Maternal Employment, Migration, and Child Development[1]

Haiyong Liu
Department of Economics
East Carolina University

Thomas Mroz
Department of Economics and Carolina Population Center
University of North Carolina at Chapel Hill

Wilbert Van der Klaauw
Department of Economics and Carolina Population Center
University of North Carolina at Chapel Hill

Current version: November 2003

[1]We have benefited from helpful comments from participants at the 2002 SITE meetings at Stanford, and seminar participants at the University of Wisconsin-Madison.

Abstract

In this paper we analyze the roles and interrelationships between school inputs and parental inputs in affecting child outcomes in the U.S. We investigate the interactions among and endogeneity of these inputs in the production of child outcomes by specifying and estimating a behavioral model of parent's decisions that can affect these outcomes. We focus on two important dimensions of school and parental input decisions: the parents' choice of which school attendance area to live in, and the mother's decision to work as a proxy for maternal time directly devoted to child education. Parents receive utility from consumption, leisure, and the child's achievement and they maximize expected utility. In making location and employment decisions, parents take into account the distribution of impacts of these decisions on their child's educational development, modeled through a production function for child outcomes. The environment in which these decisions are made is characterized by uncertain future wages and job prospects for both parents, and uncertainty in the child's future educational outcomes. Besides school quality, residential location decisions are influenced by local labor market conditions, housing and moving costs and geographic preferences. Using data from the National Longitudinal Survey of Youth, we integrate information on household migration, maternal employment decisions, and parental wage rates with observations on child outcomes over a 10 year period. Our statistical model follows directly from the theoretical framework. We relax many functional form assumptions that have been imposed by previous researchers who have studied how parents and schools can affect a child's development. Our preliminary findings are that we are able to explain and reject several of the counter-intuitive estimation results found in the literature on the determinants of children's school performance.

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 offer, the mother chooses how much time to devote to labor market activities. She does this with the understanding that more hours spent in the labor market could cut back on her time input to education production. 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 put 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 on 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 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 will investigate the following policy issues: 1) the effects of exogenous changes in several school quality measures on mother's labor supply and child outcomes; 2) the effects of a change in mothers' wages on labor supply, schooling choices, and outcomes. This study will provide a unique perspective on designing educational policies as well as social programs targeted to encourage maternal employment.

2 Background

The literature on the impact 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 both extensive and have been extensively reviewed 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 child outcome measure 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. Similar findings were reported by Dearden et al (2002) using data from Britain.

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 a parental response 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 together with 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 proposed 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 will explicitly address the endogeneity of the mother's work decision.

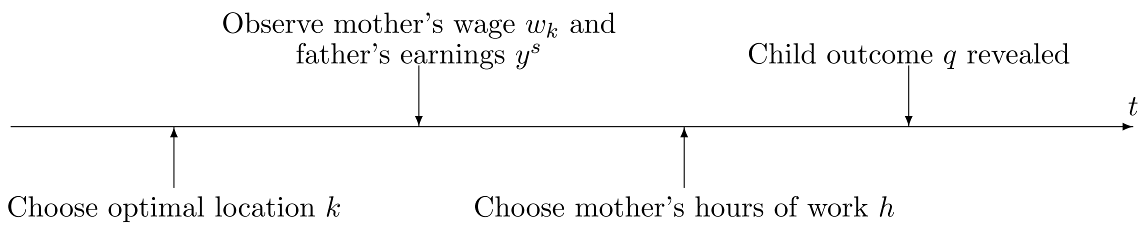
3 The model

In this section we present our theoretical model, as well as the estimation method that we will 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 will correspond 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 has preferences over consumption, x_i, the child's cognitive achievement, q_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 hours of work as a discrete choice variable equal to zero (h_0), part-time (h_1), or full-time (h_2), and will adopt the following specification of the utility function:

$$\begin{aligned} U^i(x_i, q_i, h_i, k_i) \quad = \quad & \frac{(x_i + \gamma_{0i})^{\gamma_1}}{\gamma_1}[1 + \alpha_0 MS_i + \alpha_1 \cdot \mathbf{1}(h_i > 0)] \; + \frac{\alpha_2}{\gamma_3}(\frac{q_i}{\bar{q}_i^{\alpha_3}} + \gamma_2)^{\gamma_3} \\ & + \alpha_{4ai} \cdot \mathbf{1}(h_i > 0) + \alpha_{4bi} \cdot \mathbf{1}(h_i = h_2) + \sum_{j=0}^{2} \mathbf{1}(h_i = h_j)\varepsilon_{h_j,i} \end{aligned}$$

$$+m_{k_{-1}k} + \sum_{j=1}^{4} \alpha_{5j} \mathbf{1}[R(k_i) = j] + \xi_{k,i}, \tag{2}$$

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 the child's cognitive achievement score q_i 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 special case where the mother's score is 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 geographic census regions, as in

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

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 will 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:

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

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

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

$$x_i = w_{ki}h_i + y_{ki}^s + y_i^o - c_{ki}, \tag{4}$$

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_{ki}^s are the earnings of the father. Household non-labor income is denoted by y_i^o, and c_{ki} explicitly measures the average housing cost in location k.[2]

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

$$q_i = f_q(t_i, S_k, T - h_i, Z_i, MS_i, \mathbf{1}(k \neq k_{-1}), x_i, \mu_i^c, \varepsilon_{q,i}), \tag{5}$$

where t represents the child's age 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 earnings 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, her race and marital status, and also includes the child's gender.

The specification allows for a (one-period) reduction in the child's performance due to a move. To capture other omitted educational expenditures, we include x_i, the net consumption level of the household, given by equation (4). The implicit assumption here is that these additional educational expenditures are a fixed proportion of consumption. Unobserved heterogeneity in the child's intellectual endowments is captured by the inclusion of μ_i^c. This

[2]Given that moving costs are unobserved, to avoid identification problems, instead of modeling both psychic and monetary moving costs, we only incorporate the former.

term represents the child's unobserved endowment at the age at which they enter our data set (age 5 or 6). The stochastic component $\varepsilon_{q,i}$ represents all other unobserved factors influencing the child's outcome at each age, and is assumed to be independently distributed across individuals and periods. The exact specification of f_q and the distribution of $\varepsilon_{q,i}$ will be discussed later.

Similar to the treatment of child achievement scores, the mother's hourly wage rates will be discretized and the wage offer distribution function in location k will be specified as a flexible parametric function of the variables describing local labor market conditions ($\overline{W_{ki}}$), mother characteristics in Z_i and the unobserved trait μ_i, as in

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

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 will be 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

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

where Ω is a set of state variables, $\Omega = \{t, S_k, Z, k_{-1}\}$, μ represents the vector (μ, μ^s, μ^c), and $x(h)$ and $q(h)$ are the values of consumption and the child's achievement score when the mother works h hours as implied by (4) and (5), and we have dropped the individual subscript, i, to simplify notation.

This requires the calculation of an expectation because the child's educational outcome is stochastic (though influenced by hours 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

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

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

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

With ε_h being independently, and identically distributed extreme value errors, the probability of working h hours conditional on wage and earnings draws w_k and y_k^s equals

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

At the time of the location decision, while each woman knows the distributions of her wage offer and the husband's earnings at each location $k = 1, 2, \cdots, K$, she does not know the actual wage she will be offered at each location, nor does she know what her husband will earn. 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

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

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 all possible wage offers and husband's earnings realizations in location k yields

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

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

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

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

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

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 will be discretized and the educational production function in (5) will be specified using a flexible parametrization of its corresponding "hazard rate." More specifically, we will follow Gilleskie and Mroz (2002) 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

$$\Pr(q_i > p | q_i \geq p) = \frac{\exp\left[f_{vi} + f_b(p)\right]}{1 + \exp\left[f_{vi} + f_b(p)\right]}, \tag{16}$$

where q_i is the observed score of child i, $f_b(p)$ represents the "baseline" hazard at achievement level p and f_{vi} captures the effect of covariates.

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

$$f_b(v) = \sum_{n=1}^{3} \left\{\theta_{qn}[-\log(P - p)]^n\right\}, \tag{17}$$

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

$$\begin{aligned} f_{vi} &= f_v(t_i, S_k, T - h_i, Z_i, MS_i, \mathbf{1}(k \neq k_{-1}), x_i, \mu_i^c) \\ &= \beta_{q0} + \beta_{q1} t_i + \beta_{q2} S_k + \beta_{q3a} \mathbf{1}(h_i = h_1) + \beta_{q3b} \mathbf{1}(h_i = h_2) \\ &\quad + \beta_{q4} Z_i + \beta_{q5} MS_i + \beta_{q6} \mathbf{1}(k \neq k_{-1}) + \beta_{q7} x_i + \rho_q(\mu_i^c), \end{aligned} \tag{18}$$

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 $p_{\mu,j},\ j = 1, \ldots, J$.

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 will 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 parent characteristics by including interactions between $f_b(p)$ and several determinants of f_{vi}. This flexibility will allow us to estimate the degree of substitutability or complementarity across all inputs.

We will similarly model the mother's wage offer distribution, discretizing the wage data into G groups, with the probability density function $Pr(w_k = g)$ of having a wage offer g similarly defined in term of the corresponding conditional probabilities

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

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

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

and $\overline{W_{ki}}$ is the median local earnings for women with the same education level, race, and age.

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

$$f_{yi} = \beta_{y0} + \beta_{y1}\overline{W^s_{ki}} + \beta_{y2}Z^s_i + \rho_y(\mu^s_i) \text{ , and} \tag{22}$$

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

where $\overline{W^s_{ki}}$ is the median local wage rate for men with the same education level, race, and age.

3.4 Estimation

As discussed below, our data set will include longitudinal information on the place of residence, the mother's work decision and wage, the husband's earnings (if married), the child achievement test score, as well as all covariates described in our model, for each child-mother pair 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^*_{kit} and husband's earnings y^{s*}_{kit} at the location k^* that was chosen, and 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 different values, the likelihood contribution for child i is given by

$$\begin{aligned} L_i \; = \; & \sum_{j=1}^{J} \Bigg\{ \Pr(\mu_j) \prod_{t=t_0}^{T_i} \Bigg\{ \Pr(k_{it}^*|\Omega_{it}, \mu_j) \Bigg\{ \sum_{g=1}^{G} [\Pr(w_{kit} = g|\Omega_{it}, \mu_j, k_{it}^*) \\ & \Pr(h_{it}^*|\Omega_{it}, w_{kit}, \mu_j, k_{it}^*) \Bigg\}^{\mathbf{1}(h_{it}^*=0)} \cdot \Bigg\{ \Pr(w_{kit} = w_{kit}^*|\Omega_{it}, \mu_j, k_{it}^*) \\ & \Pr(h_{it}^*|\Omega_{it}, w_{kit}^*, \mu_j, k_{it}^*) \Bigg\}^{\mathbf{1}(h_{it}^*\neq 0)} \cdot \Pr(y_{kit}^s = y_{kit}^{s*}|\Omega_{it}, \mu_j, k_{it}^*)] \cdot \Pr(q_{it}|\Omega_{it}, h_{it}^*, y_{kit}^{s*}, \mu_j, k_{it}^*)] \Bigg\} \Bigg\}. \end{aligned}$$

The sample likelihood function then is given by

$$L = \prod_{i=1}^{N} L_i. \tag{24}$$

The size of the choice set we will 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 $D(k_i^*) \subseteq C$ be a subset consisting of $\widetilde{K}$ elements. The sampling method is to select the chosen alternative plus $\widetilde{K}-1$ non-chosen alternatives drawn with probability p at random from the set C. Let $\Pr[D(k_i^*)]$ be the probability that $D(k_i^*)$ will be drawn, given the observed choice k_i^* and the vector of state variables Ω. Then

$$\Pr[D(k_i^*)] = \begin{cases} p^{\widetilde{K}-1}(1-p)^{K-\widetilde{K}} & \text{if } D(k_i^*) = \{k_i^*, \ldots\} \subseteq C \\ 0 & \text{otherwise.} \end{cases}$$

As shown by McFadden, the choice probabilities $\Pr(k_{it}^*|\Omega_{it}, \mu_j)$ in equation (15) can then be written as follows:

$$\Pr(k_i = k_i^*|\Omega_i, \mu_i) = \frac{\exp\left\{\frac{\widetilde{V}_k(\Omega_i,\mu_i)}{b_2} + \ln[\Pr(D(k_i^*))]\right\}}{\sum_{k' \in D(k_i^*)} \exp\left\{\frac{\widetilde{V}_{k'}(\Omega_i,\mu_i)}{b_2} + \ln[\Pr(D(k'))]\right\}}. \tag{25}$$

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 (25) 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 we conducted a small scale Monte-Carlo study to test the sensitivity of parameter estimates to the size of the sampling set $\widetilde{K}$. First, we generated a conditional logit model with 10 exogenous variables, 10,000 choices, and 7,000 observations, which is comparable to the empirical model we will use in the study. Second, different sizes of subsets were drawn from the full choice set. For each subset size, the standard errors for corresponding parameters were obtained by repeating the sampling and estimating 100 times. We found 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 will be a child observed between ages 5 and 15 in 1986 or later, with up to six possible time-period specific observations per child. 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 will 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, will serve as our child outcome measure.

In our model households make residential location choices based on local labor market conditions, housing prices, geographic preference, 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 will 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 will then be characterized by the median wage in each locality corresponding to her/his education level, race, and age. That is, the median wage will be 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 will use to measure school inputs at each location will be measured at the smallest geographic unit currently available in case of the NLSY, the county level[3]. The choice set from which families choose their residential location in each period therefore is the set of all 3,181 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 indicator 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 will measure these variables relative to the average annual earnings of college

[3]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.

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 were 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 function also depends upon characteristics of the county 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.

To describe the wage offers available to the individual women in each locality, we use a discrete distribution with 10 points of support, and model the probability that each woman has an offered wage from each of the ten 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, 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. As not all husbands in the sample had positive earnings, one of the ten masspoints corresponded to zero earnings. In addition to the local median male wage rate in the county, the equation includes the husband's race, and education level as explanatory variables.

Table 1 contains summary statistics for the data used in our analysis. Table 2 contains the data on school district characteristics broken down into 24 different subsets of counties. These 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.

5 Estimation Results

I. Point Estimates of Utility and Production Function Parameters.

Tables 3A, 3B, 3C, and 4D contain parameter estimates for the production function, utility function, wage offer, and heterogeneity distribution parameters. The estimates in these tables 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 wage offer distribution and the distribution of spousal earnings. 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 child's years of observation.

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.

II. Simulations of Effects

A. The Production Function Marginal Effects

To interpret the implications of the model estimates we simulate outcomes using the estimated parameters. Table 4A 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 (50 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 4A 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 eighty five percent. The negative effect of higher teacher salaries falls by about a quarter, and the impact of expenditures per pupil falls by more than eighty 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 regions becomes negative and significant after controlling for the endogenous behaviors, while without the endogeneity controls it appears that a move significantly increases a child's test score. The positive impacts of the mother working part time or full-time (instead of not working) both become larger in absolute value but become negative after controlling for the endogeneity of these parental decisions. Without endogeneity controls, it would appear that a mother's working full-time could increase her child's percentile score by 3 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 5.5 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.

B. Simulations of Changes in Location Patterns

Forthcoming.

6 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 8 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 part-time or 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 become small and insignificant determinants of the child outcomes, and are much smaller than is implied by estimation methods that do not allow for possible endogeneity biases.

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.

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.

Dearden, L., J. Ferri and C. Meghir, ‘The Effect of School Quality on Educational Attainment and Wages’, *Review of Economics and Statistics*, February 2002, Vol. 84(1).

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

Gilleskie, D. and T. Mroz, ‘Estimating the Effects of Covariates on Health Expenditures’, manuscript, UNC-Chapel Hill, 2000.

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.

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.

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, 'Towards a Unified Approach for Modeling the Production Function for Cognitive Achievement', forthcoming in *Economic Journal*, 2002.

Appendix

Appendix A 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). 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). 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.

Appendix B. Defining regions in this study

24 regions are constructed among 3141 counties or county equivalents in the United States by the following three steps: Step 1. The counties within each Census Region (West, Midwest, Northeast and South) are ordered by drop-out rates and then divided into three groups with equal size of population; Step 2. The counties within each of these 12 groups of counties are further ordered by the amount of expenditure by pupil and divided into two sub-groups with equal size of population; Step 3. The drop-out rates, relative teacher salary, and relative expenditure per pupil are then calculated for these 24 sub-groups using weighted medians of counties within each sub-group.

Table 1

Summary Statistics
Number of Children: 6964
Number of Children-Years: 12387
(means over children-years)

Variable	Mean	Standard deviation
Age of mother	33.481	0.341
Age of child	10.497	2.198
Married	0.583	0.493
High school	0.482	0.500
More than high school	0.328	0.470
Non-white	0.568	0.495
Boy	0.501	0.500
Dropout rate	0.063	0.011
Expenditure per pupil(1)	0.124	0.039
Teacher salary(1)	0.801	0.175
Yearly housing cost(2)	0.664	0.132
Move	0.113	0.317
Net income(2)	2.379	2.503
Hourly wage rate(3)	1.012	0.783
Father's yearly earnings(2)	2.603	2.065
Local median hourly wage rate(3)	1.001	0.274
Part-time work(4)	0.312	0.463
Full-time work(4)	0.439	0.496
PIAT Math score	0.497	0.285
# of young children in HH (0-5)	0.493	0.718
# of old children in HH (6-17)	2.171	1.082
Mother's AFQT score	0.450	0.281

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

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	6.15	9.23	0.760	0.116	0.152	0.997
		High	6.93	10.40	0.721	0.152	0.152	0.898
	Middle	Low	10.09	15.14	0.789	0.120	0.136	0.850
		High	8.77	13.16	0.782	0.136	0.136	0.715
	High	Low	13.21	19.81	0.857	0.111	0.134	0.786
		High	15.10	22.65	0.905	0.134	0.134	0.676
Midwest	Low	Low	5.46	8.19	0.818	0.087	0.117	0.519
		High	6.10	9.16	0.819	0.117	0.117	0.558
	Middle	Low	9.20	13.81	0.709	0.095	0.113	0.588
		High	10.04	15.05	0.687	0.113	0.113	0.594
	High	Low	14.99	22.48	0.676	0.089	0.106	0.602
		High	13.31	19.97	0.800	0.106	0.106	0.715
South	Low	Low	10.07	15.10	0.728	0.076	0.109	0.543
		High	9.51	14.27	0.821	0.109	0.109	0.636
	Middle	Low	12.29	18.44	0.706	0.071	0.113	0.537
		High	12.80	19.20	0.757	0.113	0.113	0.655
	High	Low	17.14	25.72	0.694	0.081	0.110	0.606
		High	16.33	24.50	0.788	0.110	0.110	0.594
West	Low	Low	9.51	14.26	0.675	0.081	0.104	0.861
		High	9.36	14.04	0.922	0.104	0.104	0.652
	Middle	Low	15.02	22.54	0.789	0.078	0.096	0.856
		High	14.93	22.40	0.743	0.096	0.096	0.650
	High	Low	17.35	26.02	0.779	0.090	0.098	0.937
		High	16.92	25.38	0.727	0.098	0.098	0.507
Overall Sample Mean			12.09	18.09	0.768	0.098	0.112	0.659

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	0.034	0.184
Age of mom (in 10 years)	-0.200	0.019
child age (in 10 years)	0.343	0.256
child age squared	-0.051	0.124
Married	0.143	0.018
Mother education high school	0.176	0.017
Mother education more than HS	0.330	0.021
Non-white	-0.303	0.016
Boy	0.034	0.012
Dropout rate	-0.141	0.490
Expenditure per pupil	-0.095	0.580
Teacher salary	0.053	0.157
Move	-0.067	0.018
Mother part-time work	-0.079	0.065
Mother full-time work	-0.179	0.069
Consumption (in $10,000's)	0.010	0.004
Dropout rate × Log(*P*-*p*)	-0.192	0.803
Expense per pupil × Log(*P*-*p*)	-0.023	0.183
Teacher salary × Log(*P*-*p*)	-0.115	0.069
Mother part-time × Log(*P*-*p*)	-0.102	0.031
Mother full-time × Log(*P*-*p*)	-0.112	0.032
Dropout rate× Child age	-0.155	1.047
Expense per pupil × Child age	-0.077	0.497
Teacher salary×Child age	-0.009	0.108
Part-time × Child age	-0.003	0.027
Full-time × Child age	0.075	0.035
Mother's AFQT score	0.313	0.077
Mother's AFQT × Log(*P*-*p*)	0.562	0.046
Loading on 1st order heterogeneity factor	-0.038	0.240
Loading on 2nd order heterogeneity factor	-0.312	0.580
Loading on 3rd order heterogeneity factor	-0.479	0.366
1st order of baseline hazard [log(*P*-*p*)]	-1.236	0.289
2nd order of baseline hazard [log(*P*-*p*)]^2	-0.195	0.546
3rd order of baseline hazard [log(*P*-*p*)]^3	0.069	0.357
4th order of baseline hazard [log(*P*-*p*)]^4	0.062	0.075

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	1.927	Std. Error
1st order discrete factor loading in reserve	2.192	1.034
2nd order discrete factor loading in reserve	-0.906	0.602
Power of consumption (γ_1)	0.100	0.633
Married (α_0)	0.038	0.045
Any work (α_1)	0.280	0.070
Relative power of mother's AFQT score (α_3)	0.776	0.035
Scale on child's score / mother's AFQT score (α_2)	0.988	0.576
Reserve child's score / mother's AFQT score (γ_2)	3.249	1.252
Power of child's score / mother's AFQT score (γ_3)	0.404	0.003
Intercept in part-time leisure	1.194	0.256
1st order discrete factor loading in part-time leisure	-1.501	0.184
2nd order discrete factor loading in part-time leisure	0.228	0.395
3rd order discrete factor loading in part-time leisure	2.713	0.922
# of young children (0-5) in part-time leisure	0.069	0.690
# of old children (6-17) in part-time leisure	0.011	0.029
Married in $f_{4a}(.)$	-0.002	0.013
Intercept in full-time leisure	-3.185	0.020
1st order discrete factor loading in full-time leisure	-1.625	0.192
2nd order discrete factor loading in full-time leisure	-2.040	0.392
3rd order discrete factor loading in full-time leisure	1.815	0.684
# of young children (0-5) in full-time leisure	0.941	0.011
# of old children (6-17) in full-time leisure	0.707	0.045
Married in $f_{4b}(.)$	-0.673	0.034
Moving psychic cost (any move)	-3.729	0.077
Additional moving psychic cost across states	-2.596	0.958
Additional moving psychic cost across census regions	-1.097	0.799
Inverse of parameter **b** Gumbel error on working choice	2.262	0.272
Inverse of parameter **b** Gumbel error on location choice	1.228	0.312
Dummy for census region – Midwest(1)	-0.680	0.302
Dummy for census region – South	-0.296	0.129
Dummy for census region – West	-0.209	0.119

Notes: (1). Northeast is the base region.

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

Variable	Estimates	Std. Err.
Intercept	-1.652	0.113
Local wage rate	0.979	0.094
Local wage rate squared	-0.221	0.028
Local wage rate × discrete factor	-0.270	0.102
High school (mother)	0.375	0.030
More than high school (mother)	0.630	0.037
Non-white	0.223	0.020
Mother's AFQT score	0.334	0.098
Mother's AFQT × Log(*G*-*g*#)	0.463	0.060
Loading on 1st order heterogeneity factor	-0.644	0.182
Loading on 2nd order heterogeneity factor	-0.892	0.123
1st order of baseline hazard [log(*G*-*g*)]	-1.077	0.150
2nd order of baseline hazard [log(*G*-*g*)]^2	-0.173	0.151
3rd order of baseline hazard [log(*G*-*g*)]^3	-0.073	0.044

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	-2.181	0.097
Local median earnings	0.042	0.003
Local median earnings × discrete factor	-0.067	0.004
High school (father)	0.606	0.039
More than high school (father)	1.265	0.045
Non-white	-0.625	0.033
Loading on 1st order discrete factor	-2.848	0.195
Loading on 2nd order discrete factor	-1.725	0.162
1st order of baseline hazard [log(*G'*-*g'*)]	-1.602	0.149
2nd order of baseline hazard [log(*G'*-*g'*)]^2	0.001	0.157
3rd order of baseline hazard [log(*G'*-*g'*)]^3	-0.147	0.046

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	3.226	0.077
Probability parameter at 1/3	-4.025	0.136
Probability parameter at 2/3	5.021	0.152

Table 3F

Probability Distribution of Heterogeneity Types

Heterogeneity factor	Probability
0	0.449
1/3	0.368
2/3	0.157
1	0.026

Table 4A

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)
Age of mother	–0.029***	–0.053***	–0.069***
	(0.007)	(0.009)	(0.007)
Age of child	0.067***	0.044***	0.028***
	(0.011)	(0.008)	(0.008)
Married	0.015***	0.012***	0.029***
	(0.006)	(0.004)	(0.004)
High school (mother)	0.041***	0.030***	0.036***
	(0.007)	(0.004)	(0.003)
More than high school	0.060***	0.053***	0.067***
	(0.008)	(0.005)	(0.004)
Non-white	–0.055***	–0.048***	–0.062***
	(0.006)	(0.004)	(0.003)
Boy	0.013***	0.013***	0.007***
	(0.005)	(0.003)	(0.002)
Dropout rate	–0.970***	–0.936***	–0.123
	(0.233)	(0.174)	(0.101)
Teacher salary	–0.016	–0.011	–0.012
	(0.014)	(0.012)	(0.008)
Expenditure per pupil	0.183***	0.187***	0.033
	(0.061)	(0.045)	(0.028)
Move	0.018**	0.016***	–0.013***
	(0.007)	(0.006)	(0.004)
Part-time work	0.037***	0.036***	–0.050***
	(0.006)	(0.005)	(0.003)
Full-time work	0.024***	0.031***	–0.055***
	(0.006)	(0.005)	(0.005)
Net income	0.008***	0.005***	0.001**
	(0.001)	(0.001)	(0.001)
AFQT score (mom's)	0.280***	0.270***	0.253***
	(0.012)	(0.007)	(0.006)
Intercept	0.382***	–	–
	(0.032)	–	–

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

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	0.162***	0.204***	0.221***
	(0.020)	(0.018)	(0.028)
High school	0.061**	0.122***	0.139***
	(0.026)	(0.015)	(0.043)
More than high school	0.208***	0.221***	0.207***
	(0.030)	(0.021)	(0.020)
Non-white	0.121***	0.129***	0.302***
	(0.018)	(0.012)	(0.022)
Mother AFQT score	0.618***	0.606***	0.475***
	(0.037)	(0.026)	(0.027)
Intercept	0.218***	–	–
	(0.032)	–	–

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

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 median wage	0.032***	0.035***	0.028***
	(0.002)	(0.002)	(0.002)
High school	0.423***	0.520***	0.493***
	(0.063)	(0.034)	(0.038)
More than high school	1.291***	1.291***	1.114***
	(0.071)	(0.049)	(0.048)
Non-white	–0.806***	–0.755***	–0.550***
	(0.047)	(0.032)	(0.035)
Intercept	1.443***	–	–
	(0.077)	–	–

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