), 251–66.

RICKMAN, B., AND C. PARKER, "Alternative Wages and Teacher Mobility: A Human Capital Approach," *Economics of Education Review* 9 (1990), 73–79.

Rivken, S., E. Hanushek, and J. Kain, "Teachers, Schools and Academic Achievement," National Bureau of Economic Research paper, 1998.

Rust, J., "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55 (1987), 999–1033.

———, "An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Quit," mimeo, 1996b.

Stinebrickner, T., "An Empirical Investigation of Teacher Attrition," *Economics of Education Review* 17 (1998), 127–136.

———, "Estimation of a Duration Model in the Presence of Missing Data," *Review of Economics and Statistics* 81 (1999), 529–542.

———, "Serially Correlated Wages in a Dynamic, Discrete Choice Models," *Journal of Applied Econometrics* 15 (2000), 595–624.

———, "A Dynamic Model of Teacher Labor Supply," *Journal of Labor Economics* 19 (2001), 196–230.

———, "An Analysis of Occupational Change and Departure from the Labor Force: Evidence of the Reasons that Teachers Leave," *Journal of Human Resources*, forthcoming.

Strauss, R., and E. Sawyer, "Some New Evidence on Teacher and Student Competencies," *Economics of Education Review* 5 (1986), 41–48.

Theobald, N., and M. Gritz, "The Effects of School District Spending Priorities on the Exit Paths of Beginning Teachers Leaving the District," *Economics of Education Review* 15 (1996), 11–22.

van der Klaauw, W., "Expectations and Career Decisions: An Analysis of Teaching Careers Using Expectations Data," working paper, 1996a.

———, "Female Labor Supply and Marital Status Decisions: A Life Cycle Model," *Review of Economic Studies* 63 (1996b), 199–236.

Wolpin, K., "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy* 92 (1984), 852–74.