Maternal Employment, Migration, and Child Development[1]

Haiyong Liu
East Carolina University[2]

Thomas Mroz
University of North Carolina at Chapel Hill

Wilbert van der Klaauw
Federal Reserve Bank of New York

December 2006

Abstract

We analyze the roles and interrelationships between school inputs and parental inputs in affecting child development through the specification and estimation of a behavioral model of household migration and maternal employment decisions. We integrate information on these decisions with observations on child outcomes over a 13-year period from the NLSY. We find that the impact of our school quality measures diminish by a factor of 2 to 4 after accounting for the fact that families may choose where to live in part based on school characteristics and labor market opportunities. The positive statistical relationship between child outcomes and maternal employment reverses sign while remaining statistically significant after controlling for its possible endogeneity. Our estimates imply that when parental responses are taken into account, policy changes in school quality end up having only minor impacts on child test scores.

[1]We have benefited from helpful comments from participants at the 2002 SITE meetings at Stanford, participants at the 2003 PET meetings at Duke, participants at the 2005 Conference on Structural Models in Labor, Aging and Health, and seminar participants at the University of Essex, University of Wisconsin-Madison, Northwestern University, Johns Hopkins University, University of Pennsylvania, Laval University and Ohio State University.

[2]Corresponding author: Department of Economics, Brewster A-439, Tenth Street, East Carolina University, Greenville, NC 27858, email: liuh@ecu.edu

1 Introduction

Understanding the impacts of school inputs and parental behavior on children's educational development can provide crucial information to both policy makers and parents as they make decisions about the allocation of resources to children. The literature providing estimates of educational production functions, however, provides little consensus about the magnitude or even the direction of the impact of many school and family inputs to children's development. Two of the most recent reviews of the literature on the impact of school resources and school "quality" on educational outcomes, by Hanushek et al. (1998) and Krueger (1998), present conflicting interpretations of the literature. Similarly, the extensive literature on the impact of parental inputs on educational production as surveyed by Haveman and Wolfe (1995) reports widely varying effect estimates.

One potential reason for the ambiguity in the literature appears, at a first glance, to be due to the fact that it is impossible to collect data on all of the relevant inputs to the educational production function. Some studies have detailed measures of school inputs but almost no information about parental inputs (e.g. Hanushek, 1996, Krueger, 1999). Other studies have good information on what parents do for their children with only limited information on the environment in the schools the children attend (e.g. Harvey, 1999; Moore and Driscoll, 1997). There are also experimental impact studies where children have been randomly placed in different schooling environments (Krueger, 1999). Researchers have used all of these different types of data resources to estimate the impacts on child educational outcomes of the inputs to educational production.

A recent paper by Todd and Wolpin (2002) addresses explicitly several reasons for why studies using different data sources could give rise to different estimates of the impacts of school inputs on child outcomes. Their basic interpretation of the literature is that researchers have failed to use a coherent, common theoretical framework in the interpretation of their empirical results. Todd and Wolpin posit that the education production process uses inputs that are determined by both parents and schools and that the levels of these inputs are influenced by each child's ability to use the inputs productively. If the data are not rich enough to include measures of both parents' and schools' inputs and child background characteristics, and if there are correlations among these different types of inputs, then the observed inputs will be correlated with the implicit error terms. The resulting endogeneity biases will make it difficult to compare estimates based on different data sets. Todd and Wolpin also point out that even randomly assigned, experimentally determined inputs do not resolve such problems. If parents and schools can adjust their input decisions in response to random assignments, then experimentally estimated impacts will convolute the true production impacts of the experimentally-assigned inputs with the impacts of adjustments to other inputs made by parents and schools in response to the experimentally-induced input decisions. While the resulting estimates may still measure useful policy effects, they are not estimates of educational production function parameters.

The problems of estimating an educational production function are not solely due to data limitations. At one extreme is the lack of rigor in defining what researchers would like to learn about educational production that Todd and Wolpin highlight; researchers have been unclear about what they would like to estimate. It is important to define precisely what one can estimate. Additionally, researchers have focused on simple functional forms for the production function. This can give rise to unstable estimates that can vary widely and depend crucially on the sample under consideration.

We attempt to address a wide variety of these issues. We start by specifying a formal model of parental decisions about the choice of schools and parental involvement with their children. We implement this formal model by assuming that parents choose their place of residence in part because of employment opportunities and in part because of the characteristics of the schools where they choose to reside. We distinguish between various measures of school quality to assess school inputs, while maternal employment is the main parental input considered. Parents make their place of residence decisions based on location-specific wage offer distributions and school quality measures, but before knowing the exact child outcome and wage offers they might receive. After making a place of residence decision, both parents receive wage offers. While the husband is assumed to always accept the wage, the mother chooses how much time to devote to labor market activities at that wage rate. She does this with the understanding that more hours spent in the labor market could cut back on her inputs to the education production function. Even though she knows the average effect that her work behavior would have on her child's educational outcomes, she does not know what the child's actual educational outcomes will be.

An important unanswered question from the literature on education production is whether parental and school inputs are substitutes or complements to each other in the process of education production. Moreover, if substitutability and complementarity are co-existing, it is not clear which one is dominant among families at different socio-economic strata. If the substitutability of two inputs is prevalent it would be expected that a mother spends less time with her children once she puts them in a school with a perceived better quality, holding everything else constant. Similarly, the mother should be more actively involved in educating her children by sacrificing leisure time or working fewer hours in the labor market if she perceives the school quality to be poor. At the extreme, parents may choose to take the sole responsibility of educating their children, reflected by the increasing popularity of home schooling (Behrman and King, 2001). If the hypothesis holds that parental inputs and school inputs are substitutes, then results from studies of the effects of teacher and school quality might be biased downward due to the lack of control for parental inputs. On the other hand, some empirical evidence seems to support the existence of complementarities between school and parental inputs in the production of education. For instance, several educational reform programs actually mandate that the parents be more involved (McMillan, 2001). If this complementarity effect dominates, impact estimates of parental inputs from regressions that do not properly control for school quality are likely to be biased upward.

In summary, this study combines an economic model of how parents respond to the child's school environment with the parents' choice of school environment to provide a more complete understanding of the relationship between parental inputs and school inputs, and their age-specific and cumulative impacts on child development. This study investigates the following policy issues: 1) the effects of exogenous changes in several school quality measures on mother's labor supply decisions, 2) the effects of a change in mothers' wages on labor supply, schooling choices, and outcomes, and 3) the impacts of family and school characteristics on children's test scores. This study provides a unique perspective on designing educational policies as well as social programs targeted to encourage maternal employment.

2 Background

The literatures on the impacts of school resources and school 'quality' on educational outcomes on the one hand and on the impact of parental inputs on educational production on the other are extensive and have been reviewed in detail by others. Here we focus on some recent discussions of potential reasons for the disparate empirical findings in the literature.

One important source of variation across studies has been the measure of child outcome used. For example, a frequently cited paper by Card and Krueger (1992) argued that long-term labor market outcomes (such as earnings) are better output measures than test scores. Unlike several influential studies based on test scores, their paper reported significant school input effects. More evidence by Betts (1995, 1996) using NLSY data, however, suggests that the choice of outcome variable is not the reason for the difference in findings.

A second potential source of variability in estimates concerns the data source used. It has been argued that because of input endogeneity issues, as well as measurement error issues, that using student level data is likely to lead to biased estimates; grouped or more aggregated data presumably are less likely to suffer from such problems (Loeb and Bound, 1996; Hanushek et al, 1996). In addition, as pointed out by Todd and Wolpin (2002), one can expect input effect estimates based on experimental data to be very different from those based on non-experimental data, because they generally would estimate entirely different causal effects. The first would include short-term parental responses to the experimental input assignment, while those based on non-experimental data usually estimate production function parameters.

Another important source of heterogeneity across studies is the identification strategy used to deal with the endogeneity of input choices in estimating their effects on educational outcomes. Input choices are likely to depend on a child's innate ability. Evaluations in which the endogeneity of school and parental inputs are ignored are likely to produce unreliable results which could vary considerably with the set of variables included in the regression. The choice of econometric method to solve the endogeneity problem is also likely to be a factor. As argued by Card (1995, 1999), who investigated the variability in estimated returns to years of schooling across studies, if input effects vary across students, schools or locations, or if these effects are nonlinear, different evaluation strategies usually identify local average treatment effects (Imbens and Angrist, 1994) which apply to different sub-populations.

While we believe that all these alternative explanations for the large variation in estimates in the literature have merit and undoubtedly will play a role, it is our view that the specification of the educational production function and the fact that many inputs are often ignored or missing have not received sufficient attention in the literature. In particular, if parental inputs are chosen jointly or in response to school inputs, then this raises serious questions about the validity of instruments that underly most current identification approaches to deal with the endogeneity of school characteristics.

Besides ample evidence of an effect of maternal employment on child outcomes (e.g., Blau and Grossberg, 1992; Parcel and Menaghan, 1994; Bernal, 2003), there exists a growing literature documenting that parents influence and respond to school quality differences. In Tennessee's STAR class size experiment, for example, because of complaints by the parents of children who had been randomly assigned to larger classes, a teacher aide was subsequently assigned randomly to such classes (Krueger, 1998). Second, studies of housing values indicate that parents value better schools, as differences in school expenditures and average test scores have been found to be factored into housing prices (e.g. Black, 1999). In making location choices parents have been found to respond strongly to school quality differences as well as employment opportunities, and both incentives appear to be strongly related (Blanchard and Katz, 1992; Rapaport, 1997; Nechyba and Strauss, 1998; Bayer, 2000; Ferreyra, 2001, Kennan and Walker, 2002). Hedges and Greenwald (1996) have suggested that the interaction between school inputs and family inputs may explain the small school input effects found in many studies. They propose that the increase in female labor force participation rates and the rising prevalence of single parent households may have offset the positive effects of increased spending and declines in class sizes, to produce no overall improvements in outcomes.

Our approach for investigating the interaction between and endogeneity of school and family inputs is to specify and estimate a structural model of school choice and maternal employment decisions, taking into account the dependence of both choices on parental preferences and financial constraints as well as the child's innate ability. As in Bayer (2000), who estimates a static equilibrium model of household location and school choice decisions, this allows us to estimate an educational production function while controlling for the non-random sorting of households across locations and schools. In addition this approach explicitly addresses the endogeneity of the mother's work decision.

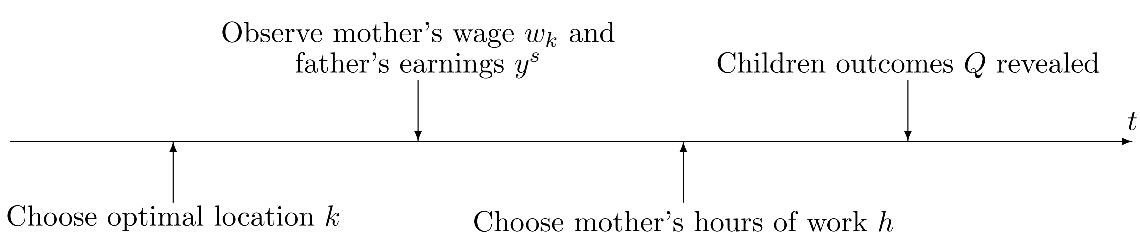
3 The model

In this section we present our theoretical model, as well as the estimation method that we use to estimate it.

3.1 Introduction

In each period a decision is made on where to live, how much to consume, and how many hours the mother will work. These decisions are made taking into account their expected effect on the distribution of the child's cognitive achievement. At the beginning of each period t (where each period corresponds to a given child's age), a family first makes a location decision on the basis of the expected utility of living in each location. Locations are characterized by a set of school quality indicators, local labor market conditions, moving and housing costs and their geographic location. The choice of location $k \in \{1, 2, 3, \cdots, K\}$ in each period depends on the wage offer distributions in each location, the child achievement score distribution (given school quality), as well as the costs of moving. After choosing a location of residence, the father (if present) receives an earnings offer y^s and the mother receives a wage offer w. While fathers are always assumed to accept their earnings offer in each period, which could be zero, the mother makes an employment decision h. Given the location and work choice, school and family inputs produce a child's cognitive achievement score q.

The time line of this model is described as follows:



Consider a one-period model in which a family i with N_i children has preferences over consumption, x_i, a vector of the children's cognitive achievements, $Q_i = (q_{i1}, \ldots, q_{iN_i})$, the mother's hours devoted to non-market activities, l_i, and the geographic location of residence, k_i. l_i is defined to include maternal time directly devoted to child education, other household production activities, and leisure. Explicitly, it satisfies

$$T = h_i + l_i, \tag{1}$$

where T is total time available to the mother, and h_i represents mother's hours spent working in the market. It is therefore equivalent to state preferences in terms of h, instead of l.

Let those preferences be represented by the utility function,

$$U^i = U^i(x_i, Q_i, h_i, k_i).$$

In our empirical analysis we model the mother's hours of work as a discrete choice variable equal to zero (h_0), part-time (h_1), or full-time (h_2) and adopt the following specification of the utility function:

$$\begin{aligned} U^i(x_i, Q_i, h_i, k_i) \quad = \quad & \frac{(x_i + \gamma_{0i})^{\gamma_1}}{\gamma_1}[1 + \alpha_0 MS_i + \alpha_1 \cdot \mathbf{1}(h_i > 0)] \; + \frac{\alpha_2}{\gamma_3}\sum_{n=1}^{N_i}(\frac{q_{in}}{\bar{q}_i^{\alpha_3}} + \gamma_2)^{\gamma_3} \\ & + \alpha_{4ai} \cdot \mathbf{1}(h_i > 0) + \alpha_{4bi} \cdot \mathbf{1}(h_i = h_2) + \sum_{j=0}^{2} \mathbf{1}(h_i = h_j)\varepsilon_{h_j,i} \\ & + m_{k_{-1}k} + \sum_{j=1}^{4} \alpha_{5j} \mathbf{1}[R(k_i) = j] + \xi_{k,i}, \end{aligned} \tag{2}$$

where MS is an indicator for the marital status of the mother, and $\mathbf{1}(\cdot)$ is the indicator function equal to 1 if the argument is true and 0 if not. Note that in household i the child n's cognitive achievement score q_{in} is assumed to enter as a relative score, measured relative to a power function of the mother's Armed Forces Qualification Test(AFQT) score, $\bar{q}_i$. This allows high achieving mothers to have different standards in evaluating child success than lower achieving mothers, and it includes as a special case the mother's score being irrelevant (when $\alpha_3 = 0$). The parameter α_{4ai} measures the disutility of working part-time, when compared to not doing any market work, and α_{4bi} measures the additional disutility from working full-time. A corresponding change in the marginal utility of consumption when working compared to not-working is captured by α_1.

The specification also allows for direct geographic preferences for living in each of the four census regions of the U.S., as captured by the parameters α_{5j} $(j = 1, \ldots, 4)$, where the $R(k) \in \{1, 2, 3, 4\}$ denotes the region corresponding to location k. $m_{k_{-1}k}$ measures the psychic cost associated with moving from the previous period's location k_{-1} to k, which is further allowed to differ between moves within and across state borders and geographic census regions, as in

$$(3) \quad m_{k_{-1}k} = \begin{cases} \delta_0 + \delta_1 \mathbf{1}[ST(k) \neq ST(k_{-1})] + \delta_2 \mathbf{1}[R(k) \neq R(k_{-1})] & \text{when } \; k \neq k_{-1} \\ 0 & \text{otherwise.} \end{cases}$$

The function $ST(\cdot)$ maps each residential location into its corresponding state in the U.S. The stochastic components $\varepsilon_{h,i} = (\varepsilon_{h_0,i}, \varepsilon_{h_1,i}, \varepsilon_{h_2,i})$ are maternal evaluations of the unobserved attributes of the three employment states. They are assumed to be independently and identically distributed over different working hours and across individuals and time periods, with an extreme value distribution with variance parameter b_1. In addition, $\xi_{k,i}$ represents random variation in the evaluation of the unobserved attributes of community k, which is assumed to be independently and identically distributed over different communities, individuals and time. It follows an extreme value distribution with variance parameter b_2.

In the empirical model, we introduce two additional sources of heterogeneity in preferences. First, we specify the parameters measuring the disutility from working, α_{4ai} and α_{4bi}, as functions of the mother's marital status, the number of children in the household between the ages of 0 and 5 (n_{yi}), the number of children between the ages of 6 and 17 (n_{oi}), as well as a time-invariant mother-specific heterogeneity component μ_i^m:

$$\alpha_{4ai} = f_{4a}(MS_i, \mu_i^m, n_{yi}, n_{oi}) \qquad \alpha_{4bi} = f_{4b}(MS_i, \mu_i^m, n_{yi}, n_{oi}).$$

Second, we allow for individual heterogeneity in the "reserve consumption" value γ_{0i} by specifying $\gamma_{0i} = f_0(\mu_i^m)$.

The husband's earnings are assumed to be stochastic and location specific. The presence of a father directly influences preferences for leisure and consumption. In addition, as described below, it directly affects child outcomes. This mostly static, empirical framework implicitly treats the parents' prior family formation, fertility and education behavior as exogenous in the analysis of current choice behavior. In addition we ignore capital markets, by assuming that parents do not save or borrow. Hence, the budget set when residing in location k can be expressed as

$$x_i = w_{ki}h_i + y^s_{ki} + y^o_i - c_{ki}, \tag{4}$$

where w_{ki} represents the wage offered to the mother in location k, which is observed only after a person has decided to locate in k, and y^s_{ki} are the earnings of the father. Household non-labor income is denoted by y^o_i, and c_{ki} explicitly measures the average housing cost in location k.[1]

As will be discussed in greater detail later, the child quality measure q_{in} we use is discrete and is modeled using a very flexible parametric specification for the child quality production function, defined as

$$q_{in} = f_q(t_{in}, b_{in}, S_k, T - h_i, Z_i, MS_i, \mathbf{1}(k \neq k_{-1}), \mu^c_i, \varepsilon_{q_n,i}), \tag{5}$$

where t represents the child's age, b is a indicator for the child gender, and S_k is a vector of school characteristics in location k believed to influence child outcomes, including school quality measures such as per pupil expenditures and the average teacher salary, relative to average earning in other professions. It also includes the average high school dropout rate, measuring community level differences across school districts, to capture school quality and peer effects. Z_i is a vector of mother's characteristics in the current period including the mother's education, age, AFQT score, race and marital status.

The specification allows for a (one-period) reduction in the child's performance due to a move. Unobserved heterogeneity in the child's intellectual endowments is captured by the inclusion of μ^c_i. This term represents the child's unobserved endowment at the age at which they enter our data set (age 5 or 6). The stochastic component $\varepsilon_{q_n,i}$ represents all other unobserved factors influencing the child's outcome at each age, and it is assumed to be independently distributed across individuals and periods. The exact specification of f_q and the distribution of $\varepsilon_{q_n,i}$ will be discussed later.

Similar to the treatment of child achievement scores, the mother's hourly wage rates are discretized and the wage offer distribution function in location k is specified as a flexible

[1]Since moving costs are unobserved, to avoid identification problems, we do not attempt to measure monetary moving costs.

parametric function of the variables describing local labor market conditions ($\overline{W_{ki}}$), mother characteristics in Z_i and the unobserved trait μ_i^m, as in

$$w_{ki} = f_w(\overline{W_{ki}}, Z_i, \mu_i^m, \varepsilon_{w,i}), \tag{6}$$

where $\varepsilon_{w,i}$ captures stochastic changes over time in the offered wage rate. While each woman is assumed to know the wage offer distribution associated with each location, she is assumed to receive only one wage offer each period and only for the chosen location of residence. The specification of f_w and the distribution of $\varepsilon_{w,i}$ will be discussed later.

New earnings realizations for the husband are modeled in a similar way. They are specified as a flexible parametric distribution of local labor market variables for men, the husband's characteristics, and the spouse's unobserved ability endowment μ_i^s, where $\varepsilon_{y,i}$ are i.i.d. earnings shocks, as in

$$y_{ki}^s = f_y(\overline{W_{ki}^s}, Z_i^s, \mu_i^s, \varepsilon_{y,i}). \tag{7}$$

3.2 Value Functions

Location and work decisions are made sequentially with a family first making a location decision. This is followed by the location specific realizations of the husband's earnings and the wife's wage offers, and then the mother makes her work decision. After each of these decisions has been made the child's achievement score is realized. When deciding on her hours of market work, the mother takes into account the effect of her work decision as well as of the location-specific school quality level on the expected child outcome. Similarly, when the location decision is made, the effect of this decision on subsequent wage offer distributions, work decisions and the effect of these on the distribution of child outcomes also are taken into account. In doing so, both parental choice decisions will depend on the child's innate ability as in Becker and Tomes (1976). The solution to this stochastic optimization problem can be obtained by solving backwards, first for the optimal work decision for each possible location choice, husband earnings realization and wage offer, and second, by determining the optimal location choice.

If the mother ends up in location k, with the husband earning y_k^s, the mother receiving wage draw w_k, and learning the utility shocks associated with each work decision ($\varepsilon_h =$

$(\varepsilon_{h_0}, \varepsilon_{h_1}, \varepsilon_{h_2})$), the expected utility associated with each hours of work choice can be defined as

$$V_k^h(\Omega_k, w_k, y_k^s, \mu, \varepsilon_h, \xi_k) = E[U(x(h), Q(h), h, k)|\Omega_k, w_k, y_k^s, \mu, \varepsilon_h, \xi_k], \tag{8}$$

where Ω_k is a set of state variables, $\Omega_k = \{t_1, t_2, \ldots, t_N, S_k, Z, k_{-1}\}$, μ represents the vector (μ^m, μ^s, μ^c), and $x(h)$ and $Q(h)$ are the values of consumption and the children's achievement scores when the mother works h hours as implied by (4) and (5). We have dropped the individual subscripts, i, to simplify notation.

This requires the calculation of an expectation because the children's educational outcomes are stochastic (though influenced by mother's hours of work choice) and realized after the work decision is made. With a finite number of outcome scores, for each hours of work choice, the expected utility to the family after observing her wage draw, her husband's earnings and utility shocks, and integrating over all possible test score outcomes equals

$$\begin{aligned} V_k^h(\Omega_k, w_k, y_k^s, \mu, \varepsilon_h, \xi_k) &= \sum_{n=1}^{N}\sum_{p=1}^{P}\{\Pr(q_n = p\ |\Omega_k, w_k, y_k^s, \mu, h, k)\, U\left[x(h), q(h), h, k\right)]\} \\ &= \overline{V}_k^h(\Omega_k, w_k, y_k^s, \mu) + \varepsilon_h + \xi_k, \end{aligned} \tag{9}$$

where the second equality follows from the additive separability of the utility function in both errors, combined with the assumed independence (conditional on μ) of (ε_h, ξ_k) and ε_q.

This implies that given the choice of a location k, a wage rate offer w_k and husband's earnings realization y_k^s, the optimal work decision can be defined as

$$h_k = argmax\{\overline{V}_k^h(\Omega_k, w_k, y_k^s, \mu) + \varepsilon_h\}. \tag{10}$$

With ε_h being independently, and identically distributed extreme value errors, the probability of woman from household i working h_j $(j = 0, 1, 2)$ hours conditional on wage and earnings draws w_k and y_k^s equals

$$\Pr(h_{ki} = h_j|\Omega_{ki}, w_{ki}, y_{ki}^s, \mu_i) = \frac{\exp\left[\frac{\overline{V}_k^{h_j}(\Omega_{ki}, w_{ki}, y_{ki}^s, \mu_i)}{b_1}\right]}{\sum_{j'=0}^{2}\exp\left[\frac{\overline{V}_k^{h_{j'}}(\Omega_{ki}, w_{ki}, y_{ki}^s, \mu_i)}{b_1}\right]}. \tag{11}$$

At the time of the location decision, while each woman knows the <u>distributions</u> of her wage offer and the husband's earnings at each location $k = 1, 2, \cdots, K$, she knows neither the actual wage she will be offered nor what her husband will earn at each location. Equally important, she does not know the realizations of the utility shocks associated with her future work decision or the shocks to the test score outcome discussed above. For any given offered wage and husband's earnings level in location k, the expected maximum utility associated with choosing location k is given by

$$
\begin{aligned}
(12) \qquad V_k(\Omega_k, w_k, y_k^s, \mu, \xi_k) &= E \max_h V_k^h(\Omega_k, w_k, y_k^s, \mu, \varepsilon_h, \xi_k) \\
&= b_1 \log \left\{ \sum_{h'=0}^{2} \exp\left(\frac{\overline{V}_k^{h'}\left[\Omega_k, w_k, y_k^s, \mu\right]}{b_1} \right) \right\} + \xi_k \\
&= \overline{V}_k(\Omega_k, w_k, y_k^s, \mu) + \xi_k,
\end{aligned}
$$

where the expectation is with respect to the extreme value distributed utility shocks ε_h, b_1 is the variance of these shocks, and the maximization is subject to the budget constraint (4).

When making a residential location decision, the actual wage offers and husband earnings realizations in each possible destination are unknown. Integrating the value function defined in equation (12) over G possible (discrete) wage offers $(\bar{w}_g, g = 1, \ldots, G)$ and G possible husband's earnings realizations $(\bar{y}_{g'}^s, g' = 1, \ldots, G)$ in location k yields

$$
\begin{aligned}
(13) \quad EV_k(\Omega_k, \mu, \xi_k) &= \sum_{g=1}^{G} \sum_{g'=1}^{G'} \left[\Pr(w_k = \bar{w}_g) \cdot \Pr(y_k^s = \bar{y}_{g'}^s) \cdot \overline{V}_k(\Omega_k, w_k, y_k^s, \mu) \right] + \xi_k \\
&= \widetilde{V}_k(\Omega_k, \mu) + \xi_k.
\end{aligned}
$$

At the time of the location decision ξ_k is known, and the optimal location decision for the agent can be defined as

$$
(14) \qquad k = argmax\{\widetilde{V}_k(\Omega_k, \mu) + \xi_k\}.
$$

Assuming $i.i.d.$ extreme value errors for the ξ_k, the probability to the researcher that location k is chosen by family i is therefore given by

$$
(15) \qquad \Pr(k_i = k | \Omega_i, \mu_i) = \frac{\exp\left[\frac{\widetilde{V}_k(\Omega_{ki}, \mu_i)}{b_2}\right]}{\sum_{k'=1}^{K} \exp\left[\frac{\widetilde{V}_{k'}(\Omega_{k'i}, \mu_i)}{b_2}\right]},
$$

where $\Omega_i = \{(\Omega_{ki}, L_{ki}); k = 1, .., K\}$, and L_{ki} represents the local labor market conditions at location k for individual i.

While we do make extensive use of multinomial logit formulations for describing the utility maximizing choices, it is important to note that our estimation strategy does not require the strong independence of irrelevant alternatives assumption. In particular, our use of person and family specific unobservables in the utility and production functions allows for the very real possibility that there can be correlation in utilities across all possible choices.

3.3 Wage, Spousal Earnings and Educational Production Functions

Empirically, the child achievement score q is discretized and the educational production function in (5) is specified using a flexible parametrization of its corresponding "hazard rate." More specifically, we follow Gilleskie and Mroz (2004) in modeling the probability of an advance to a higher level, conditional on reaching a given achievement level, (i.e. one minus the hazard rate) using the logit function

$$\Pr(q_{in} > p | q_{in} \geq p) = \frac{\exp\left[f_{qin} + f_{bq}(p)\right]}{1 + \exp\left[f_{qin} + f_{bq}(p)\right]}, \tag{16}$$

where q_{in} is the observed score of child n in household i, $f_{bq}(p)$ represents the "baseline" hazard at achievement level p and f_{qin} captures the effect of covariates.

The baseline hazard is specified as a third-degree polynomial in the current (potential) achievement level p as follows:

$$f_{bq}(v) = \sum_{\eta=1}^{3} \left\{\theta_{q\eta}[\log(P-p)]^{\eta}\right\}, \tag{17}$$

where P is the highest score level. The covariate component is further specified as

$$\begin{aligned} f_{qin} &= f_q(t_{in}, S_k, T - h_i, Z_i, MS_i, \mathbf{1}(k \neq k_{-1}), \mu_i^c) \\ &= \beta_{q0} + \beta_{q1} t_i + \beta_{q2} b_i + \beta_{q3} S_k + \beta_{q4a} \mathbf{1}(h_i = h_1) + \beta_{q4b} \mathbf{1}(h_i = h_2) \quad (18) \\ &\quad + \beta_{q5} Z_i + \beta_{q6} MS_i + \beta_{q7} \mathbf{1}(k \neq k_{-1}) + \rho_q(\mu_i^c) + \text{higher order terms} + \text{interactions}, \end{aligned}$$

where S_k is the vector of school inputs defined earlier, Z_i is a vector of characteristics of the mother expected to influence the child's performance, and $\rho_q(\mu_i^c)$ is the estimated effect of the unobserved heterogeneity component μ_i^c. Empirically, it is approximated by a third degree polynomial in μ^c, where μ^c has a discrete distribution taking on J different values, with corresponding probabilities $\Pr(\mu_j^c),\ j = 1, \ldots, J$.

Note that this specification is very flexible, allowing effects of covariates to differ at different levels of the child's achievement score. In the estimations we also relax several of the separability assumptions implicit in (16) to allow for interactions of the baseline hazard, the age of the child, unobserved heterogeneity and school and parental characteristics by including interactions between $f_{bq}(p)$ and several determinants of f_{vin}. This flexibility allows us to estimate the degree of substitutability or complementarity across all inputs.

We similarly model the mother's wage offer distribution, discretizing the wage data into G groups, with the probability density function $\Pr(w_k = \bar{w}_g)$ of having a wage offer $\bar{w}_g$ similarly defined in term of the corresponding conditional hazard probabilities

$$\Pr(w_k > \bar{w}_g | w_k \geq \bar{w}_g) = \frac{\exp[f_{wi} + f_{bw}(g)]}{1 + \exp[f_{wi} + f_{bw}(g)]} \text{, where} \tag{19}$$

$$f_{wi} = \beta_{w0} + \beta_{w1}\overline{W_{ki}} + \beta_{w2}Z_i + \rho_w(\mu_i^m) \tag{20}$$

$$+\text{higher order terms} + \text{interactions,}$$

$$f_{bw}(g) = \sum_{\eta=1}^{3} \{\theta_{w\eta}[\log(G-g)]^{\eta}\}, \tag{21}$$

and $\overline{W_{ki}}$ is the median local earnings for women with the same education level, race, and age. We allow for interactions between the baseline hazard f_{bw} and maternal characteristics.

Finally, the spouse's earnings distribution is modeled similarly using G' discrete values for y^s, and with f_{wi} and f_{bw} in equations (20) and (21) replaced by

$$f_{yi} = \beta_{y0} + \beta_{y1}\overline{W_{ki}^s} + \beta_{y2}Z_i^s + \rho_y(\mu_i^s) \tag{22}$$

$$+\text{higher order terms} + \text{interactions , and}$$

$$f_{by}(g) = \sum_{\eta=1}^{3} \{\theta_{y\eta}[\log(G'-g)]^{\eta}\}, \tag{23}$$

where $\overline{W_{ki}^s}$ is the median local wage rate for men with the same education level, race, and age. Higher order terms and interactions between baseline hazard f_{by} and the characteristics of the spouse are also included in f_{yi}.

3.4 Estimation

As discussed below, our data set includes longitudinal information on the place of residence, the mother's work decision and wage, the husband's earnings (if married), the children's achievement test scores, as well as all covariates described in our model, for each household i in our sample. While we observe for each child-year observation the location of residence k^*_{it} and the mother's work choice h^*_{it}, we only observe the mother's wage draw, w^*_{it} and husband's earnings y^{s*}_{it} at the location k^* that was chosen. We do not observe any wage when the woman is not engaged in market work. It is important to note that our maximum likelihood framework, which accounts for correlated person and family specific unobservable components, controls for endogeneity and self-selection issues in wages, work, and place of residence.

Given the optimal decision rules (10) and (14), and specifying a discrete multinomial distribution for the unobserved heterogeneity vector μ_i, where μ_i can take on J sets of different values, the likelihood contribution for household i is given by

$$
\begin{aligned}
(24) \qquad L_i \;=\; & \sum_{j=1}^{J} \Bigg\{ \Pr(\mu_j) \prod_{t=t_0}^{T_i} \{ \Pr(k_{it} = k^*_{it} | \Omega_{it}, \mu_j) \\
& \left\{ \sum_{g=1}^{G} [\Pr(w_{it} = \bar{w}_{gt} | \Omega_{k^*it}, \mu_j, k^*_{it}) \Pr(h_{it} = h^*_{it} | \Omega_{k^*it}, w_{it}, \mu_j, k^*_{it}) \right\}^{\mathbf{1}(h^*_{it}=0)} \\
& \{ \Pr(w_{it} = w^*_{it} | \Omega_{k^*it}, \mu_j, k^*_{it}) \Pr(h_{it} = h^*_{it} | \Omega_{k^*it}, w^*_{it}, \mu_j, k^*_{it}) \}^{\mathbf{1}(h^*_{it} \neq 0)} \\
& \Pr(y^s_{it} = y^{s*}_{it} | \Omega_{k^*it}, \mu_j, k^*_{it})] \prod_{n=1}^{N_i} \Pr(q_{int} = q^*_{int} | \Omega_{k^*it}, h^*_{it}, y^{s*}_{it}, \mu_j, k^*_{it})] \Bigg\} \Bigg\}.
\end{aligned}
$$

The sample likelihood function then is given by

$$
(25) \qquad L = \prod_{i=1}^{I} L_i.
$$

The size of the choice set we consider (which is the set of all counties in the U.S.), makes direct implementation of the likelihood maximization procedure impractical. To deal with the large size of the choice set we apply a random sampling procedure proposed by McFadden (1978). Let $C = \{1, \ldots, K\}$ be the full choice set, and let $D \subseteq C$ be a subset consisting of $\widetilde{K}$ elements. The sampling method is to select the chosen alternative plus $\widetilde{K} - 1$ non-chosen alternatives randomly drawn from the set C. Let $\pi(D|k^*_{it})$ be the probability that D will be drawn, given the observed choice k^*_{it}. Then

$$
\pi(D|k^*_{it}) = \begin{pmatrix} K-1 \\ \widetilde{K}-1 \end{pmatrix}^{-1}.
$$

As shown by McFadden, the choice probabilities $\Pr(k_{it} = k^*_{it}|\Omega_{it}, \mu_j)$ in equation (24) will then be re-writen as follows:

$$\Pr(k_{it} = k^*_{it}|D, \Omega_{it}, \mu_j) = \frac{\exp\left\{\frac{\widetilde{V}_{k^*}(\Omega_{it},\mu_j)}{b_2} + \ln \pi(D|k^*_{it})\right\}}{\sum\limits_{k' \in D} \exp\left\{\frac{\widetilde{V}_{k'}(\Omega_{it},\mu_j)}{b_2} + \ln \pi(D|k')\right\}}. \tag{26}$$

Consistency of the resulting maximum likelihood estimators relies on the Independence of Irrelevant Alternatives (IIA) property of the error terms in the discrete choice model. It can be demonstrated that the estimator described in equation (26) possesses the IIA property, but only conditional on a given heterogeneity type μ_j. Unconditionally, we cannot rely on the IIA property to guarantee consistency. To assess the properties of the estimator in this case, Liu (2004) conducts a Monte-Carlo study to test the sensitivity of parameter estimates to the size of the sampling set $\widetilde{K}$. For a simulated model with the similar size of the empirical model in our study, he finds that coefficient estimates and the standard errors were nearly identical to each other for choice sample sizes greater than 20. This finding was insensitive to the pseudo R-square value computed from the full choice set model.

4 Data Description

Our primary data source is the Geocode version of the NLSY79 data set and its Child-Mother Supplement. The NLSY79 began in 1979 with a national sample of 12,686 young adults between the ages of 14 and 21. It included a nationally representative sample of 6,111 youths, an over-sample of 5,295 blacks, Hispanics, and economically disadvantaged whites, and a supplemental sample of 1,280 persons in the military in September 1978. Interviews with the military sub-sample were suspended after 1984 and for economically disadvantaged non-Hispanic whites after 1990. Following most of the literature based on NLSY data, we exclude the disadvantaged non-Hispanic white over-samples, because they were selected on the basis of potentially endogenous variables. The Black and Hispanic over-samples are included, and race and ethnicity indicators are included as explanatory variables to capture differences in preferences and opportunities. We restrict our analysis further to cases for which we have mother-child data.

Beginning in 1986, the NLSY-Child collected data on all of the children born to the female

NLSY respondents. The NLSY-Child sample (through 1998) supplies data on children with mothers between the ages of 33 and 40 at the end of 1997. Children under the age of 15 comprise the majority of this sample. The NLSY-Child biennially interviewed both mother and child. The unit of observation in our analysis is a household, in which at least one child is observed between ages 5 and 15 at the time of interview. Children in the data can have up to six possible time-period specific observations. Because of the structure of the data set, we model some child outcomes at ages 6, 8, 10, 12 and 14, and other children's outcomes at ages 5, 7, 9, 11, 13, and 15. Consequently, each period in our empirical model corresponds to a 2-year interval. The NLSY-Child contains a set of cognitive and behavioral assessments of each child at these ages. We exploit information on each child's performance on the Peabody Individual Achievement Test (PIAT) in mathematics. This test is among the most widely used academic achievement assessment instruments that have demonstrably high re-test reliability, concurrent validity, and good psychometric properties (Markwardt, 1989). These scores, nationally normed by age and measured in terms of percentile scores, serve as our child outcome measure.

In our model households make residential location choices based on local labor market conditions, housing prices, geographic preferences, and the quality of the school associated with that location. Regarding the first, we assume the relevant geographic labor market to be the county of residence. Using the Geocodes in the NLSY, matched with US Census 5-Percent Public-Use Microdata Samples (PUMS 5%) county-level data, we measure labor market conditions in the county by the median wage rate by education level, age, race and sex. We distinguish among at least three levels of education (less than high school, high school, more than high school), and compute the median wage by single year of age for each of two race categories (white, non-white). The local labor markets that each mother and father in the sample face are then characterized by the median wage in each locality corresponding to her/his education level, race, and age. That is, the median wage is used as a location specific determinant of the distribution of wages that the individual expects to face in that location.

The school quality indicators we use to measure school inputs at each location are measured at the smallest geographic unit currently available in case of the NLSY, the county level[2]. The choice set from which families choose their residential location in each period therefore is the set of all 3,141 counties in the U.S. The 1990 Census school district special tabulation (see Appendix A) provides comprehensive geographic information on school district boundaries, permitting us to link the sample households to the school districts where they resided, assuming little temporal variation in the boundaries.

Two school quality indicators we consider are per-pupil average school expenditures and average teacher salaries. Both are measured as (weighted) averages across all public schools in each county. We measure these variables relative to the average annual earnings of college educated males age 27-40 in the county. A third school and neighborhood characteristic we consider is the average high school dropout rate in each county. These quality measures are extracted from the School District Data Book (SDDB) and Common Core Data (CCD) files. In particular, CCD is a comprehensive, annual, national statistical database of information concerning all public elementary and secondary schools (approximately 95,000) and school districts (approximately 17,000). The special Geocode data, combined with information from the SDDB also provide us with a measure of local housing costs. In particular, a special tabulation from the SDDB offers information on the median housing rent and the median housing value at the school district level.

While ideally we would have wanted to model the residential location decision at the school attendance area, it is important to point out that our data set contains finer geographic information than is commonly used to analyze the relationship between school inputs and child outcomes (for example, Card and Krueger 1992).

For the estimation of the production function, we group the percentile scores of the children's math tests into 10 discrete cells (5:1-10, 15:11-20, ...,95:91-100), and model the determinants of how a child progresses from one cell to the next highest cell through the logit "hazard" model specified in section 3. The child's score depends upon characteristics of the child such as his/her age and gender, as well as characteristics of the mother such as her age, marital status, schooling level, part-time and full-time work status, AFQT score, and the family's income after deducting a measure of the local cost of housing. This production

[2] Individual observations in the standard Geocode version of the NLSY have county Federal Information Processing Standards (FIPS) codes as smallest geographic indicator of their residential location.

function also depends upon characteristics of the school district where the family resides, explicitly the high school dropout rate, average teacher salary as a fraction of the average annual earnings of college educated persons aged 27 to 40, and per child school expenditures as a fraction of the average earnings of individuals in the county. Monte-Carlo estimates reported by Gilleskie and Mroz (2004) indicate that little information is lost by using a discretized outcome rather than the more continuous one.

To describe the wage offers available to the individual women in each locality k at t, w_{it}, we use a discrete distribution with 10 points of support (i.e. $w_{it} = \bar{w}_{1t}, \bar{w}_{2t}, \ldots, \bar{w}_{10t}$) and model the probability that each woman has an offered wage from each of the 10 categories. As in the child production function model, we use a "hazard" model to describe transitions across categories. For arguments to these hazard models, we allow for detailed interactions between individual level covariates and local area wages. In particular, for each of three levels of education (less than high school, high school, and more than high school), we compute the median wage by single year of age for each of two race categories (white, non-white). For each woman, we assign to her in each locality the median wage corresponding to her education level, race, and age. We use these median wages as explanatory variables for the hazard model of the wage offer distribution. We also allow for separate education level effects and effects of the mother's AFQT score on the probabilities of each of these discrete wage offers.

A specification with 10 points of support was similarly adopted for the spouse's earnings equation. In addition to local median male yearly earnings in the county, the equation includes the father's race, and his education level as explanatory variables.

Table 1A and Table 1B contain summary statistics for the sample data on parents and children respectively. An interesting phenomenon is that the mean education levels and hourly wage rates of the mothers in our sample increased over years. This suggests that the existence of self selection on women's decision to enter parenthood at young ages. Table 2 contains the data on school district characteristics broken down into 24 different study regions, which are used to conduct a series of policy simulation exercises. These study regions of residence are defined first by four geographic breakdowns (Northeast, Midwest, South, and West). Within each geographic area we make a further division by three levels of the school dropout rate within each region. We then divide each of these 12 geographic-dropout rate regions into two cells on the basis of the level of per student school expenditures in the school districts. As shown by Table 2, there is considerable regional variability in average values of these school district characteristics.

5 Estimation Results

I. Point Estimates of Utility and Production Function Parameters.

Tables 3A, 3B, 3C, 3D, and 3E contain parameter estimates for the production function, utility function, wage offer, and heterogeneity distribution parameters. The estimates in this table assume that there is unobserved heterogeneity that can influence the production function, the "reserve" consumption in the utility function, the utility costs of full- and part-time work, the mother's wage offer distribution, and the earnings distribution for the mother's spouse. We specify four points of support for this unobserved heterogeneity. We permit up to a third order polynomial in the value of the heterogeneity to influence all of the model components, so this is close to a non-parametric specification of the unobservable heterogeneity in this semi-parametric model. We model this heterogeneity as constant throughout the years of observation on the mother's family.

Given the complex interactions between most of the covariates in the economic model, it is quite difficult to give a simple interpretation to each of the point estimates. Consider, for example, the impact of the mother having a high school degree on the wage hazard model presented in Table 3C. The interpretation of this parameter is how a woman being a high school graduate instead of a dropout influences the propensity to move at a given wage value to a higher wage category holding constant the median wage paid to women in her locality with the same original education level. The complete effect of having a high school degree on her wage level is much more complicated than this. To examine the effect of having a high school diploma instead of a being a dropout, one would need to consider this impact in conjunction with how the higher median wages offered to high school graduates affects the wage offer distribution. Given the various interactions between the explanatory variables incorporated in the model, it is difficult to put any simple interpretation on the parameter estimates. Therefore, we focus on estimated marginal effects of these explanatory variables using simulation methods described in the following subsection.

II. Simulations of Effects

A. The Production Function, Wage and Spouse Earnings Marginal Effects

In Table 4A, we estimate the production function using OLS with a series of specifications. The coefficient estimates of school district characteristics and maternal work decisions are fairly robust. To interpret the implications of the model estimates we simulate outcomes using the estimated parameters. Table 4B presents three sets of marginal, ceteris paribus, estimates of the impacts of characteristics on the child's PIAT Math test score. These simulated estimates abstract from any parental location or work decisions that might change in response to the changes in the covariates. The first column contains OLS estimates, and it can be interpreted as a standard OLS regression. Columns 2 and 3 contain estimates based upon the "hazard" model for the child scores with controls for unobserved heterogeneity. We simulated how expected test scores would change in response to varying explanatory variables one at a time, normalized to a unit change in the characteristic as is the case with the OLS estimates.

The estimates used for the simulation results in Column 2 only incorporate the functional form used for the production function used in this analysis. While they do not incorporate any controls for selection or endogeneity, they use exactly the same form of the heterogeneity distribution as we used in the structural model. The simulations in Column 2 correspond directly to the OLS estimates, and the estimates underlying Column 2 were obtained by estimating the production function separately from the rest of the economic model. The estimates used to define the marginal effects in Column 3 were estimated as part of the structural model that incorporated the endogeneity of the location decisions and the endogeneity of the mothers' hours of work decisions. The standard errors of the marginal effects are calculated using a parametric bootstrap (100 replications) simulated using the estimated covariance matrix of the parameter estimators.

A comparison of the OLS estimates in Column 1 and the simulated derivatives in Column 2 isolates the impact of using a more flexible functional form for the estimation of the child outcome production function. For the most part, the estimated derivatives are quite close for these two models, with most of the absolute differences being less than one standard error of the OLS model. The standard errors in Column 2 are almost always smaller than those in Column 1, suggesting that the nonlinear production function provides more accurate estimates than does OLS.

The estimates of the marginal effects presented in the last column in Table 4B come from an estimation model based exactly on the structural model with controls for endogeneity of the parental choices including the location decisions. After controlling for the endogeneity of these variables, most of the impacts of the exogenous variables are fairly close to those from the two models that do not control for the endogeneity of the production function inputs. The impacts of the potentially endogenous variables, however, do change significantly. The negative impact of the local dropout rate on expected child outcome diminishes by more than 74 percent. The negative effect of higher teacher salaries falls by about 20 percent, and the impact of expenditures per pupil falls by 44 percent. The estimated effects of school district characteristics on a child's test score, diminish considerably after using the structural model to control for endogeneity of the location and work decisions.

The impact of the family changing residing counties becomes negative and statistically significant after controlling for the endogenous behaviors, while without the endogeneity controls it appears that a move significantly increases a child's test score. Specifically, without endogeneity controls, the estimates using OLS and nonlinear functional form suggest that moving improves child test score by 1.1 and 1.6 percent respectively. The estimates obtained from the full structural model with endogeneity controls indicate that moving actually causes a 5.8 percent drop in child scores.

The positive impacts of the mother working part time or full-time (instead of not working) both become negative after controlling for the endogeneity of these parental decisions. Without accounting for endogeneity, it would appear that a mother's working full-time could increase her child's percentile score by 2.2 points. After controlling for endogeneity the estimated impact of a mother working full-time implies an expected decline in the child's percentile score of 1.2 points. The estimates from the structural model with endogeneity controls provide significantly different estimated impacts of the effects of school district characteristics and parental work decisions on children's expected test scores.

Table 4C contains the marginal effects of schooling, race, and the mother's AFQT score on the mothers offered wage distribution. The impact of being non-white on offered wages increases after controlling for the selection into the labor force, while the magnitude of the mothers AFQT score falls by over 20 percent. Table 4D contains the marginal effects on the fathers earnings. In the full model, the effects on earnings of higher levels of schooling are slightly lower than in the model without the full set of controls. The estimated effect of race is also smaller in the complete model. In general these estimates are less sensitive to the endogeneity controls than those for the mothers offered wage.

B. Simulations of Changes in Location Patterns

Using the point estimates from the full structural model and the optimal decision rules, we are able to simulate a set of residential location choice, maternal work decision, and child outcome for an observed household at each period. We then map the simulated choices about residing counties into 24 study regions; these are defined in the data description section. The simulated frequencies in the second column of Table 5 demonstrate that the structural model fits the observed distribution of residential location decisions, as categorized in terms of the 24 different regions, quite well. A primary reason why this model appears to perform so well is that we treat the initial residential location decision (when the child was either age 5 or age 6) as exogenous. This, in conjunction with the relatively large utility costs of migration that we estimate makes it relatively easy to fit the marginal distribution of residences. The third and fourth columns of Table 5 represent actual and simulated average employment rates among 24 study regions respectively. The sample and simulated average child test scores are presented in the fifth and sixth columns of Table 5 provide. For most regions, the model predicts average employment decisions and child scores fairly accurately.

Table 6 presents the simulated overall average moving rate, employment percentage, and child scores, given three hypothetical scenarios. In the third column, we increase per pupil school expenditures in each county by 25%. The actual average expenditure levels for each region are presented in Table 2. It appears that parents are re-optimizing their migration and maternal employment decisions, given the exogenous changes to school district characteristics. After increasing the expenditure levels for each of the low-expenditure regions, there is a reduction in the migration rate and an increase in the employment rate. Due to the dominating effect of declined maternal time inputs, the simulated average child test scores are lower than the baseline level by 0.4 percentage point, although the magnitude of induced migrations is fairly small. The fourth column represents a hypothetical 25% increase of high school dropout rates in each county. We observe reduced moving rate and increased employment percentage. However, due to the overall rising dropout rates, we find that the resulting average child test scores drop by 1.1 percentage points. Finally, the fifth column represents a 25% across the board increase in wage rates. We find the scale of the migration rate reduction is similar to those of the previous two experiments, while the rise in employment rate is significantly greater than the results of in columns 3 and 4. Consequently, the average child test score falls between the corresponding results associated with increasing expenditure per pupil and increase dropout rates.

6 Discussion about Model Identification

In order to discuss identification issues, it is useful to characterize the structural model estimated in this paper as consisting of three main components: (1) the child outcome equation as a function of two sets of endogenous inputs, maternal employment and school inputs, (2) the mother's employment choice equation and (3) the school input equation, which corresponds to the residential location decision. Identification of the effect of both sets of inputs on child outcomes relies to a considerable extent on our assumption that a household's place of residence when the first child reached age 5 or 6 is exogenous to the subsequent residence locations, work decisions, and child outcomes. Because moves are costly, with costs depending on the geographical distance between the current and new location, local labor market conditions and school characteristics in any given period will depend on the family's initial location. Moreover, there are stochastic, age-varying exogenous changes over time in the distribution of local male and female labor market conditions in each location, and these interact with the family's initial location to generate possibly substantial variation in current labor market and school conditions. For example, those who face relatively improved labor market conditions in their initial location of residence will be less likely to move. In addition to cross-sectional variation in the family's initial location and time-variation in geographic labor market conditions for men and women, the treatment of time-varying fertility and marital status decisions as exogenous provides an additional sources of identification. In future work, however, it would be important to explore relaxing these latter exogeneity assumptions.

Regarding the key identification condition, the exogeneity of the initial local location when the oldest child reaches age 5 or 6, we considered two alternatives to this assumption but concluded that neither was preferable to the straightforward one we use in this study. First, we considered using a more "equilibrium" approach where the place of "first residence" in the model would be modeled as depending only on comparisons among the 3141 U.S. counties using county data from 1980 without explicitly conditioning on the woman's initial location and without adjustments for moving costs. That is, we model $\Pr(k_{i1} = k^*_{i1} | S_{1-r}, Z_{i,1-r}, \mu_i)$, where period $1 - r$ corresponds to the calendar year 1980 and S_t represents the set of all $S_{kt}, k = 1, \ldots, K$. This model would be driven by interactions between individual-specific exogenous variables such as education, age, and family background, and the location-specific measures of economic, social, and school opportunities. Unobserved heterogeneity correlated with the study period's unobserved heterogeneities would provide a link to control for the mother's endogenous initial location. The fact that county characteristics do change between 1980 and the start of the collection of the Child Supplement data in 1986 means that this approach could provide some exogenous variation in place of location that would not be perfectly related to the characteristics of the place of residence when the oldest child was age 5 or 6. The drawback of this approach is that such a model would likely provide a poor fit to the actual distribution of places of residence as it would not exploit the inertia in migration processes.

Second, we considered using the mother's place of residence at the time of the initial 1979 interview as an exogenous initial condition. One could then model sequentially annual migration decisions with potential moving costs related to the prior place of residence in an approximate form that mimics the dynamic decision process, yielding $\Pr(k_{i1} = k^*_{i1} | k^*_{i,-r}, S_{1-r}, S_{2-r}, \ldots, S_1, Z_{i,1-r}, Z_{i,2-r}, \ldots, Z_{i,1}, \mu_i)$. This would mitigate the main problem with the first alternative approach, as we would condition on the location in 1979 and the inertia would be explicitly incorporated. However, all the annual migration decisions after 1979 would depend on the timing of fertility outcomes and marital events. It is not clear that controlling for the endogeneity of the "first residence" location should be more important than controlling for the endogeneity of marital status and the timing of fertility. Rather than use a piecemeal approach to attempt to solve these extensive "initial condition" problems, we postpone dealing with them to future research on this topic.

Panel data relationships like those examined in this study can implicitly provide many additional identification conditions than one might infer by simply counting the number of contemporaneous exogenous variables (e.g. instruments) excluded from the structural equation of interest. There are two primary reasons for this. First, consider the case examined by Bhargava (1991) when one is willing to impose structural parameter stability over the time dimension. As Bhargava (1991) demonstrated, every lag of each instrument could have a separate effect on the "contemporaneous" value of the endogenous explanatory variables. This time dimension for the exogenous time varying instruments creates a multiplicity of "instruments" associated with each "exclusion restriction," resulting in significantly more variables to control for endogeneity. He demonstrates that over-identification can be obtained under quite weak conditions. In this study the exogenous time-varying variables include the mothers', children's, and fathers' ages, marital status, family composition, and age induced variations in the distribution of local labor marker conditions and school characteristics, and many of these factors do not vary deterministically through time. These time-varying exogenous factors interact with each observation's initial place of residence to generate numerous exclusion restrictions.

A second source of additional identification arises in the context of dynamic nonlinear models. Mroz and Surette (1998) discuss this in more detail. Their discussion exploits the fact that variations in the time-ordering of the exogenous variables provide even higher degrees of over-identification than would be obtained by a simple reference to Bhargava's (1991) observation as discussed above. The basic idea underlying their additional identification argument is that in dynamic nonlinear models of the type used here the impact a lagged exogenous variable, say at $(t-s)$, on a current endogenous variable, at t, depends crucially on the entire time series of all exogenous variables before t-s and between t-s and t. In the presence of endogenous determinants that evolve over time, the impact of an exogenous variable dated $(t-s)$ on an endogenous determinant at t will vary depending on the values of the same exogenous variable dated at $(t-s-1)$ and at $(t-1)$. Each additional exogenous variable will typically increase dramatically the degree of identification.

By using an explicit sequential dynamic modeling framework, one can incorporate all such interactions that depend on the precise timing and sequencing of the values of the varying exogenous variables. The maximum likelihood approach we use here automatically incorporates these interactions among the time series properties of the sets of exogenous variables. They do so efficiently, without one having to resort to including numerous time-varying interactions of the exogenous variables in an arbitrary fashion, as would be the case with a more static instrumental variables approach.

It is important to note that this is the only study of the determinants of child test score outcomes we are aware of to consider the endogeneity of both mother's work behavior and school characteristics. Even though we treat as exogenous the mother's place of residence when the oldest child is age 5 or 6, contrasts of the empirical results with and without the endogeneity controls are striking. Clearly the results reported here will not provide the last word on the estimation of the effects of mother's work and school characteristics on child outcomes. But the framework used here does lay a strong foundation that can be extended to incorporate many additional, potentially important modeling issues and help us to understand better this important topic.

7 Conclusions

Estimating the educational production function as part of a structural model provides significantly different estimates of the production process. For the most part, the impacts of the school district characteristics diminish by factors of 2 to 4 after controlling for the fact that families may be choosing where to live because of the school district characteristics and labor market opportunities. We also find that the impacts on child outcomes of having moved and working full-time (as opposed to not working) to change signs and remain statistically significant after controlling for the possible endogeneity of these decisions. One interpretation of these changes in the estimated production function impacts is that families whose children would anyways perform quite well tend to choose to live in school districts with the highest levels of productive inputs and work more. This is a standard endogeneity of inputs argument. Inputs to the educational production process remain significant determinants of the child outcomes, but they are much smaller than is implied by estimation methods that do not allow for possible endogeneity biases.

When we turn to the estimates of the overall effects of changes in characteristics on child outcomes, a somewhat different story emerges. Since parents can re-optimize by choosing different school districts and hours of work, many of the benefits (or detriments) to students from changing school district characteristics end up having only minor impacts on the child test scores. While we have only considered a small number of simulations, these results do suggest that it is important to recognize that parents' decisions about work and school districts can offset improvements in the school environments. Additionally, parents can substitute into better school districts in order to make up for the fact that full-time work can leave the parents with less time to spend helping their children learn. The presence of such substitution possibilities might help explain some of the disparate estimates of the educational production function found in the literature.

Table 1A
Summary Statistics for Parents

Variable	Mean (Standard Deviation)							
	1986	**1988**	**1990**	**1992**	**1994**	**1996**	**1998**	**Overall**
Age of mother	26.540	28.543	30.060	31.800	33.600	35.452	37.340	32.594
	(1.941)	(1.938)	(2.116)	(2.178)	(2.188)	(2.183)	(2.186)	(4.003)
Married	0.509	0.507	0.539	0.568	0.584	0.602	0.620	0.570
	(0.500)	(0.500)	(0.499)	(0.495)	(0.493)	(0.490)	(0.486)	(0.495)
High school	0.516	0.525	0.534	0.525	0.494	0.468	0.447	0.497
	(0.500)	(0.500)	(0.499)	(0.500)	(0.500)	(0.499)	(0.497)	(0.500)
More than high school	0.159	0.182	0.243	0.293	0.342	0.393	0.441	0.313
	(0.366)	(0.386)	(0.429)	(0.455)	(0.475)	(0.489)	(0.497)	(0.464)
Non-white	0.647	0.646	0.617	0.576	0.564	0.538	0.505	0.575
	(0.478)	(0.478)	(0.486)	(0.494)	(0.496)	(0.499)	(0.500)	(0.494)
Mother’s AFQT score	0.354	0.357	0.403	0.433	0.448	0.472	0.495	0.433
	(0.267)	(0.268)	(0.271)	(0.273)	(0.276)	(0.279)	(0.283)	(0.279)
Yearly housing cost(1)	0.499	0.500	0.504	0.507	0.508	0.511	0.518	0.508
	(0.137)	(0.136)	(0.135)	(0.136)	(0.138)	(0.140)	(0.140)	(0.138)
Move	0.135	0.164	0.143	0.067	0.152	0.091	0.086	0.115
	(0.342)	(0.371)	(0.350)	(0.250)	(0.359)	(0.288)	(0.281)	(0.319)
Net income(1)	1.215	1.375	1.741	2.003	2.117	2.548	2.981	2.105
	(1.878)	(1.810)	(2.033)	(2.568)	(2.432)	(2.993)	(3.168)	(2.610)
Mother’s hourly wage rate(2)	0.661	0.715	0.781	0.799	0.909	0.927	1.032	0.858
	(0.422)	(0.768)	(0.606)	(0.528)	(0.862)	(0.967)	(1.038)	(0.805)
Father’s yearly earnings(1)	1.813	1.963	2.266	2.342	2.543	2.844	3.236	2.541
	(1.409)	(1.548)	(1.752)	(1.959)	(2.099)	(2.312)	(2.555)	(2.122)
Local median hourly wage rate for women(2)	0.809	0.829	0.923	0.977	1.021	1.078	1.102	0.990
	(0.393)	(0.347)	(0.381)	(0.419)	(0.407)	(0.430)	(0.428)	(0.419)
Part-time work(3)	0.353	0.323	0.329	0.312	0.289	0.271	0.283	0.304
	(0.478)	(0.468)	(0.470)	(0.464)	(0.453)	(0.445)	(0.451)	(0.460)
Full-time work(3)	0.298	0.387	0.416	0.425	0.462	0.499	0.529	0.444
	(0.457)	(0.487)	(0.493)	(0.495)	(0.499)	(0.500)	(0.499)	(0.497)
# of young children in HH (0-5)	1.134	0.710	0.652	0.647	0.546	0.437	0.379	0.604
	(0.945)	(0.803)	(0.781)	(0.784)	(0.738)	(0.691)	(0.650)	(0.788)
# of old children in HH (6-17)	1.127	1.613	1.641	1.828	1.946	2.002	2.011	1.793
	(0.933)	(0.990)	(0.981)	(0.982)	(0.937)	(0.945)	(0.960)	(0.996)
Sample size	1075	992	1409	1688	1858	1859	1739	10620

Notes:
(1) In $10,000's of 1990 dollars
(2) In $10's of 1990 dollars
(3) See appendix for the definition of part-time and full-time

Table 1B
Summary Statistics for Children

Variable	Mean (Standard Deviation)							
	1986	**1988**	**1990**	**1992**	**1994**	**1996**	**1998**	**Overall**
Age of child	7.608 (2.008)	9.072 (2.323)	9.389 (2.553)	9.805 (2.656)	9.938 (2.617)	10.063 (2.615)	10.226 (2.553)	7.608 (2.008)
Non-white	0.663 (0.473)	0.656 (0.475)	0.629 (0.483)	0.588 (0.492)	0.570 (0.495)	0.543 (0.498)	0.506 (0.500)	0.663 (0.473)
Boy	0.528 (0.499)	0.515 (0.500)	0.504 (0.500)	0.499 (0.500)	0.498 (0.500)	0.502 (0.500)	0.509 (0.500)	0.528 (0.499)
Dropout rate	0.064 (0.011)	0.063 (0.011)	0.063 (0.011)	0.063 (0.011)	0.063 (0.011)	0.062 (0.010)	0.062 (0.010)	0.064 (0.011)
Expenditure per pupil(1)	0.126 (0.039)	0.124 (0.038)	0.124 (0.038)	0.124 (0.038)	0.124 (0.039)	0.124 (0.040)	0.123 (0.040)	0.126 (0.039)
Teacher salary(1)	0.760 (0.187)	0.766 (0.181)	0.771 (0.195)	0.769 (0.190)	0.772 (0.185)	0.774 (0.187)	0.770 (0.189)	0.760 (0.187)
PIAT Math score	0.451 (0.259)	0.429 (0.252)	0.445 (0.258)	0.461 (0.261)	0.477 (0.266)	0.511 (0.272)	0.524 (0.271)	0.491 (0.259)
Sample size	1493	1796	2515	3056	3300	3202	2873	18235

Notes:
(1) Measured as a proportion of yearly income of college-graduated prime-aged males. See appendix for details.

Table 2
Study Regions Definitions and
School and other Location Characteristics

Census region	Dropout divisions	Expend-iture per pupil	Median dropout rate		Median teacher salary(1)	Median expenditure per pupil(1)		Median Annual Housing Cost
			Raw data	Increased for simulation		Raw data	Increased expend-iture for simulation	
Northeast	Low	Low	4.75	5.93	0.780	0.143	0.179	0.997
		High	4.82	6.03	0.867	0.205	0.257	0.898
	Middle	Low	5.46	6.82	0.840	0.124	0.155	0.850
		High	5.52	6.89	0.840	0.179	0.224	0.715
	High	Low	6.44	8.05	0.827	0.136	0.170	0.786
		High	6.88	8.60	0.976	0.181	0.226	0.676
Midwest	Low	Low	4.60	5.75	0.873	0.059	0.074	0.519
		High	4.56	5.70	0.863	0.130	0.163	0.558
	Middle	Low	5.47	6.84	0.782	0.082	0.102	0.588
		High	5.60	7.00	0.856	0.133	0.166	0.594
	High	Low	6.51	8.14	0.755	0.088	0.111	0.602
		High	6.80	8.50	0.758	0.167	0.209	0.715
South	Low	Low	5.65	7.06	0.708	0.094	0.117	0.543
		High	5.37	6.72	0.792	0.155	0.194	0.636
	Middle	Low	6.28	7.85	0.718	0.093	0.117	0.537
		High	6.43	8.04	0.773	0.150	0.187	0.655
	High	Low	7.74	9.68	0.749	0.102	0.128	0.606
		High	7.75	9.69	0.774	0.152	0.190	0.594
West	Low	Low	5.46	6.83	0.792	0.097	0.121	0.861
		High	5.56	6.95	0.959	0.146	0.182	0.652
	Middle	Low	6.51	8.14	0.817	0.094	0.118	0.856
		High	6.76	8.45	0.885	0.131	0.163	0.650
	High	Low	7.48	9.35	0.803	0.115	0.144	0.937
		High	8.18	10.22	0.847	0.136	0.170	0.507
Overall Sample Mean			6.26	7.83	0.801	0.124	0.155	0.664

Note:
(1) Measured as a proportion of yearly income of college-graduated prime-aged males. See appendix for details.
(2) In $10,000’s of 1990 dollars

Table 3A

Parameter Estimates from the Full Model
Production Function Parameters

Variable	Estimate	Std. Err.
Intercept	-0.430	0.337
Age of mom (in 10 years)	0.058	0.345
child age (in 10 years)	0.021	0.135
child age squared	-0.064	0.020
Married	0.109	0.016
Mother education high school	0.147	0.018
Mother education more than HS	0.353	0.023
Non-white	-0.312	0.016
Boy	0.042	0.013
Dropout rate	-0.152	3.757
Expenditure per pupil	1.555	0.915
Teacher salary	-0.135	0.207
Move	-0.330	0.022
Mother part-time work	0.178	0.070
Mother full-time work	0.200	0.071
Dropout rate × Log(*P*-*p*)	-0.231	1.260
Expense per pupil × Log(*P*-*p*)	-0.229	0.327
Teacher salary × Log(*P*-*p*)	-0.010	0.082
Mother part-time × Log(*P*-*p*)	-0.147	0.035
Mother full-time × Log(*P*-*p*)	-0.210	0.036
Dropout rate× Child age	-0.784	2.895
Expense per pupil × Child age	-0.675	0.747
Teacher salary×Child age	-0.120	0.151
Part-time × Child age	0.030	0.036
Full-time × Child age	0.111	0.042
Mother's AFQT score	0.347	0.086
Mother's AFQT × Log(*P*-*p*)	0.539	0.053
Loading on 1st order heterogeneity factor	-0.351	0.211
Loading on 2nd order heterogeneity factor	-1.845	0.516
Loading on 3rd order heterogeneity factor	0.949	0.327
1st order of baseline hazard [log(*P*-*p*)]	4.738	0.148
2nd order of baseline hazard [log(*P*-*p*)]^2	9.140	0.063
3rd order of baseline hazard [log(*P*-*p*)]^3	6.873	0.029
4th order of baseline hazard [log(*P*-*p*)]^4	1.551	0.012

Note:
P is the highest level of discretized scores and *p* is any given discretized score level.

Table 3B

Parameter Estimates from the Full Model

Utility Function Parameters

Variable	Estimates	Std. Err
Intercept in reserve	2.765	0.490
1st order discrete factor loading in reserve	-3.317	1.355
2nd order discrete factor loading in reserve	1.744	1.094
Power of consumption (γ_1)	0.003	0.002
Married (α_0)	0.163	0.061
Any work (α_1)	0.004	0.003
Relative power of mother's AFQT score (α_3)	1.005	0.757
Scale on child's score / mother's AFQT score (α_2)	2.432	5.817
Reserve child's score / mother's AFQT score (γ_2)	1.709	4.251
Power of child's score / mother's AFQT score (γ_3)	0.136	0.530
Intercept in part-time leisure	0.427	0.338
1st order discrete factor loading in part-time leisure	-1.499	0.639
2nd order discrete factor loading in part-time leisure	0.169	0.964
3rd order discrete factor loading in part-time leisure	1.460	0.843
# of young children (0-5) in part-time leisure	0.094	0.030
# of old children (6-17) in part-time leisure	-0.003	0.006
Married in $f_{4a}(.)$	0.008	0.014
Intercept in full-time leisure	-2.088	0.155
1st order discrete factor loading in full-time leisure	-0.245	1.075
2nd order discrete factor loading in full-time leisure	3.231	0.984
3rd order discrete factor loading in full-time leisure	-1.079	25.172
# of young children (0-5) in full-time leisure	0.542	0.057
# of old children (6-17) in full-time leisure	0.170	0.032
Married in $f_{4b}(.)$	-0.104	0.100
Moving psychic cost (any move)	-3.739	1.147
Additional moving psychic cost across states	-1.629	0.481
Additional moving psychic cost across census regions	-0.570	0.170
Inverse of parameter **b** Gumbel error on working choice	3.227	0.270
Inverse of parameter **b** Gumbel error on location choice	1.969	0.578
Dummy for census region – Midwest(1)	-0.369	0.145
Dummy for census region – South	-0.224	0.108
Dummy for census region – West	-0.073	0.101

Note: Northeast is the base region.

Table 3C
Parameter Estimates from the Full Model
Mother's Wage Rate Function Parameters

Variable	Estimates	Std. Err.
Intercept	-1.859	0.135
Local wage rate	0.912	0.102
Local wage rate squared	-0.236	0.029
Local wage rate × discrete factor	0.050	0.122
High school (mother)	0.328	0.033
More than high school (mother)	0.548	0.043
Non-white	0.238	0.023
Mother's AFQT score	0.405	0.200
Mother's AFQT × Log(*G*-*g*#)	-1.633	0.151
Loading on 1st order heterogeneity factor	-1.232	0.223
Loading on 2nd order heterogeneity factor	-0.223	0.211
1st order of baseline hazard [log(*G*-*g*)]	0.058	0.058
2nd order of baseline hazard [log(*G*-*g*)]^2	-0.664	0.132
3rd order of baseline hazard [log(*G*-*g*)]^3	-0.247	0.076

Note: *G* is the highest level of discretized wage rates; *g* is any given discretized wage rate level.

Table 3D
Parameter Estimates from the Full Model
Father's Earning Function Parameters

Variable	Estimates	Std. Err.
Intercept	-1.658	0.102
Local median earnings	0.040	0.003
Local median earnings × discrete factor	0.000	0.000
High school (father)	-0.019	0.004
More than high school (father)	0.352	0.039
Non-white	0.809	0.046
Loading on 1st order discrete factor	-0.618	0.033
Loading on 2nd order discrete factor	-3.650	0.221
1st order of baseline hazard [log(*G'*-*g'*)]	-0.402	0.181
2nd order of baseline hazard [log(*G'*-*g'*)]^2	-1.357	0.207
3rd order of baseline hazard [log(*G'*-*g'*)]^3	-0.964	0.213

Note: *G'* is the highest level of discretized earnings; *g* is any given discretized earning level.

Table 3E

Parameter Estimates from the Full Model:
Those Defining the Probabilities for the
Discrete Heterogeneity Points

Variable	Estimates	Std. Err.
Probability parameter at 0	1.552	0.129
Probability parameter at 1/3	-1.932	0.130
Probability parameter at 2/3	3.777	0.163

Table 3F

Probability Distribution of Heterogeneity Types

Heterogeneity factor	Probability
0	0.201
1/3	0.277
2/3	0.369
1	0.153

Table 4A

Production Function Estimates of Marginal Effects with Different OLS Specifications

Variable	OLS 1	2	3	4	5	6
Age of mother	–0.027***	–0.020***	–0.025***	–0.018***	–0.026***	–0.019***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Age of child	0.020***	0.017**	0.016**	0.014**	0.019**	0.016**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Married	0.025***	0.010**	0.028***	0.012***	0.026***	0.011***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
High school (mother)	0.036***	0.035***	0.039***	0.036***	0.035***	0.034***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
More than high school	0.063***	0.059***	0.066***	0.060***	0.062***	0.058***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Non-white	–0.058***	–0.056***	–0.056***	–0.054***	–0.056***	–0.054***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Boy	–0.001	–0.001	–0.001	–0.001	–0.001	–0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Dropout rate	–	–	–0.802***	–0.774***	–0.790***	–0.759***
	–	–	(0.180)	(0.179)	(0.180)	(0.179)
Teacher salary	–	–	–0.013	–0.009	–0.015	–0.010
	–	–	(0.010)	(0.010)	(0.010)	(0.010)
Expenditure per pupil	–	–	0.158***	0.146***	0.165***	0.151***
	–	–	(0.047)	(0.047)	(0.047)	(0.047)
Move	0.011**	0.013**	0.011**	0.014**	0.011**	0.013**
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Part-time work	0.027***	0.024***	–	–	0.027***	0.024***
	(0.005)	(0.005)	–	–	(0.005)	(0.005)
Full-time work	0.021***	0.013***	–	–	0.022***	0.014***
	(0.005)	(0.005)	–	–	(0.005)	(0.005)
Net income	–	0.007***	–	0.008***	–	0.007***
	–	(0.001)	–	(0.001)	–	(0.001)
AFQT score (mom’s)	0.273***	0.257***	0.274***	0.257***	0.270***	0.255***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Intercept	0.255***	0.280***	0.308***	0.329***	0.298***	0.318***
	(0.017)	(0.017)	(0.022)	(0.022)	(0.022)	(0.023)

Note: Standard errors are in parentheses.
* Statistically significantly at the .90 level.
** Statistically significantly at the .95 level.
*** Statistically significantly at the .99 level.

Table 4B

Production Function Estimates of
Marginal Effects with Comparisons to OLS and
"Hazard" Models without Endogeneity Controls

Variable	OLS	Marginal Effects: Production function only (no endogeneity Controls)	Marginal Effects: Full model (with selection and endogeneity controls)
Exogenous Variables			
Age of mother	–0.026***	–0.053***	–0.063***
	(0.005)	(0.009)	(0.005)
Age of child	0.019**	0.024***	0.019***
	(0.007)	(0.008)	(0.005)
Married	0.026***	0.015***	0.019***
	(0.004)	(0.004)	(0.005)
High school (mother)	0.035***	0.030***	0.026***
	(0.005)	(0.004)	(0.003)
More than high school	0.062***	0.053***	0.062***
	(0.006)	(0.005)	(0.004)
Non-white	–0.056***	–0.048***	–0.056***
	(0.005)	(0.004)	(0.003)
Boy	–0.001	0.013***	0.007***
	(0.004)	(0.003)	(0.002)
AFQT score (mom's)	0.270***	0.270***	0.332***
	(0.009)	(0.007)	(0.008)
Intercept	0.298***	–	–
	(0.022)	–	–
Endogenous Variables			
Dropout rate	–0.790***	–0.752***	–0.216*
	(0.180)	(0.174)	(0.111)
Teacher salary	–0.015	–0.011	–0.012***
	(0.010)	(0.012)	(0.004)
Expenditure per pupil	0.165***	0.186***	0.091***
	(0.047)	(0.045)	(0.028)
Move	0.011**	0.016***	–0.058***
	(0.006)	(0.006)	(0.004)
Part-time work	0.027***	0.036***	–0.009***
	(0.005)	(0.005)	(0.003)
Full-time work	0.022***	0.031***	–0.012***
	(0.005)	(0.005)	(0.004)

Note: Standard errors are in parentheses.
* Statistically significantly at the .90 level.
** Statistically significantly at the .95 level.
*** Statistically significantly at the .99 level.

Table 4C

Mother's Wage Equation Estimates of
Marginal Effects with Comparisons to OLS and
"Hazard" Models without Endogeneity Controls

		Marginal Effects	
Variable	**OLS**	**Wage function only (no endogeneity Controls)**	**Full model (with selection and endogeneity controls)**
Local median wage	0.158***	0.204***	0.224***
	(0.024)	(0.018)	(0.024)
High school	0.029	0.122***	0.141***
	(0.030)	(0.015)	(0.039)
More than high school	0.196***	0.220***	0.212***
	(0.035)	(0.021)	(0.021)
Non-white	0.114***	0.129***	0.302***
	(0.022)	(0.012)	(0.022)
Mother AFQT score	0.652***	0.606***	0.475***
	(0.045)	(0.026)	(0.027)
Intercept	0.245***	–	–
	(0.037)	–	–

Note: Standard errors are in parentheses.
* Statistically significantly at the .90 level.
** Statistically significantly at the .95 level.
*** Statistically significantly at the .99 level.

Table 4D

Father’s Earning Equation Estimates of
Marginal Effects with Comparisons to OLS and
“Hazard” Models without Endogeneity Controls

		Marginal Effects	
Variable	**OLS**	**Earning function only (no endogeneity Controls)**	**Full model (with selection and endogeneity controls)**
Local yearly median earnings	0.027***	0.035***	0.029***
	(0.002)	(0.002)	(0.002)
High school	0.371***	0.520***	0.495***
	(0.071)	(0.034)	(0.038)
More than high school	1.271***	1.291***	1.134***
	(0.080)	(0.049)	(0.048)
Non-white	–0.844***	–0.755***	–0.545***
	(0.052)	(0.032)	(0.035)
Intercept	1.598***	–	–
	(0.083)	–	–

Note: Standard errors are in parentheses.
* Statistically significantly at the .90 level.
** Statistically significantly at the .95 level.
*** Statistically significantly at the .99 level.

Table 5

Goodness of fit: Simulated Region of Residence,
Employment Rate, and PIAT Math Score

			Location		**Employment**		**Math Score**	
			Raw data	**Simulated**	**Raw data**	**Simulated**	**Raw data**	**Simulated**
All regions			-	-	75.1445	75.6247	0.4910	0.4904
Census region	**Dropout divisions**	**Expenditure per pupil**						
Northeast	Low	Low	1.460	1.508	80.000	68.021	0.590	0.548
		High	1.412	1.403	76.667	74.329	0.531	0.498
	Middle	Low	1.733	1.846	75.543	75.628	0.555	0.517
		High	3.107	3.206	69.394	78.159	0.541	0.514
	High	Low	2.580	2.514	63.504	73.597	0.444	0.479
		High	3.183	3.162	64.201	74.967	0.449	0.449
Midwest	Low	Low	3.597	3.661	78.272	73.119	0.584	0.558
		High	3.795	3.983	84.119	73.408	0.568	0.542
	Middle	Low	2.834	2.720	79.402	79.615	0.591	0.535
		High	5.716	5.490	72.652	77.776	0.483	0.495
	High	Low	4.529	4.573	69.647	75.541	0.470	0.494
		High	3.380	3.371	64.624	75.471	0.487	0.465
South	Low	Low	7.599	7.554	77.323	76.072	0.486	0.483
		High	6.620	6.532	78.378	72.407	0.463	0.470
	Middle	Low	8.409	8.204	79.731	77.805	0.423	0.467
		High	6.638	6.508	77.872	77.212	0.464	0.462
	High	Low	6.902	6.991	75.580	80.013	0.422	0.443
		High	6.733	6.933	76.084	74.152	0.430	0.465
West	Low	Low	2.034	2.079	70.370	77.364	0.483	0.502
		High	2.476	2.841	77.947	76.881	0.477	0.528
	Middle	Low	2.985	2.867	72.871	79.884	0.501	0.464
		High	3.503	3.465	73.925	77.295	0.468	0.500
	High	Low	8.136	7.919	69.560	70.943	0.449	0.458
		High	0.640	0.669	82.353	74.434	0.305	0.432

Table 6
Simulated Impacts on Migration, Employment, and Math Scores

	Raw data	Simulated			
		Baseline	**Increase expenditure by 25%**	**Increase dropout rate by 25%**	**Increase wage rate by 25%**
Biyearly migration rate	11.34	11.92	11.42	11.41	11.41
Employment rate	75.14	75.62	75.70	75.71	76.14
Math score	0.491	0.490	0.486	0.479	0.483

Bibliography

Bayer, P., 'Tiebout Sorting and Discrete Choices: A New Explanation for Socioeconomic Differences in the Consumption of School Quality', manuscript, Yale University, 2000.

Becker, G. and N. Tomes, 'Child Endowments and the Quantity and Quality of Children', *Journal of Political Economy*, vol. 75(4), part I, August, 352-365, 1976.

Behrman, J.R. and E.M. King, 'Household schooling behaviors and decentralization', *Economics of Education Review*, 20, 321-341, 2001.

Bernal, R., 'Employment and Child Care Decisions of Mothers and the Well-Being of their Children', New York University mimeo, 2003.

Betts, J., 'Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth', *Review of Economics and Statistics*, 77, 231-50, 1995.

Betts, J., 'Is There a Link Between School Inputs and Earnings? Fresh Scrutiny of an Old Literature', in Gary Burtless, ed. *Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success*, 141-91, Washington, D.C.: Brookings Institution, 1996.

Bhargava, A., 'Identification and Panel Data Models with Endogenous Regressors', *Review of Economic Studies*, 58(1) 129-140, 1991.

Black, S.E., 'Do Better Schools Matter? Parental Valuation of Elementary Education', *Quarterly Journal of Economics*, May 1999.

Blanchard, O.J. and L.F. Katz, 'Regional Evolutions', *Brookings Papers on Economic Activity*, 1, 1-37, 1992.

Blau, F.D. and A.J. Grossberg, 'Maternal Labor Supply and Children's Cognitive Development', *Review of Economics and Statistics*, vol. 74(3), 474-81, August, 1992.

Card, D., 'Earnings, Schooling and Ability Revisited', in S. Polachek ed. *Research in Labor Economics*, 14, 1995, 23-48.

Card, D., 'The Causal Effect of Education on Earnings', in *handbook of Labor Economics*, Vol. 3A, North-Holland, 1999.

Card, D. and A. B. Krueger, 'Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States', *Journal of Political Economy*, 100(1), 1-40, 1992.

Ferreyra, M.,' Estimating the Effects of Public School Vouchers in Multi-District Economies', manuscript, University of Wisconsin-Madison, 2001.

Gilleskie, D. and T. Mroz, 'A Flexible Approach for Estimating the Effects of Covariates on Health Expenditures', *Journal of Health Economics*, 23(2), 391-418, 2004.

Hanushek, E.A., 'School Resources and Student Performance', in Gary Burtless, ed. *Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success*, 43-73. Washington, D.C.: Brookings Institution, 1996.

Hanushek, E.A., 'Conclusions and Controversies about the Effectiveness of School Resources', *Economic Policy Review*, 4(1) New York: Federal Reserve Bank of New York, March 1998.

Hanushek, E.A., S.G. Rivkin and L.L. Taylor, 'Aggregation and the Estimated Effects of School Resources', *The Review of Economics and Statistics*, 78(4), 611-27, 1996.

Harvey, E., 'Short-Term and Long-Term Effects of Early Parental Employment on Children of the National Longitudinal Survey of Youth', *Developmental Psychology*, 35, 445-459, 1999.

Haveman,R. and B. Wolfe, ' The Determinants of Children's Attainments: A Review of Methods and Findings', *Journal of Economic Literature*, 33, 1829-1878, 1995.

Hedges, L., R. Greenwald, 'Have Times Changed? The relationship Between School Resources and Student Performance', in Gary Burtless, ed. *Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success*, Washington, D.C.: Brookings Institution, 1996.

Imbens, G. and J. Angrist, 'Identification and Estimation of Local Average Treatment Effects', *Econometrica*, 62, 1994.

Kennan, J. and J.R. Walker, 'The Effect of Expected Income on Individual Migration Decisions', manuscript, University of Wisconsin-Madison, 2002.

Krueger, A.B., 'Reassessing the View That American Schools Are Broken', *Economic Policy Review*, 4(1), New York: Federal Reserve Bank of New York, March 1998.

Krueger, A.B., 'Experimental estimates of education production functions', *Quarterly Journal of Economics*, 497-532, 1999.

Liu, H, 'Using random sampling to reduce the choice set of a conditional Logit model estimated by the discrete factor method', manuscript, East Carolina University, 2004.

Loeb, S. and J. Bound, 'The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts', *Review of Economics and Statistics*, 78(4), 653-664, 1996.

McFadden, D., 'Modeling the choice of residential location', in A. Karlqvist, L. Lundqvist, F. Snichars, and J. Weibull, eds *Spatial Interaction Theory and Planning Models*, Amsterdam: North Holland, 75-96, 1978.

McMillan, R., 'Parental involvement, competition, and public school performance: An empirical analysis', manuscript, University of Toronto, 2001.

Moore, K.A. and A.K. Driscoll, 'Low-wage maternal employment and outcomes for children: A study', *The Future of Children*,7(1), 122-127, 1997.

Nechyba, T.J. and R.P. Strauss, 'Community choice and local public services: A discrete choice approach', *Regional Science and Urban Economics*, 28(1), 51-74, 1998.

Parcel, T.L. and E. Menaghan, 'Early parental work, family social capital and early childhood outcomes', *American Journal of Sociology*, vol. 99(4), 972-1009, 1994.

Todd, P.E. and K.I. Wolpin, 'On the specification and estimation of the production function for cognitive achievement', *Economic Journal*, vol. 113, 485-515, 2003.

Appendices

Summaries of Major Data Sources.

1. *The National Longitudinal Survey of Youth*

The original NLSY began in 1979 with a national sample of 12686 young adults between the ages of 14 and 21. It included a nationally representative sample of 6,111 youths, an over-sample of 5,295 blacks, Hispanics, and economically disadvantaged whites, and a supplemental sample of 1,280 persons in the military in September 1978. Interviews with the military sub-sample were suspended after 1984 and for economically disadvantaged non-Hispanic whites after 1990. In this study, we exclude these economically disadvantage non-Hispanic whites and focus on the mothers who had mother-child data from 1986 to 1998.

2. *The National Longitudinal Survey of Youth - Children Sample—*

Beginning in 1986, the NLSY-Child collected data on all of the children born to the female NLSY respondents. The NLSY-Child sample (through 1998) supplies data on children with mothers between the ages of 33 and 40 at the end of 1997. Children under the age of 15 comprise the majority of this sample. The NLSY-Child contains a set of cognitive and behavioral assessments. The NLSY-Child sample biennially interviews both mother and child.

3. *Top 100 Database of Key Demographic Items, School District Data Book*

This is a compact file of key demographic data items, drawn mainly from 1990 Census school district special tabulation. Expenditures per pupil are obtained by counties from this data set.

4. *Census of Population and Housing, 1990 [United States]: Public Use MicroData Sample: 5-percent Sample*

To construct our relative measurements of teacher salary and expenditure per pupil, we select college-graduated white males, who were between 27 and 38 years old and working full-time (35+ hours a week and 40+ weeks a year). The smallest geographical indentifier in the 5-percent PUMS data is Public Use Microdata Area (PUMA), which could include partial, single or multiple counties. We aggregate all PUMA's to 908 study areas, and any given county exclusively belong to one of them. In this selected sample, the relative median teacher salary in a county is calculated by dividing the median annual wage income of male public non-postsecondary teachers by the median annual wage income of males with occupations other than these teachers. These non-postsecondary teachers include pre-kindergarten and kindergarten teachers, elementary school teachers, and secondary school teachers). These grade-dependent teacher salaries then are used in the education production function estimated in the paper. Similarly, the relative expenditure per pupil is measured by the nominal expenditure per pupil relative to the median annual wage income of all male in the selected MicroData sample.

The empirical model also uses the median wage rates of females in this data set by age, education attainment, and ethnic groups. These median wage rates are treated as important elements of local wage distribution.

5. *USA Counties*

USA Counties is a database produced by the U.S. Bureau of Census. It contains statistical data from the Census Bureau, other federal agencies, and private organizations. High school drop-out rates in 1990 are obtained from this database.