A Migration Study of Mother's Work, Welfare Participation, and Child Development

Haiyong Liu[*]
Department of Economics
East Carolina University

September 2007

Abstract

This paper investigates how women's migration and labor supply behaviors respond to changes in welfare policies and labor market conditions, controlling for endogenous initial residence and unobserved heterogeneity. It also traces out how these responses influence educational inputs and child outcomes. The simulation results show that poor and low-educated single women with children do change their residential locations in response to changes in welfare policies and labor market conditions. The magnitude of this response in the form of migration is rather modest. More importantly, however, such policy changes often have large and important impacts on particular at-risk groups.

JEL Classification Codes: I21, J22, J60, R23.

---

[*] Department of Economics, East Carolina University, Greenville, NC 27858 (Tel: 1-252-328-1083 Fax: 1-252-328-6743 E-mail: liuh@ecu.edu).

1 Introduction

This paper investigates whether parents base their migration decisions on welfare benefits, labor market conditions, and quality of school districts, how women's migration and labor supply behaviors respond to changes in Aid to Families with Dependent Children (AFDC) policies and labor market conditions, and how these responses affect their children's achievement outcomes.

Understanding how at-risk individuals change their migration and labor supply behaviors in response to reforms in welfare policies can provide crucial information to policy makers as they design assistance programs to achieve certain desired outcomes. With the passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996, state governments in the United States have much greater power in reforming their welfare programs. In particular, researchers from different disciplines have investigated whether changes in Aid to Families with Dependent Children (AFDC) policies have caused women to adjust their labor supply and residential location decisions. However, the literature on these issues is far from definite. For instance, few studies analyze the interaction between welfare migration and labor supply decisions. In addition, previous studies reach little consensus on the extent to which inter-state benefit differentials induce welfare recipients to move across states to seek more assistance. [1] For instance, Gramlich and Laren (1984) and Blank (1988) find substantial migration flows induced by cross-section benefit differentials, while Walker (1994), Kennan and Walker (2001; 2003), and Levine and Zimmerman (1999) show virtually no evidence to support the existence of welfare migration. A recent paper by Meyer (2000) addresses explicitly several reasons why studies using different approaches could give rise to different estimates about the magnitude of welfare migration. He emphasizes that not controlling for important migration determinants tends to bias estimates on the magnitude of welfare migration. For instance, Blank (1988) finds significant welfare-induced migration by estimating a model with joint decisions about migration and AFDC participation. One main problem in her analysis is that instead of modeling work choice explicitly she uses hours worked as an independent variable. Given the fact that high-benefit states tend to have high wage rates, one cannot infer if her findings are due to wage earnings or welfare benefits.

Given the fact that the AFDC program was established in a bid to provide financial assistance to children who were deprived of support due to death, incapacity, or absence of parent (Burke, 1997), it is surprising that our understanding about the impacts of these welfare-induced behavioral changes on children's achievement outcomes is very limited. Researchers who study educational production functions show that migration and labor supply decisions have several channels to interact with child development (Hanushek, 2004; Blau and Grossberg, 1992). For instance, it would be expected that welfare-induced migration might improve children's immediate financial well-being. On the other hand, moving could also have negative impacts on children's achievement outcomes because of disruption of regular schooling. Changes in school inputs associated with a move would also affect child outcomes. It is, however, not clear what is the net impact on these children whose mothers decide to move. A joint decision of changing employment status and relocation further complicates a modeling of the educational production process. 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. Even two of the most recent reviews of the literature (Hanushek *el al.*, 1998; Krueger, 1998) present conflicting interpretations of the literature.

In seeking to answer the research questions posed at the beginning of this section, this paper makes an important advance by developing a conceptual model that characterizes the simultaneity and endogeneity of decision making about migration, welfare program participation, and labor supply. Furthermore, the model recognizes that all of these decisions could impact their children's achievement outcomes. We implement this model by assuming that parents choose a residential location in part because of income prospects, such as employment opportunities and welfare benefits (if they are eligible), and in part because of the characteristics of the school districts where they choose to live. To estimate the structure of her preferences and the stochastic child achievement production process, we apply semi-parametric maximum likelihood procedures to matched mother-child data from the National Longitudinal Study of Youth of 1979 (NLSY). 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 5 after controlling for the fact that families may be choosing where to live because of the school district characteristics, labor market opportunities, and state welfare policies. The small magnitude of these estimated effects of school quality measures is in line with the findings by Hanushek *el al.* (1998). Note that the regression models in Hanushek *el al.* (1998) do not account for the potential endogeneity of school inputs. In addition, they almost have no control for parental inputs. This may suggest that competing biases offset each other in these cross-sectional models. Furthermore, we find that the impacts on child outcomes of maternal working behaviors to change signs and remain statistically significant after controlling for the possible endogeneity of these decisions. All else equal, full-time employment of the mother is predicted to reduce a child's performance on a mathematics test on average by 6.3 percentile points, while part-time employment reduces it by 4.6 percentile points.

We conduct simulation exercises among a sample of low-educated single women to examine their behavioral responses to a series of exogenous changes in welfare policies and labor market conditions. The simulation results show that poor and low-educated single women with children do change their residential locations in response to changes in welfare policies and labor market conditions. The magnitude of this response in terms of migration, however, is very modest. In particular, reducing a state's welfare benefits level only induces net out-migration by 0.1 percentage point among the sample of low-educated single women, which is equivalent to an increase in the probability of migration by less than 1%. How do these results compare with previous estimates of the magnitude of welfare-induced migration? Levine and Zimmerman (1999) use a natural experiment design, which exploits the categorical eligibility of the welfare programs, and find little evidence for welfare-induced migration. Walker (1994) also uses an experiment design, albeit among residents in contiguous states, and finds that high welfare benefits fail to increase migration propensities among poor young women. Using a model with joint decisions about migration and AFDC participation, in contrast, Gramlich and Laren (1984) and Blank (1988) find substantial migration flows induced by cross-sectional variations in welfare benefits. A shortcoming of these models is that it does not adequately account for endogeneity of migration and other important aspects of migration decisions, including labor market conditions and school quality. More importantly, we find changes in welfare benefits tend to have large impacts on child development outcomes of the at-risk families that migrate in response to welfare magnets. For example, increases in a state's welfare benefits can significantly increase the fraction of in-migrants who newly decide to enter welfare. Similarly, the impacts on the children of those women who would move out of a state in the presence of work requirements are large. On average, using New York as an example, children's math scores would fall by 5.8 percentile points due to changes of their mothers' work and migration decisions, which are induced by exogenous welfare benefit hikes. The estimates of these causal effects on child development indicate that the unintended outcomes of a social program reform could be quite significant. These findings have important implications for the optimal design of social programs as the welfare reform continues in the U.S.

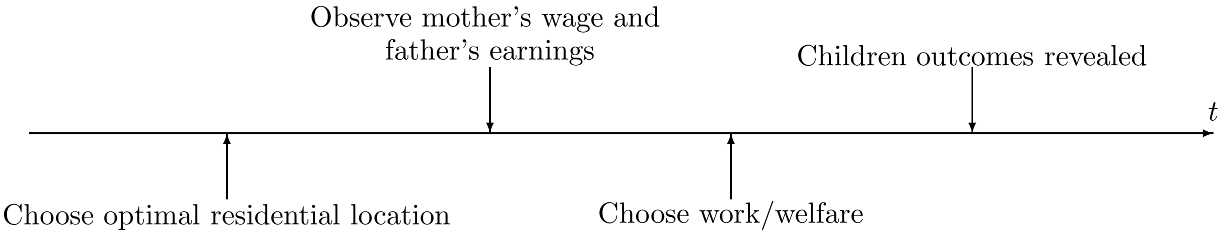
2 Model

This paper extends the migration model developed by Liu, Mroz, and van der Klaauw (2006) to include welfare participation decisions. The migration decision is characterized as part of a two-period stochastic utility optimization problem, in which a mother's hours of work, welfare participation, and place of residence are chosen to maximize the expected household utility subject to a budget constraint incorporating tax, housing cost, and transfer programs. The basic assumption is that parents choose their residential location in part because of income prospects, such as employment opportunities and welfare benefits, and in part because of the impact the characteristics of the school district they choose can have on their children.

2.1 Time Line

In each period a decision is made on where to live, how many hours the mother will work, and whether to participate in welfare. At the beginning of each period t, a household first makes a location decision based on the expected utility of living in each locality, $k = 1, \ldots, K$. Prior to the location decision being made, family incomes and education production function are stochastic. Locations are characterized by a set of school quality measures, local labor market conditions, moving and housing costs, welfare policies, and its geographic region. The family first makes a location decision on the basis of the expected utility of living in each location. After choosing a location of residence, which simultaneously determines her child's consumption of school quality as well as other local economic parameters, the father (if present) receives an earning offer y^s and the mother receives a wage offer w. While the father always is assumed to accept the earnings offered to them in each period, which could be zero or positive amount, the mother makes an employment decision h. Besides employment decisions, the single mothers will also decide whether to participate in the AFDC program if they are eligible. The mother recognizes that the more time spent in the labor market could reduce her time inputs to the education production process. However, the parents cannot completely determine her child's outcome in the sense that they only know the distribution of her child's possible outcomes and how this distribution varies with her work and welfare decisions.

The time line of this model is described as follows.



2.2 Utility function

To simplify notation, we drop all subscripts t in the subsection. Consider a one-period model in which a family i with N_i children has preferences over consumption, x_{ki}, a vector of the children's cognitive achievements, $\mathbf{Q}_i = (q_{i1}, \ldots, q_{iN_i})$, the mother's hours devoted to non-market activities, l_i, the family's choice about AFDC receipt a_i, and the geographic location of residence, k_i. Leisure time, 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 the total time available to the mother. It is therefore equivalent to state preferences in terms of h, instead of l. Let those preferences be represented by a direct utility function as follows

$$U_i = U_i(x_{ki}, \mathbf{Q}_i, h_i, a_i, k_i).$$

Empirically, the utility function is specified as follows

$$\begin{aligned} U^i(x_{ki}, \mathbf{Q}_i, h_i, a_i, k_i) \quad = \quad & \frac{(x_{ki} + \gamma_{0i})^{\gamma_1}}{\gamma_1}[1 + \alpha_0 m_i + \alpha_1 \cdot \mathbf{1}(h_i \neq h_0)] \\ & + \frac{\alpha_2}{\gamma_3} \sum_{n=1}^{N_i} (\frac{q_{in}}{q_i^{\alpha_3}} + \gamma_2)^{\gamma_3} + \sum_{j=1}^{5} [(\alpha_{4ji} + \varepsilon_{ji}) \cdot \mathbf{1}(d_i = j)] \\ & + \alpha_5 a_i + v_{k_{-1}k} + \alpha_6 \cdot \lambda_{ki} + \sum_{j=1}^{4} \alpha_{7j} \cdot \mathbf{1}[R(k_i) = j] + \epsilon_{ki}, \end{aligned} \tag{2}$$

where α's and γ's are parameters to be estimated, m_i is a binary indicator for the marital status of the mother, and $\mathbf{1}(\cdot)$ is an index function. Note that α_0 and α_1 capture the corresponding changes in the marginal utility of consumption when married compared to single and when working compared to not-working respectively. In household i the child n's cognitive test score q_{in} is assumed to be 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 their children's success than lower achieving mothers, and it includes the mother's score being irrelevant (when $\alpha_3 = 0$) as a special case. We also allow hours of work decisions to influence utility discretely through part-time and full-time work utility costs α_{4ji} that depends on employment state $j = 1, \ldots, 5$. ε_{ji} is the family's evaluation of the attributes of the work/welfare status $d_i = j$. In addition, ε_{ji} is assumed to be independently and identically distributed over different hours of work and welfare choices and to follow an Extreme Value distribution with variance parameter b_1. α_5 measures the disutility associated with welfare participation, accounting for non-participation among eligible women. Following Moffitt (1983), this "stigma" term is separable. $v_{k_{-1}k}$, λ_{ki}, and $R(k_i)$ are terms related with one-time moving cost. They will be defined later in Section 2.5.

Additional sources of heterogeneity are introduced into utility function (2). First, We specify the parameters measuring the disutility from work/welfare choice, α_{4ji}, as functions of the number of children between the ages of 0 and 5 (n_y), the number of children between the ages of 6 and 17 (n_o), as well as a time-invariant mother-specific heterogeneity component μ_i^m:

$$\alpha_{4ji} = \begin{cases} 0 & \text{if } j = 1, 2 \\ f_{4\alpha_1}(\mu_i^m, m_i, n_{yi}, n_{oi}) & \text{if } j = 3, 4 \\ f_{4\alpha_2}(\mu_i^m, m_i, n_{yi}, n_{oi}) & \text{if } j = 5, \end{cases}$$

Second, the "reserve" consumption, γ_{0i}, is defined as a second order polynomial function of mother's heterogeneity μ_i^m and father's heterogeneity μ_i^s : $\gamma_{0i} = f_{\gamma_0}(\mu_i^m, \mu_i^s)$.

2.3 Educational production function

The child quality q_{in} is measured by child n's Peabody Individual Achievement Test (PIAT) Mathematics score in household i. The PIAT is among the most widely used academic achievement assessment instruments that have demonstrably high re-test reliability and concurrent validity (Markwardt, 1989). The scores used in this paper are based on age-specific national norms.[2]

Empirically, we adopt a hazard function representation of the child quality production function as described below (Gilleskie and Mroz, 2004). This hazard model assumes that at each point in the range of discretized child outcomes, the child has a particular probability of advancing to a higher level, conditional upon reaching the current level. The hazard specification does not impose arbitrary assumptions about error distributions. Gilleskie and Mroz (2004) have demonstrated that this estimation strategy provides accurate and precise estimates of derivatives of expected outcomes for a wide range of types of explanatory variables. In addition to a more flexible production function specification, the approach is computationally attractive in our setting as the discretization of outcomes implies that the problem of integrating over child outcomes is replaced by the calculation of a weighted sum. The hazard function in this model is defined by a logit function of the sum of a level-varying baseline hazard and a set of covariates. The probability of an "advance" to a higher level, conditional on reaching the current level, p, is given by

$$\Pr(q_{in} > p | q_{in} \geq p) = \frac{\exp\left[f^c_{qin} + f^b_q(p)\right]}{1 + \exp\left[f^c_{qin} + f^b_q(p)\right]}, \tag{3}$$

where q_{in} is the observed score of child n in household i, and the covariate component f^c_{qin} is given by

$$f^c_{qin} = f_q(\mathbf{E}_i, \mathbf{H}_i, \mathbf{S}_{ki}, \mu^c_i), \tag{4}$$

where $\mathbf{E}_i$, $\mathbf{H}_i$, and $\mathbf{S}_{ki}$ are vectors of child's characteristics, mother's characteristics, school characteristics respectively. Detailed description of these variables is given in Section 4. 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). Empirically, the baseline hazard at score p is specified as follows.

$$f^b_q(p) = \sum_{\eta=1}^{3} \{\theta_{q\eta}[\log(P - p)]^\eta\}, \tag{5}$$

where η is the order of polynomials, and P is the highest score level. We have also relaxed several of the separability assumptions in this framework to allow for interactions of baseline hazard, age of the child, and school and parent characteristics by interacting elements of f^c_{qin} and $f^b_q(p)$.

2.4 AFDC/Work decisions and budget constraint

For simplicity, we assume that mother's work choice (h_i) is made in conjunction with welfare participation (a_i). This joint decision, d_i, then is restricted to the following set:

$$d_i = \begin{cases} 1 & : \text{not working } (h_i = h_0) \text{ and not on AFDC } (a_i{=}0) \\ 2 & : \text{not working } (h_i = h_0) \text{ and on AFDC } (a_i{=}1) \\ 3 & : \text{working part-time } (h_i = h_1) \text{ and not on AFDC } (a_i{=}0) \\ 4 & : \text{working part-time } (h_i = h_1) \text{ and on AFDC } (a_i{=}1) \\ 5 & : \text{working full-time } (h_i = h_2) \text{ and not on AFDC } (a_i{=}0). \end{cases}$$

However, in each period women might have to choose from a subset of the above five choices, depending on her circumstances. We further impose that married women do not have the option of receiving AFDC. In practice, married couples were eligible to receive cash assistance under the AFDC-Unemployed Parent (AFDC-UP) program. However, the AFDC-UP program adopted stricter eligibility requirements and, in turn, had smaller enrollments. In the full sample of 10620 person years, only 55 married women, whose spouses were present, received AFDC benefits. The paper focuses on the basic AFDC program and the participation in AFDC-UP is not modeled. In addition, we do not allow women who worked full-time to receive welfare benefits because in the data very few women who worked full-time received AFDC benefits. The eligibility rule of AFDC is further described in the Appendix.

The net consumption is expressed as

$$x_{ki} = w_{ki}h_i + y^s_{ki} + y_{ki} - c_k - \tau_{ki} + a_i B_{ki}, \tag{6}$$

where w_{ki} represents a wage draw at location k, which is unknown to the mother prior to the migration decision being made. For simplicity, we assume that wage rates are the same for a certain individual, no matter whether she works for full-time or part-time. However, she does know the distribution of wage draws in each location k. h_i represents mother's hours spent on working in the labor market during a given year. The empirical model discretizes hours of work (in thousands) into no work ($h_i = h_0$), work part-time ($h_i = h_1$), and work full-time ($h_i = h_2$). A woman is defined as a full-time worker if she works more than 30 hours per week for more than 45 weeks in a year. If she works and is not a full-time worker, she is defined as a part-time worker. y_{ki} denotes non-labor income, which includes household non-labor income and earnings of others in the household i aside from the mother and her husband. c_k explicitly measures the average housing cost in location k. τ_{ki} is the sum of federal and state tax liabilities for individual i, given her marital status, number of dependents, wage income and non-labor income. If individual i participates in AFDC (a_i=1) she could receive AFDC benefit B_{ki}, which is a function of her income, family size, and AFDC rules in location k. Detailed information about the calculation of tax liabilities and AFDC benefits is provided in the Appendix.

Similar to the empirical specification of education production function, hourly wage rates are discretized into G groups, and the probability density function of having a wage offer $\bar{w}_g$ is estimated in the similar fashion as child outcomes. The specification, in particular, is given by

$$\Pr(w_{ki} > \bar{w}_g | w_{ki} \geq \bar{w}_g) = \frac{\exp[f^c_{wi} + f^b_w(g)]}{1 + \exp[f^c_{wi} + f^b_w(g)]}, \tag{7}$$

where

$$f^c_{wi} = f_w(\overline{W_{ki}}, \mathbf{H}_i, \mu^m_i), \tag{8}$$

and

$$f^b_w(g) = \sum_{\eta=1}^{3} \{\theta_{w\eta}[\log(G - g)]^{\eta}\}. \tag{9}$$

$\overline{W_{ki}}$ is the mean wage rate for women with the same education level, ethnicity, and age in each locality, so this formulation allows for a different wage distribution in each locality. The specification of f^c_{wi} also allow for interactions between the baseline hazard f^b_w and maternal characteristics.

The father's earnings are assumed to be stochastic and location specific. The static framework used in this paper also implicitly treats the parents' prior family formation, marital status, fertility, and education as exogenous variables. For simplicity we ignore capital markets by assuming that parents do not save or borrow. To account for the varying living costs across different locations, we use the Cost of Living Index for the American States to construct state-level consumer price indices, which are used to convert all non-housing monetary amounts to 1990 values (See Appendix for details)[3].

Furthermore, the father's earnings distribution is modeled similarly using G' supporting points for y^s, and with f^c_{wi} and f^b_w in equations (7) replaced by (8), and (9)

$$f^c_{yi} = f_y(\overline{W^s_{ki}}, \mathbf{H}^s_i, \mu^s_i), \tag{10}$$

and

$$f_y^b(g') = \sum_{\eta=1}^{3} \{\theta_{y\eta}[\log(G' - g')]^{\eta}\}, \tag{11}$$

where $\overline{W_{ki}^s}$ is the median local wage rate for men with the same educational attainment, race, and age. μ_i^s denotes the fathers unobserved ability endowment. $\mathbf{H}_i^s$ is a vector of husband's characteristics. Higher order terms and interactions between baseline hazard f_y^b and the characteristics of the spouse are also included in f_{yi}^c.

2.5 Migration decision

Given that moving costs are unobserved, to avoid identification problems, we define the moving cost as a disutility component that incorporates both the psychic and monetary costs associated with location changes. We specify two specific moving costs ($v_{k_{-1}k}$ and α_6) that directly affect the family's utility. $v_{k_{-1}k}$ is the explicit moving cost from residential location choice in $t-1$, k_{-1}, to current location choice k, which is approximated by a simple function of distance, i.e.

$$v_{k_{-1}k} = \begin{cases} \phi_0 + \phi_1 \frac{D(k_{-1},k)}{D(k_{-1},k)+1000} & \text{for } k_{-1} \neq k \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where ϕ's are parameters to be estimated and $D(k_{-1}, k)$ is the distance from k_{-1} to k in miles. This specification assumes that the utility cost of moving is increasing at a diminishing rate as the covered distance rises. λ_{ki} is a binary indicator of migration across states, which occurs when k belongs to a state that is different from the residing state in period $t-1$. Therefore, besides the distance-depending cost $v_{k_{-1}k}$, an extra disutility term α_5 is incurred as the family moves across the states. $R(\cdot)$ maps each location k_i into its corresponding census region in the U.S. Therefore, the specification allows for direct geographic preferences for residing in each of four census regions, including Northeast, Midwest, South, and West, which are measured by the parameters $\alpha_{7j}, j = 1, \ldots, 4$ respectively. For identification, α_{71} is normalized into zero.

Maternal evaluation of the unobserved attributes of location k is ϵ_{ki}, which is assumed to be independently and identically distributed over different locations, individuals, and time periods. It is assumed to follow an Extreme Value distribution with variance parameter b_2.

3 Estimation

As described in the previous section, place of residence and work/welfare decisions are made sequentially with a household first making a location decision, followed by the realization of a wage offer and the mother's subsequent work/welfare decision, after which the child's achievement score is realized. When making her work/welfare choice, the mother takes into account the effect of her work/welfare decision as well as of the location-specific school quality level on the expected child outcome. Similarly, when the place of residence decision is being made, the effect of this decision on subsequent work/welfare decisions and the effect of both on the expected child outcome are taken into account. The solution to this stochastic optimization problem can be obtained by solving backwards: first, making the optimal work decision for each possible location choice and wage offer; second, making the optimal location choice. To simplify notation, the subscripts i and t are dropped.

If the mother ends up in location k, with the spouse earning y_k^s she receives wage draw w_k and learns the utility shocks associated with each work/welfare decision, i.e. $\varepsilon_d = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_5)$. The expected utility associated with each work/welfare choice can be defined as

$$V_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu, \varepsilon_d, \epsilon_k) = E\{U[x(d), \mathbf{Q}(d), d, k] | \mathbf{\Omega}_k, w_k, y_k^s, \mu, \varepsilon_d, \epsilon_k\}, \tag{13}$$

where $\mathbf{\Omega}_k$ is an information set, $\mathbf{\Omega}_k = \{\mathbf{E}, \mathbf{H}, \mathbf{S}_k, k_{-1}\}$. $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{S}_k$ are vectors of child's characteristics, mother's characteristics, grade-specific school characteristics in location k respectively. μ is the heterogeneity vector, $\mu = \{\mu^m, \mu^c, \mu^s\}$. $x(d)$ and $\mathbf{Q}(d)$ are the values of consumption and the children's achievement scores respectively when the mother chooses work/welfare decision set d.

This is an expected maximization problem because the child's educational outcome is stochastic (though influenced by work/welfare choices). For each work/welfare choice, the expected utility, after integrating over all possible test score outcomes, is given by

$$\begin{aligned} V_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu, \varepsilon_d, \epsilon_k) &= \sum_{n=1}^{N} \sum_{p=1}^{P} \{\Pr(q_n = p | \mathbf{\Omega}_k, w_k, y_k^s, \mu)\, U[x(d), q(d), d, k]\} \\ &= \overline{V}_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu) + \varepsilon_d + \epsilon_k \end{aligned} \tag{14}$$

where the second equality follows the additive separability of the utility function (2) and assumed independence (conditional on μ) of both error terms $(\varepsilon_d, \epsilon_k)$.

Given location choice k and wage rate offer w_k the optimal work/welfare decision d^* can be defined as

$$d^* = argmax\{\overline{V}_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu) + \varepsilon_d\}, \tag{15}$$

The probability of household i making work/welfare choice d^* conditional on wage and husband earning draws w_k and y_k^s to the researcher is given by

$$\Pr(d_{ki} = d^* | \mathbf{\Omega}_{ki}, w_{ki}, y_{ki}^s, \mu_i) = \frac{\exp\left[\frac{\overline{V}_k^{d^*}(\mathbf{\Omega}_{ki}, w_{ki}, y_{ki}^s, \mu_i)}{b_1}\right]}{\sum\limits_{d'=1}^{5} \exp\left[\frac{\overline{V}_k^{d'}(\mathbf{\Omega}_{ki}, w_{ki}, y_{ki}^s, \mu_i)}{b_1}\right]}. \tag{16}$$

When the location decision is being made, the woman does not know the actual wage offer at each location, nor does she know the realizations of the utility shocks associated with her future work/welfare decisions. For any given wage offer in location k, with ε_d being *i.i.d.* Extreme Value errors, the value function conditional on location choice k and wage offer w_k is given by

$$\begin{aligned} V_k(\mathbf{\Omega}_k, w_k, y_k^s, \mu) &= \underset{d}{EMAX}[V_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu, \varepsilon_d, \epsilon_k)] \\ &= b_1 \log\left\{\sum_{d=1}^{5}\left[\exp\left(\frac{\overline{V}_k^d(\mathbf{\Omega}_k, w_k, y_k^s, \mu)}{b_1}\right)\right]\right\} + \epsilon_k \\ &= \overline{V}_k(\mathbf{\Omega}_k, w_k, y_k^s, \mu) + \epsilon_k, \end{aligned} \tag{17}$$

where b_1 is a parameter defining the variance of the Extreme Value distribution and the equality follows from the additive separability of the utility function in both errors, combined with the assumed independence (conditional on μ) of ε_d and ϵ_k.

When making a residential decision, the wage offer and husband earning in each possible destination k are unknown. Integrating the expected utilities defined in equation (17) over G possible wage offers $(\bar{w}_g, g = 1, \ldots, G)$ and G' possible realizations of husband earnings $(\bar{y}_{g'}^s, g' = 1, \ldots, G')$ in location k yields

$$\begin{aligned} EV_k(\mathbf{\Omega}_k, \mu, \epsilon_k) &= \sum_{g=1}^{G}\sum_{g'=1}^{G'}\left[\Pr(w_k = \bar{w}_g) \cdot \Pr(y_k^s = \bar{y}_{g'}^s) \cdot \overline{V}_k(\mathbf{\Omega}_k, w_k, y_k^s, \mu)\right] + \epsilon_k \\ &= \widetilde{V}_k(\mathbf{\Omega}_k, \mu) + \epsilon_k. \end{aligned} \tag{18}$$

At this time, we assume the agent knows ϵ_k for all locations (but not the utility shocks, ε_d, which are associated with work/welfare decisions within locations), so she chooses

$$k^* = argmax\{\widetilde{V}_k(\mathbf{\Omega}_k, \mu) + \epsilon_k\}. \tag{19}$$

Assuming *i.i.d.* Extreme Value distributed ϵ_k, the probability of the mother in household i choosing location k^* to the researcher is given by

$$\Pr(k_i = k^* | \mathbf{\Omega}_i, \mu_i) = \frac{\exp\left[\frac{\widetilde{V}_{k^*}(\mathbf{\Omega}_{k^*i}, \mu_i)}{b_2}\right]}{\sum_{k'=1}^{K} \exp\left[\frac{\widetilde{V}_{k'}(\mathbf{\Omega}_{k'i}, \mu_i)}{b_2}\right]}, \tag{20}$$

where $\mathbf{\Omega}_i = \{(\mathbf{\Omega}_{ki}, \mathbf{M}_{ki}); k = 1, \ldots, K\}$, and $\mathbf{M}_{ki}$ represents the local labor market conditions at location k for parents of household i.

Unfortunately, direct estimation of (20) is computationally costly because the choice set for locations is very large. In particular, this location choice set consists of 3141 counties in the United States. However, it is still possible to estimate the model consistently. The IIA property implies that the conditional probabilities of choosing from a subset of the full choice set equal the choice probabilities when the choice set equals the subset. McFadden (1978) shows that the full choice set can be replaced with the chosen and a probability sample of the non-chosen alternatives. In addition, it is necessary to include weighting terms unless the sampling method is simple random sampling. Let D_i denote the subset of location choices for household i and $\pi(D_i|k^*)$ represent the probability of the set D_i being drawn conditional on k^* is the chosen alternative.

Hence, equation (20) is re-written as follows

$$\Pr(k_i = k^* | D_i, \mathbf{\Omega}_i, \mu_i) = \frac{\exp\left[\frac{\widetilde{V}_{k^*}(\mathbf{\Omega}_{k^*i}, \mu_i)}{b_2} + \ln \pi(D_i|k^*)\right]}{\sum_{k' \in D_i} \exp\left[\frac{\widetilde{V}_{k'}(\mathbf{\Omega}_{k'i}, \mu_i)}{b_2} + \ln \pi(D_i|k')\right]}. \tag{21}$$

More details about the sampling scheme are discussed in the Appendix.

3.1 Initial Residential Locations

The literature on welfare migration indicates that states that provide high benefits in one period often provide high benefits in subsequent periods. A cross-section study without adequately controlling for the endogeneity of current location might yield downward biased estimates. On the other hand, a longitudinal study can take advantage of the fact that there exist "exogenous" variations in benefits within states over time. These variations can be used to address the endogeneity of residential choices. Nevertheless, the location in which the household is initially observed when the oldest child in the family reaches age 5 or 6 may still be endogenous. Following Heckman (1981), we modify the empirical model by taking into account the transition between the mother's birth place, which is assumed to be exogenous, and initially observed location in the mother-child data. For the first period the family's residential location is observed ($t = t_{0i}$), we estimate a conditional logit model that includes interactions among location-specific characteristics, family background characteristics, unobserved heterogeneity terms, and dummies for different types of geographical movements. Given the magnitude of choice set in terms of place of residence, we use the random sampling scheme described in Section 3. Let the value, for household i, of choosing location k at the first period $t = t_{0i}$ be specified as follows

$$V_k(\mathbf{Z}_{ir_i}, \mathbf{S}_{kt_{0i}}, \mu_i, e_{kit_{0i}}) = \widehat{V}(\mathbf{Z}_{ir_i}, \mathbf{S}_{kt_{0i}}, \mu_i) + \xi_{kit_{0i}},$$

where period r_i corresponds to the mother's birth year for household i, $\mathbf{Z}_{ir_i}$ denotes the mother's family background characteristics at her birth, and $\xi_{kit_{0i}}$ follows a Type I Extreme Value distribution. $\widehat{V}(\mathbf{Z}_{ir_i}, \mathbf{S}_{kt_{0i}}, \mu_i)$ is a function of location-specific characteristics of the "initial location" k, the mother's family background characteristics, dummies for different types of movements, and unobserved heterogeneity term specific to the household. Location variables include the dropout rate of the 11^{th} graders, expenditure per pupil, local average wage rate for a woman whose age, race, and educational attainment are equal to those of the mother of household i, AFDC maximum amount paid to a family with the same size in the state that k belongs to, tax liabilities for a household whose earned income is equal to the nation median value. Family characteristics include mother's education, race, and age and maternal grandmother's education. We specify three specific geographic movements, $\lambda_{kit_{0i}}$, which includes moves within the state where the mother was born, moves outside of her birth state but within her birth region, and moves outside of her birth region. The probability of observing household i in location k^* in the first period can be written as follows

$$\Pr(k_{it_{0i}} = k^* | \mathbf{\Omega}_{it_{0i}}, \mu_i) = \frac{\exp[\widehat{V}(\mathbf{Z}_{ir_i}, \mathbf{S}_{k^*t_{0i}}, \mu_i)]}{\sum_{k' \in D_{it_{0i}}} \exp[\widehat{V}(\mathbf{Z}_{ir_i}, \mathbf{S}_{k't_{0i}}, \mu_i)]}. \tag{22}$$

Because of the nature of the conditional logit model, we can only identify terms that are choice-specific and we also normalize the coefficient associated with the dummy variable for moving within a state to -1.

3.2 Likelihood function

The empirical model estimated in this study incorporates longitudinal observations of the residential places, the mother's working decision and wage (if working), the husband's earnings (if married), the family welfare participation status, the children's achievement scores for each household i in the NLSY sample. After deciding to reside in location k^*_{it} at the beginning of each period t, the wage draw $w^*_{k^*it}$ and the husband's earnings $y^{s*}_{k^*it}$ at the chosen location are observed by the researcher. We do not observe any wage information when the woman chooses not to work. For these non-workers ($h^*_{it} = h_0$), it is necessary to integrate over all possible wage offers. Furthermore, the maximum likelihood framework used in this paper accounts for correlated person and family specific unobservable components in mother's wages, husband's earnings, work decisions, welfare participation decisions, child achievement scores, and place of residence decisions. This heterogeneity is modeled as constant throughout the child's years of observation. Given the optimal decision rules (16), (21) and (22), and assuming that heterogeneity factor μ has a discrete distribution taking on J different values with corresponding probabilities $\Pr(\mu_j)$, $j = 1, \ldots, J$, the likelihood function for household i is given by

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

The sample likelihood function then is given by

$$
(24) \qquad L = \prod_{i=1}^{N} L_i .
$$

3.3 Model identification

The main components in the structural model developed in this paper are: (a) the education production function with two types of endogenous inputs: maternal work and school inputs; (b) the equation concerning mother's employment and welfare participation decision; and (c) the residential location choice, which also corresponds to the school input equation. The empirical model in this paper addresses the endogeneity of a household's initial place of residence when their oldest child was age 5 or 6, utilizing the exogenous variation of local characteristics between the mother's birth year and the beginning of the collection of the Child Supplement data in 1986. Therefore, identification of the impact of parental and school inputs on child outcomes first relies on the assumption that the mother's birth place is exogenous to the subsequent residence locations after t_{0i}, work decisions, welfare decisions, and child outcomes. The identification is partly due to the considerable cost associated with migration, which depends on the distance between the current and next location, local labor market conditions, welfare policy parameters, local tax laws, local living costs, and school characteristics in any given period. Additionally, there are stochastic, age-specific exogenous change over time in the distribution of local labor market conditions, local living costs, local tax and welfare policies in each location, and school quality measures in each location. For instance, those who face relatively improved labor market conditions or welfare benefits at the current location will be less likely to move. These time-varying situations in the data may drive different sets of decisions and, therefore, facilitate the identification of our models. In addition to the cross-sectional variation in the familys initial location and time-variation in relevant policy parameters for a fixed geographic location, the treatment of time-varying marital status and fertility outcomes as exogenous provides an additional source of identification. In future work, however, it would be important to explore relaxing these exogeneity assumptions[4].

The panel data employed in this study provide a multitude of identification conditions inferred by the number of contemporaneous exogenous variables (e.g. instruments) excluded from the structural equation of interest. For instance, every lag of each exogenous variable could have an indirect effect on the "contemporaneous" value of the endogenous explanatory variables (see Bhargava, 1991). In solving the value functions implied by the structural equations, additional "instruments" associated with each interaction of these "exclusion restrictions" over time may result in significantly more variables to control for endogeneity. In this study the exogenous time-varying variables include the mothers', children's, and fathers' ages, marital status, 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. In the context of dynamic nonlinear models, the maximum likelihood estimator used in this paper allows us to incorporate these interactions among the sets of exogenous variables. These time-varying exogenous factors interact with each observation's initial place of residence to generate numerous exclusion restrictions.

4 Data

The primary data source is the Geocode version of the NLSY79 data set and its Child-Mother Supplement. The Child-Mother Supplement collects detailed information about the young women's children every other year starting in 1986. This study excludes the disadvantaged non-Hispanic white over-samples, because their selection was based on potentially endogenous characteristics. The black and Hispanic over-samples are used, and race indicators are included in all equations to account for different preferences and opportunities. The analysis is further focused on the families of female respondents in NLSY who had at least one child between ages 5 and 15 in 1986 or later. The unit of observation in this study is a household for up to six possible time-period specific observations between 1986 and 1996. Each time period in the empirical model is equivalent to a two-year interval. Because of the structure of the data set, children in the data can have up to six possible time-period specific observations. Subsequently, 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. The primary child outcome of interest is the child's percentile score on a Mathematics aptitude test, PIAT Mathematics assessment. In the NLSY79 Child-Mother Supplement, PIAT assessments are administered to all age-eligible children, who are 5 years and older. In particular, the PIAT Mathematics assessment is one of the most reliable child assessment instruments used in the Child-Mother supplement. The median local wages based on 1990 Census data are used as indicator of local labor demand conditions. Tables 1a and1b contain summary statistics for the data on parents and children respectively. Note that all monetary terms are deflated to 1990 dollars using the state-level consumer price indices that are based on the Cost of Living Index for the American States((Berry *et al.*, 2000)).

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's score "moves" from one cell to the next highest cell through the logit hazard model described in Section 2.2. The child's score depends upon characteristics of the child such as its 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, and AFQT score. This production function also depends upon characteristics of the school district where the family resides, such as the dropout rate for 11^{th} graders, average teacher salary as a fraction of the average annual earnings of college educated persons age 27 to 40, per child school district expenditures as a fraction of the average earnings of individuals in the school districts, and the teacher/student ratio. Extracted from a variety of data sources (see the Appendix for details), these school variables are measured at the county level, which is the smallest geographic unit available in the NLSY data, and are grade specific with an exception of measurement of school expenditures. For instance, the teacher/pupil ratios are weighted averages across all public schools for every grade level of each county based on the data from the National Center of Education Statistics (NCES): Common Core Data (CCD). In particular, we calculate the means for all these characteristics for each grade between K-12 using CCD. The grade-specific school information is used for each child in the NLSY sample. For instance, a 14-year old child will be matched with school characteristics at the 9th grade.

Table 1c provides economic and demographic characteristics of the sample by employment and welfare status. It shows considerable differences between recipients and non-recipients. Women receiving AFDC benefits tended to be younger, less likely to work, and less educated. For instance, among AFDC-eligible women, 65.9% of the welfare recipients did not work while only 13.1% of the women who were not on AFDC left the labor market. However, this should not be interpreted as the disincentive effect of AFDC because women were likely to be self-selected in the welfare receipt group.

To describe the wage offers available to the individual women in each locality k at t, w_{kit}, we model the probability that each woman has an offered wage from each of the five discretized wage rates (i.e. $w_{kit} = \bar{w}_{1t}, \bar{w}_{2t}, \ldots, \bar{w}_{5t}$). As in the child production function model, we use the 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 residing county 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 and ethnicity effects, effects of the mother's AFQT score, and effects of the heterogeneity type on the probabilities of each of these discrete wage offers. Estimates for the wage equations are presented in Appendix Table 2c.

A similar approach is used to specify the father's yearly earnings equation with 6 discretized points of support, which include zero earnings. In addition to local median male yearly earnings in the county, the earning equation also includes the father's race, and his education level as independent variables.

5 Estimation Results

Appendix Table 2 contains the point estimates from our empirical model. Given the complex interactions among most of the covariates in the economic model, it is quite difficult to give a simple interpretation to each of the point estimates. For instance, the interpretation of the parameter estimate of the impact of a mother holding a high school degree on the production function presented in Appendix Table 2a is not as straightforward as expected. The interpretation of this parameter is how a woman being a high school graduate instead of a dropout affects her children's propensity to move to a higher Mathematics score category holding everything else constant. The specification of the production function allows most independent variables to interact with the score level and the child age, so the complete effect of those covariates on her child's educational production process can be much more complicated than this.

To make an inspection of these impacts possible, we simulate outcomes using the estimated parameters. Three sets of estimates of marginal effects of characteristics on the child's PIAT Mathematics test score, ceteris paribus, are reported in Table 2. The first column contains OLS estimates. columns 2 and 3 contain estimates based upon the hazard model for the child scores with controls for unobserved heterogeneity. We simulate how expected test scores would change in response to varying explanatory variables one at a time, normalized to a unit change in the characteristics 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 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 use in the structural model. The estimates used to define the marginal effects in column 3 were estimated as part of the structural model that incorporates the endogeneity of the location decisions and the endogeneity of the mothers' hours of work decisions. standard errors of the marginal effects are calculated using a parametric bootstrap (100 replications) simulated using the estimated covariance matrix of the parameter estimators.

It is noteworthy that without endogeneity controls, a mother's working part-time and working full-time appear to increase her child's percentile score by 2.7 and 2.1 respectively. After controlling for endogeneity the estimated impacts imply that a mother's working part-time and working full-time, as opposed to not working, reduce the expected percentile score by 4.6 point and 6.3 points respectively. Additionally, the impacts of the school quality measures diminish by up to a factor of 2 after accounting for the fact that families may select the residential location based on the school characteristics and labor market opportunities. One interpretation of these changes 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 while at the same time working more. This is a standard endogeneity of inputs argument.

5.1 Simulations within the Sample and Tests of Goodness of Fit

Using parameter estimates from our full empirical model, we are able to simulate the migration, work, child development, and welfare participation outcomes within the sample. Table 3a compares the observed outcomes with the simulated behaviors using the model estimates and observed exogenous variables within the sample. In general, the model predicts the means of these important outcomes quite well. Furthermore, several simulation exercises are carried out to inspect the responsiveness of individuals within the sample to four policy changes.

The baseline row in Table 3b shows the simulated outcomes using the actual policy parameters from the data, such as tax laws, AFDC rules, wage and earnings distributions, and school characteristics. In this baseline simulation, 13.8% and 74.7% of the women received AFDC benefits and worked respectively. 59.4% of working women worked full-time. In the second row of Table 3b, simulation results are presented for increasing AFDC benefits in each locality. In this case, the proportion of women who participated in AFDC increased by 0.9 percentage point while labor force participation among women fell by only 1.2 percentage points. Given the fact that the fraction of AFDC-eligible women in the full sample was small and a considerable portion of these eligible women were already recipients, the change in AFDC caseloads was small relative to the whole population. These changes in welfare participation and labor supply, however, were considerable relative to the size of AFDC-eligible non-recipients. The mean score of their children did not appear to have changed as a result of increasing benefits. This appears to be due to the fact that there were offsetting changes in residential location choices and labor supply decisions.

The rest of the contents in this table show results after imposing part-time work requirements and increasing wage rates. In particular, requiring all AFDC participants to work dramatically reduced the AFDC enrollment rate by 4.2 percentage points. A significant increase in labor force participation, 4.4 percentage points, also occurred as an outcome of this policy change. Most of these changes appear to be due to women shifting AFDC to full-time employment. Increasing wage rates by 20% in each location reduced the AFDC enrollment rate by 0.8 percentage point and increased the percentage of workers by 2.3 percentage points. As wage rates increase, a fraction of welfare women shifted from AFDC to full-time employment and many non-recipients also shifted from not working to full-time employment.

5.2 Simulations of Policy Changes on a Sample of High School Dropouts

Using the parameter estimates from the full structural model and the corresponding estimated optimal decision rules, we are able to simulate women's migration, welfare, and work decisions, as well as child outcomes for an observed household. Specifically, we first simulate baseline outcomes for a national sample of single women who were 33-year-old high school dropouts in 1988 (period 0). Each woman had an AFQT score of 0.4, one female child, and no non-labor income. In addition, the number of observations in each county of this sample is in proportion to the county's share of the U.S. population, and the number of observations in the state of New York is normalized to 10,000. We then simulate for three more periods(1990, 1992, and 1994) the choices of these women, as well as their children's math scores. Note that each period is equal to two calendar years for reasons described in the data section. After this step, we further focus on what happens to the sub-samples of women living in New Work in both 1988 and 1990, and consider policy experiments that target only New York and Massachusetts – one of its neighboring states . We refer to this group of women as "New York sample." We also refer to the welfare participants in the New York sample as "New York AFDC sample." Conceptually, the main differences between the

New York sample and the New York AFDC sample are that women in the latter tended to have tastes in favor of welfare benefits while women in the New York sample represented a more general group of women. The details of this exercise are provided in the Appendix.

A. Simulations of Changes in AFDC Maximum Benefits

First we reduce the maximum AFDC benefits in New York by 30%. The corresponding outcomes on the New York sample and the New York AFDC sample are presented in Tables 4a-c and Tables 5a-c respectively. In this case, the out-migration rate in the New York sample increased from 15.4% to 15.5%, indicating an induced out-migration of 0.13 percentage point, with a standard error of 0.12. Nevertheless, the impact of this policy change on women who stayed in New York was substantial, considering that only 22.6% of them received AFDC benefits and more than 72.6% were working. In the AFDC sample, we find the similar pattern except that women in this sample were slightly more likely to move out when the benefits in New York were reduced. Intuitively, as discussed above, women in the AFDC sample were more likely to move for better welfare benefits.

In the other experiment, AFDC benefit levels were increased by 30% in Massachusetts and the induced out-migration was even smaller (see the third rows in Table 4a and Table 5a). In the New York sample, the additional out-migration was 0.04 percentage point with a standard error of 0.04 while in the AFDC sample the induced change was 0.07 percentage point with a standard error of 0.03. A noteworthy outcome is demonstrated in this scenario. From Massachusetts's viewpoint, even though increasing AFDC benefits hardly induced new in-migrants crossing its state border, the New York residents who moved to Massachusetts tend to change their work and welfare behaviors. Specifically, only 1.07% of New York residents moved into Massachusetts and received benefits in the absence of benefit increase. In the presence of the benefit increase, more than 1.23% of New York residents migrated to Massachusetts and received benefits. This represented a 15% increase in the welfare participation rate among in-migrants from Massachusetts's perspective. Their welfare participation might merely be an incidental outcome associated with migration as the same portion of New York residents would move to Massachusetts even in the absence of benefit hike. One possible explanation is that these in-migrants who newly took AFDC benefits in Massachusetts moved for reasons other than better benefits.

B. Simulations of Changes in AFDC Work Requirements

The hypothetical policy change in this group of experiments is to impose a work requirement for all AFDC recipients in either New York or Massachusetts. Specifically, the choice of "not working and on AFDC" ($d_i = 2$) is eliminated. Imposing work requirements in New York had larger impacts on women's migration, welfare, and labor supply behaviors than other policy changes reported in this section. Specifically, the induced out-migration was 0.4 and 0.9 percentage point in the New York sample and the AFDC sample respectively. This induced out-migration, however, was still modest in absolute terms. More interesting results come from the labor market and welfare participation. Using the New York sample, we find that the fraction of women who received benefits and stayed in New York in 1994 declined from 21.5% in the baseline level to 14.6% after the imposition of work requirements in New York. The number of women who stayed in New York and worked increased by nearly 7 percentage points. These outcomes appear to be due to women shifting from welfare to full-time employment.

More importantly, using the New York sample we find that the impacts on the children of those women who would move out of New York in the presence of work requirements were large. A group of women were identified who would have stayed in New York in the absence of work requirements but moved out as the requirements were imposed. In addition to changing their place of residence, these women also adjusted their welfare and labor supply decisions accordingly. A fraction of these women received better wage offers in their new place of residence and left AFDC to start working full-time while another sub-group of women received worse wage offers in their new destinations and shifted from part-time work to not working. Consequently, the probability of the women receiving benefits declined from 45.2% to 36.4% and the probability of the women working full-time, as opposed to part-time, increased from 52.3% to 64.5%. Given the finding that the more time spent in the labor market will reduce the direct time inputs in education production, the children whose mothers newly shifted from not working to working full-time might suffer considerably in terms of their achievement outcomes, holding other educational inputs constant. Moving out of New York might further hurt these children as they migrated into school districts with worse qualities than those in New York. On average, as illustrated in Figure 1, their children's achievement test scores fell by 5.8 percentile points because of their mothers' new relocation and labor supply decisions.

C. Simulations of Changes in the Labor Market

To study the impacts of a change in the labor market, we increase the discretized wage rates by 20% in either New York or Massachusetts.

In the New York sample, we find women's welfare participation modestly responded to the wage raise. More women, in addition, started working and working full-time. In the AFDC sample, the welfare enrollment rate dropped less than 1 percentage point, while responses in the labor market were much smaller than those in the New York sample. An interpretation of the different responses in labor supply is that women in AFDC sample tended to have worse wage draws and/or bad tastes for work.

6 Conclusions

This paper provides a promising approach for analyzing how a woman's migration, welfare, and labor supply behaviors respond to changes in welfare policies and labor market conditions, and how these behaviors in turn influence her child's achievement outcomes. The results in this paper have shown that through altering parental migration and work behaviors these welfare policy changes have important impacts on at-risk children's achievement outcomes.

A limitation of current structural approaches in the welfare migration literature is that specifying an incomplete choice set faced by individuals might omit other important migration determinants and, in turn, produce biased estimates. Estimating the educational production function as part of a structural model significantly enriches the choice set faced by single mothers who have to make migration decisions for their families. Modeling work and welfare participation decisions further helps identify different incentives for migration.

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 5 after controlling for the fact that families may be choosing where to live because of the school district characteristics and labor market opportunities. This study also finds that the impacts on child outcomes of maternal working behaviors 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.

Simulations conducted by this study permit a focused view of women's behavioral responses to a series of changes in welfare policies and labor market conditions. This paper finds that the migration directly induced by welfare differentials is modest as a component of the general migration flow. In practice, people migrate for economic reasons as well as non-economic reasons. Nevertheless, poor and low-educated single women with children do change their residential locations in response to changes in welfare policies and labor market conditions. In several simulation experiments, the induced out-migration is small but statistically significant. More important are the impacts on the groups affected by policies. Those at-risk often are significantly affected. For example, increases in a state's welfare benefits can significantly increase the fraction of in-migrants who newly decide to enter welfare. Similarly, the impacts on the children of those women who would move out of a state in the presence of work requirements are large. On average, using New York as an example, their children's achievement test scores would fall by 5.8 percentile points because of their mothers' new relocation and labor supply decisions. This finding has important policy implications. Any policy reform aimed at influencing women's labor supply or welfare participation behaviors may have important effects on their children's well-being. Studies to date on the effects of welfare reforms have focused on the work (de)incentives of welfare programs or other behaviors of females. A study of the impacts of the policy reforms on children's educational outcomes will be quite important given the significant role of women in the education process of their children.

Table 1: Summary Statistics of the Sample Data

Table 1a: Parents

	Mean (Standard Deviation)						
Variable	1986	1988	1990	1992	1994	1996	Overall
Age of mother	26.540	28.543	30.060	31.800	33.600	35.452	32.594
	(1.941)	(1.938)	(2.116)	(2.178)	(2.188)	(2.183)	(4.003)
Married	0.509	0.507	0.539	0.568	0.584	0.602	0.570
	(0.500)	(0.500)	(0.499)	(0.495)	(0.493)	(0.490)	(0.495)
High school	0.516	0.525	0.534	0.525	0.494	0.468	0.497
	(0.500)	(0.500)	(0.499)	(0.500)	(0.500)	(0.499)	(0.500)
More than high school	0.159	0.182	0.243	0.293	0.342	0.393	0.313
	(0.366)	(0.386)	(0.429)	(0.455)	(0.475)	(0.489)	(0.464)
Non-white	0.647	0.646	0.617	0.576	0.564	0.538	0.575
	(0.478)	(0.478)	(0.486)	(0.494)	(0.496)	(0.499)	(0.494)
Mother's AFQT score	0.354	0.357	0.403	0.433	0.448	0.472	0.433
	(0.267)	(0.268)	(0.271)	(0.273)	(0.276)	(0.279)	(0.279)
Yearly housing cost$^{(1)}$	0.502	0.503	0.505	0.508	0.511	0.515	0.510
	(0.125)	(0.120)	(0.113)	(0.118)	(0.122)	(0.127)	(0.122)
Move	0.135	0.164	0.143	0.067	0.152	0.091	0.115
	(0.342)	(0.371)	(0.350)	(0.250)	(0.359)	(0.288)	(0.319)
Net income$^{(1)}$	1.273	1.388	1.781	2.095	2.142	2.634	2.163
	(1.890)	(1.768)	(2.020)	(2.519)	(2.357)	(2.868)	(2.546)
Mother's hourly wage rate$^{(2)}$	0.670	746	0.788	0.810	0.946	0.972	0.880
	(0.665)	(0.387)	(0.538)	(0.898)	(0.497)	(0.778)	(0.710)
Father's yearly earnings$^{(1)}$	1.848	1.972	2.296	2.380	2.555	2.880	2.573
	(1.372)	(1.474)	(1.666)	(1.842)	(1.953)	(2.143)	(1.998)
Local wage rate for women$^{(2)}$	0.823	0.839	0.927	0.987	1.031	1.094	1.001
	(0.388)	(0.348)	(0.377)	(0.428)	(0.409)	(0.438)	(0.422)
Part-time work$^{(3)}$	0.353	0.323	0.329	0.312	0.289	0.271	0.304
	(0.478)	(0.468)	(0.470)	(0.464)	(0.453)	(0.445)	(0.460)
Full-time work$^{(3)}$	0.298	0.387	0.416	0.425	0.462	0.499	0.444
	(0.457)	(0.487)	(0.493)	(0.495)	(0.499)	(0.500)	(0.497)
# of young children in HH(0-5)	1.134	0.710	0.652	0.647	0.546	0.437	0.604
	(0.945)	(0.803)	(0.781)	(0.784)	(0.738)	(0.691)	(0.788)
# of old children in HH(6-17)	1.127	1.613	1.641	1.828	1.946	2.002	1.793
	(0.933)	(0.990)	(0.981)	(0.982)	(0.937)	(0.945)	(0.996)
AFDC participation rate	0.218	0.194	0.158	0.159	0.145	0.101	0.138
	(0.413)	(0.395)	(0.364)	(0.366)	(0.352)	(0.301)	(0.345)
Sample size	1075	992	1409	1688	1858	1859	8881

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 workers.

Table 1b: Children

	Mean (Standard Deviation)						
Variable	1986	1988	1990	1992	1994	1996	Overall
Age of child	7.608	9.072	9.389	9.805	9.938	10.063	9.631
	(2.008)	(2.323)	(2.553)	(2.656)	(2.617)	(2.615))	(2.625)
Non-white	0.663	0.656	0.629	0.588	0.570	0.543	0.582
	(0.473)	(0.475)	(0.483)	(0.492)	(0.495)	(0.498)	(0.493)
Boy	0.528	0.515	0.504	0.499	0.498	0.502	0.506
	(0.499)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Dropout rate	0.064	0.063	0.063	0.063	0.063	0.062	0.063
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)
Expenditure per pupil$^{(1)}$	0.126	0.124	0.124	0.124	0.124	0.124	0.124
	(0.039)	(0.038)	(0.038)	(0.038)	(0.039)	(0.040)	(0.039)
Teacher salary$^{(1)}$	0.784	0.794	0.797	0.797	0.800	0.801	0.797
	(0.174)	(0.175)	(0.176)	(0.175)	(0.171)	(0.175)	(0.175)
Teacher student ratio	0.056	0.056	0.055	0.055	0.056	0.056	0.056
	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
PIAT Math score	0.451	0.429	0.445	0.461	0.477	0.511	0.476
	(0.259)	(0.252)	(0.258)	(0.261)	(0.266)	(0.272)	(0.266)
Sample size	1493	1796	2515	3056	3300	3202	15362

Notes: 1. Measured as proportion of yearly income of college-graduated prime-aged males(see the Appendix).

Table 1c: Distribution of Labor Supply and Welfare Participation

		Eligible	
	Not Eligible	No AFDC	AFDC
Non-workers			
Counts	1377	343	963
Row percent	51.32%	12.78%	35.90%
Column percent	21.06%	13.10%	65.87%
Mean age	32.45 (4.05)	32.5 (4.25)	30.93 (3.89)
Mean non-labor income($10,000's)	2.78 (2.59)	0.05 (0.11)	0.04 (0.09)
High school completion	0.72 (0.45)	0.58 (0.49)	0.57 (0.50)
Mean local hourly wage rate	10.26 (4.18)	9.70 (3.85)	9.57 (3.86)
Mean child math score	0.48 (0.24)	0.38 (0.23)	0.36 (0.22)
Part-time workers			
Counts	2108	617	499
Row percent	65.38%	19.14%	15.48%
Column percent	32.24%	23.56%	34.13%
Mean age	32.75 (4.04)	32.1 (4.10)	30.81 (3.89)
Mean non-labor income($10,000's)	2.76 (2.23)	0.08 (0.13)	0.05 (0.09)
High school completion	0.85 (0.36)	0.75 (0.43)	0.69 (0.46)
Mean local hourly wage rate	9.77 (3.96)	9.27 (4.63)	9.02 (4.35)
Mean child math score	0.54 (0.23)	0.43 (0.23)	0.42 (0.20)
Full-time workers			
Counts	3054	1659	-
Row percent	64.80%	35.20%	-
Column percent	46.70%	63.34%	-
Mean age	33.4 (3.74)	32.8 (3.91)	-
Mean non-labor income($10,000's)	2.37 (1.85)	0.08 (0.12)	-
High school completion	0.91 (0.28)	0.88 (0.32)	-
Mean local hourly wage rate	10.45 (4.38)	9.97 (4.24)	-
Mean child math score	0.53 (0.23)	0.47 (0.23)	-

Table 2: Production Function Estimates of Marginal Effects with Comparisons to OLS and "Hazard" models without Endogeneity Controls

	Marginal Effects[1]		
Variable	OLS	Production Function Only	Full Model
Exogenous Variables			
Age of mother	−0.0263***	−0.0532***	−0.0332***
	(0.0052)	(0.0092)	(0.0045)
Age of child	0.0200***	0.0244***	0.0180***
	(0.0072)	(0.0084)	(0.0051)
Married	0.0268***	0.0169***	0.0168***
	(0.0038)	(0.0037)	(0.0049)
High school (mother)	0.0341***	0.0325***	0.0315***
	(0.0050)	(0.0036)	(0.0028)
More than high school	0.0616***	0.0517***	0.0514***
	(0.0062)	(0.0052)	(0.0041)
None-white	−0.0557***	−0.0491***	−0.0576***
	(0.0046)	(0.0038)	(0.0026)
Boy	−0.0010	0.0134***	0.0094***
	(0.0035)	(0.0033)	(0.0017)
AFQT score (mother)	0.2710***	0.2702***	0.2704***
	(0.0089)	(0.0071)	(0.0068)
Intercept	0.2726***	–	–
	(0.0244)	–	–
Endogenous Variables			
Dropout rate	−0.7001***	−0.7513***	−0.3216***
	(0.1790)	(0.1728)	(0.1133)
Expenditure per pupil	0.1392***	0.1798***	0.0958***
	(0.0478)	(0.0472)	(0.0276)
Teacher salary	−0.0157	−0.0146	−0.0125**
	(0.0105)	(0.0103)	(0.0052)
Teacher/student ratio	0.4211**	0.4417***	0.3483***
	(0.1714)	(0.1626)	(0.1004)
Move	0.0110**	0.0166***	0.0044
	(0.0055)	(0.0054)	(0.0041)
Part-time work	0.0270***	0.0357***	−0.0461***
	(0.0047)	(0.0045)	(0.0027)
Full-time work	0.0211***	0.0308***	−0.0629***
	(0.0046)	(0.0044)	(0.0052)

Notes: 1. Standard errors of marginal impacts are in parentheses; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively.

Table 3: Simulation within Sample

Table 3a: Goodness of Fit

	Observed	Simulated
AFDC recipients		
Counts	1462 [13.77%]	1383 [13.02%]
Child outcomes		
Mean score	0.476 (0.266)[1]	0.478 (0.245)[1]
Work status		
Any work	7937 [74.7%]	8023 [75.5%]
Full-time	4713 [59.4%][2]	4821 [60.0%][2]
Migration		
Moves	1224	1362
Moves per mover	1.51	1.74

Notes: 1. Standard deviation is in parentheses. 2. Percentage of full-time workers among workers is in square brackets.

Table 3b: Distribution of Outcomes

	AFDC	Workers	Full-Time[1]	Mean Scores
Baseline	13.8%	74.7%	59.4%	0.476
Increase the AFDC benefits by 30%	14.7%	73.5%	59.3%	0.478
Introduce strict work requirements	9.6%	79.1%	54.2%	0.474
Increase wage rates by 20%	13.0%	77.0%	59.5%	0.479

Notes: The percentage of full-time workers among all workers.

Table 4: Simulation with High School Dropouts Sample I

Women Who Reside in NY in 1988 and 1990
Sample Size: 9306(100%)

Table 4a: Dist. of Location in 1994

	% In NY	% In MA	% In other US	Induced Mig.[1]
1. Baseline	84.61	4.56	10.83	–
2. NY benefit down by 30%	84.48	4.60	10.92	0.13 [0.12]
3. MA benefit up by 30%	84.57	4.62	10.81	0.04 [0.04]
4. Work requirements in NY	84.18	4.71	11.11	0.43 [0.49]
5. Work requirements in MA	84.70	4.44	10.86	−0.09 [1.19]
6. NY wage rate up by 20%	84.91	4.47	10.62	−0.30 [0.23]
7. MA wage rate up by 20%	84.53	4.67	10.80	0.08 [0.06]

Notes:

1. Induced migration is defined as the additional out-migration after policy changes in New York. Standard errors of induced migration rates are in square brackets.

Table 4b: Dist. of Location and AFDC Participation in 1994

	% In NY and on AFDC	% In MA and on AFDC	% In other US and on AFDC
1. Baseline	21.45	1.07	2.54
2. NY benefit down by 30%	19.07	1.07	2.57
3. MA benefit up by 30%	21.45	1.23	2.53
4. Work requirements in NY	14.64	1.14	2.63
5. Work requirements in MA	21.47	0.78	2.54
6. NY wage rate up by 20%	20.57	1.06	2.49
7. MA wage rate up by 20%	21.44	1.06	2.53

Table 4c: Dist. of Location and Work in 1994

	%In NY and Work	%In MA and Work	%In other US and Work
1. Baseline	71.71	4.01	9.20
2. NY benefit down by 30%	72.64	4.04	9.25
3. MA benefit up by 30%	71.67	3.98	9.18
4. Work requirements in NY	78.29	4.07	9.39
5. Work requirements in MA	71.77	4.28	9.22
6. NY wage rate up by 20%	73.33	3.92	9.02
7. MA wage rate up by 20%	71.63	4.19	9.17

Table 5: Simulation with a High School Dropouts Sample II

Women Who Reside in NY and Participate in AFDC in 1988 and 1990
Sample Size: 2680(100%)

Table 5a: Dist. of Location in 1994

	% In NY	% In MA	% In other US	Induced Mig.[1]
1. Baseline	84.25	4.55	11.19	–
2. NY benefit down by 30%	84.03	4.63	11.34	0.22 [0.15]
3. MA benefit up by 30%	84.18	4.63	11.19	0.07 [0.03]
4. Work requirements in NY	83.36	4.85	11.79	0.89 [0.51]
5. Work requirements in MA	84.33	4.48	11.19	−0.08 [1.39]
6. NY wage rate up by 20%	84.40	4.51	11.08	−0.15 [0.25]
7. MA wage rate up by 20%	84.14	4.66	11.19	0.11 [0.06]

Notes:

1. Induced migration is defined as the additional out-migration after policy changes in New York. Standard errors of induced migration rates are in square brackets.

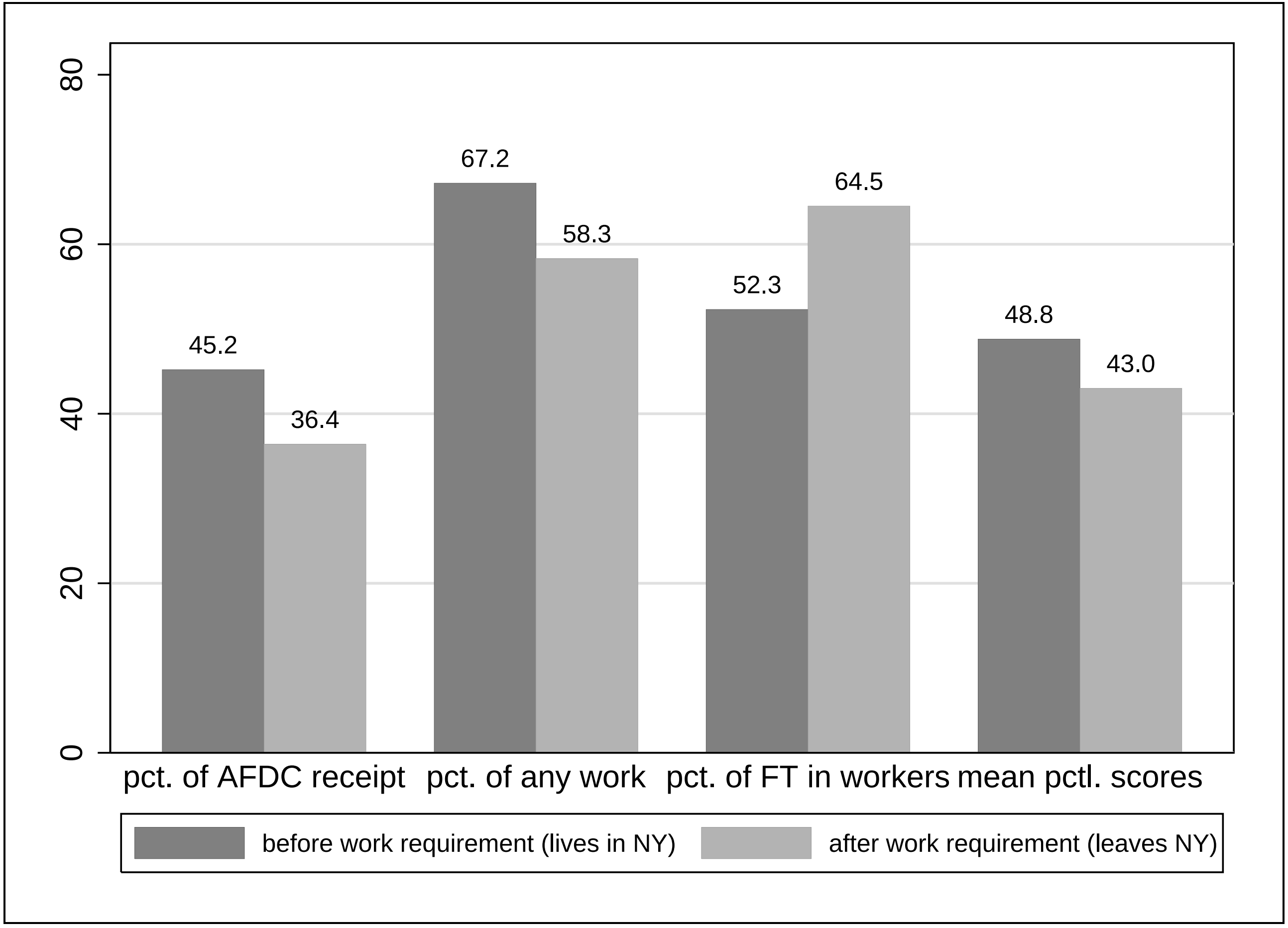
Table 5b: Dist. of Location and AFDC Participation in 1994

	% In NY and on AFDC	% In MA and on AFDC	% In other US and on AFDC
1. Baseline	31.16	1.49	3.69
2. NY benefit down by 30%	27.99	1.49	3.54
3. MA benefit up by 30%	31.16	1.60	3.69
4. Work requirements in NY	19.55	1.64	3.36
5. Work requirements in MA	31.19	1.08	3.69
6. NY wage rate up by 20%	30.30	1.49	3.58
7. MA wage rate up by 20%	31.16	1.42	3.69

Table 5c: Dist. of Location and Work in 1994

	% In NY and Work	% In MA and Work	% In other US and Work
1. Baseline	61.90	3.73	8.21
2. NY benefit down by 30%	63.10	3.77	8.28
3. MA benefit up by 30%	61.87	3.66	8.21
4. Work requirements in NY	72.20	3.81	8.54
5. Work requirements in MA	61.94	4.33	8.21
6. NY wage rate up by 20%	63.99	3.69	8.13
7. MA wage rate up by 20%	61.83	3.84	8.21

Figure 1: Changes in Women's Choices and Children's Scores



NY imposes strict work requirements
Sample: Women Who Would Leave NY in
the Presence of Work Requirements

Appendix

A. Main data sources

This paper has utilized a number of data resources.

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 select mothers who had mother-child data from 1986 to 1996 and exclude those economically disadvantage non-Hispanic whites.

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 1996) supplies data on children with mothers between the ages of 31 and 38 at the end of 1995. 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.National Center for Education Statistics(NCES): The Common Core Data(CCD)

This 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). In particular, we average the grade-specific school characteristics used in the estimation to the county level.

4. 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.

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.

6. Census of Population and Housing, 1990 [United States]: Public Use MicroData Sample (PUMS): 5-percent Sample

To construct relative measurements of teacher salary and expenditure per pupil, we select college-graduated white males, who were between 27 and 40 years old and working full-time (30+ hours a week and 45+ weeks a year). The smallest geographical identifier 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 belongs 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 with 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 males 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.

7. Urban Institute: New Federalism State Database

This database includes comprehensive state-level information on the fifty states and Washington D.C. in areas including demographic, health, child welfare, fiscal and political conditions, and social services. The maximum benefit payments for various family sizes after 1993 are obtained from this database. This data source also provides other aspects of state AFDC rules, such as payment standards, permitted deductibles, and "ratable reduction".

8. Simulated Income Tax Liabilities: TAXSIM

The values of federal and state income taxes are obtained from NBER's TAXSIM program (Feenberg and Coutts, 1993), which calculates liabilities under the U.S. Federal and State income tax laws from individual data. To prepare the income taxes used by the computer program in the paper, we generate a data matrix with 100 income levels, 2 marital statuses, and 4 family sizes for each state. The TAXSIM program then simulates the income tax liabilities for each cell in the data matrix in all states from 1986 to 1996. With these approximated tax schedules, we compute the disposable income that is used in the budget constraint of each household.

9. Cost of Living Index for the American States, 1960-2000

Berry *et al.* (2000) constructed a state cost of living index for the 48 continental American states, measured annually for the period 1960 through 2000. The consumer price index uses 100 as the median state cost of living in 2000. The indices for Hawaii, Alaska, and the District of Columbia are interpolated from Interstate Cost of Living Indices, compiled by American Federation of Teachers (AFT).

B. Formulas of Welfare Benefits

The rules that determined AFDC benefits, and particularly how those benefits changed with family income, were complex and had changed frequently over the life of the program.

For instance, since all AFDC recipients were eligible for food stamps and most qualified for Section 8 housing subsidies, some states elected to make food stamp income and housing subsidies untaxable when considering AFDC benefits. However, in general, the AFDC benefit level for a given family is determined by state-specific guarantees for various family size, the sources and amount of family income, and the sources and amount of expenditures, such as child support and in-work expenses. In the paper, the rule of yearly AFDC benefit is simplified and given as follows (Keane and Moffitt, 1998):

$$B_a = \begin{cases} Min\{P_m, r[P_s - Max(0, wh + y - D)]\} & \text{if } wh + y < P_n \text{ and } h \neq h_2 \\ 0 & \text{if otherwise,} \end{cases}$$

where B_a is the AFDC benefits received by individuals, P_m is the maximum amount paid, r is the "ratable reduction," P_s is the payment standard, wh is wage income, y is non-labor income, D is the permitted deductible, and P_n is the needs standard. The variables P_m, P_s, P_n, and r vary by state, year, and family size. The eligibility data used in the estimation are obtained from the Urban Institute and Robert Moffitt at the Johns Hopkins University.

C. Sampling of Locational Alternatives

McFadden (1978) shows that the IIA property of the logit model permits consistent estimation with only a subset of the alternatives (including the chosen alternative and a sample of rejected alternatives). Also see Arcidiacono (2001) and Bayer (2007) for other applications of this approach. 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 $\pi(D|k_i^*)$ be the probability that D will be drawn, given the observed choice k_i^*. Then

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

As shown by McFadden, the choice probabilities $\Pr(k_{it}^*|\Omega_{it}, \mu_j)$ in equation (20) can then be written as follows:

$$\text{(A.1)} \qquad \Pr(k_i = k_i^*|D, \Omega_i, \mu_i) = \frac{\exp\left\{\frac{\widetilde{V}_k(\Omega_i, \mu_i)}{b_2} + \ln \pi(D|k_i^*)\right\}}{\sum\limits_{k' \in D} \exp\left\{\frac{\widetilde{V}_{k'}(\Omega_i, \mu_i)}{b_2} + \ln \pi(D|k')\right\}}.$$

This simple random sampling strategy is characterized by the uniform conditioning property: $\pi(D|k) = \pi(D|k')$. This property implies that the weighting terms for alternative sampling bias in the logit model, $\ln \pi(D|k)$, are equal and therefore cancel out in equation (21). Consistency of the resulting maximum likelihood estimators relies on the IIA property of the error terms in the discrete choice model. Although this paper uses discrete factor method to account for unobserved heterogeneity, it can be demonstrated that the estimator described in equation (20) possesses the IIA property 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) uses the Monte-Carlo method to test the sensitivity of estimates to the sampling size $\widetilde{K}$. Using a simulated conditional logit model with 3000 choices and 5000 observations for 5 periods, which is estimated by the same discrete factor method, the simulation study indicates that coefficient estimates and the standard errors are nearly identical to each other after choice sample size is greater than 20. The size of random choice subset is set to 20 in this paper.

D. Baseline Simulations Conducted in Section 5

Step 1: We generate a national sample of single women who were 33-year-old high school dropouts in 1988 (period 0). Each woman had an AFQT score of 0.4, one female child, and no non-labor income. In addition, the number of observations in each county of this sample is in proportion to the county's share of the U.S. population, and the number of observations in the state of New York is normalized to 10,000.

Step 2: We draw a heterogeneity type for each woman in this policy relevant sample and the probability of each type is given by the estimates presented in Appendix Table 2f.

Step 3: Following the time-line described in Section 3, each woman received an *i.i.d.* Extreme Value error draw in 1990 (period 1), so that she can compare her expected utility among alternative locations and make a location choice. After choosing place of residence, this representative woman received a wage offer in the chosen location, whose distribution estimates are presented in Appendix Table 2c. Along with the wage offer, her preference shocks associated with each of all five work/welfare choices were also revealed. She then made a work/welfare decision to maximize her expected utility in the chosen location. Recall that her child's achievement outcome was still stochastic at this step. After the work/welfare decision had been made, the child outcome was realized by drawing a score from an idiosyncratic distribution estimated within the full model.

After this step, we further focus on two sub-groups. One is called the "New York sample," which consists of 9306 women who stayed in New York in both 1988 and 1990. The other sub-group is called the "New York AFDC sample," which is selected from the New York sample and contains 2680 women who originated from New York in period 1988 and stayed in New York as welfare recipients in 1990.

Step 4: For the rest of the time periods, all policy parameters in each location were fixed at their corresponding values in 1990, including tax regulations, welfare rules, and school characteristics. Using the simulated location choices in 1990 as given, we repeat Step 3 and Step 4 for two more periods, 1992 (period 2) and 1994 (period 3), recursively.

The simulated out-migration rates, AFDC participation rates, and labor market participation rates in 1994 for the New York sample and the New York AFDC sample are presented in baseline rows in Tables 4a-c and Tables 5a-c respectively.

The baseline simulation in the New York sample shows that in 1994 about 84.6% of 9306 women still stayed in New York, 4.6% of them left for neighboring Massachusetts, and the rest (10.8%) moved to the other states in the U.S. In addition, over 21.4% of the whole New York sample stayed in New York and received AFDC benefits in 1994 and 71.7% of the sample stayed and worked in New York in 1994. Women who stayed in New York were more likely to participate in welfare and less likely to work than women who left New York. Given the fact that New York has been one of high-benefit states, out-migrants in New York were more likely to leave the state for reasons other than higher welfare benefits. It appears that many women anticipated better wage offer distributions and better school districts. Consequently, women who left New York had a greater tendency to work and their children had better achievement outcomes than women who stayed.

The baseline simulation results for the New York AFDC sample can be found in the baseline rows in Tables 5a-c. Conceptually, the main differences between the New York sample and the New York AFDC sample are that women in the latter tended to have tastes in favor of welfare benefits while women in the New York sample represented a more general group of women. Women in the New York AFDC sample were more likely to be at risk of AFDC receipt. For instance, 27.0% of women in the AFDC sample were on AFDC while only 25.4% of the full New York received benefits. In addition, women in the AFDC sample also had a slightly greater tendency to move out of their original place of residence. It is also noteworthy that women in the AFDC sample were less likely to work.

For the policy experiments presented in the text, we use the same steps described in the baseline simulations to generate the New York sample and the New York AFDC sample in 1990. However, starting from 1992, we change the values of one policy parameter of interest in either New York or Massachusetts so that we can study the impacts of this specific policy change, holding everything else constant. The standard errors of induced out-migration rates are obtained by bootstrapping over 30 draws from the estimated distribution of the parameter vector. These parameters are drawn from a multivariate normal distribution with a mean vector and a covariance matrix obtained from maximum likelihood estimation.

E. Comparison between NY, MA, and Other Places in U.S.

New York and Massachusetts are selected for the purpose of policy simulations. As indicated in Appendix Table 1, the levels of AFDC benefits in New York and Massachusetts are similar, so are the median wage rates for non-white 33-39 years old women with less than high school education. However, the overall median rent in Massachusetts is 30% higher than in New York while the average school quality in New York is better than in Massachusetts. The welfare benefit levels in both states are higher than the average level for the rest of the

New England States and that for the rest of the states in the U.S.

Appendix Table 1: Comparisons of NY, MA, and Other Locations (1990)

	NY	MA	Other States in NE	Other States in US
1. Median tax rate of 1^{st} income quartiles	1.8%	3.6%	0.6%	1.9%
2. AFDC maximum payments, family of 2	\$5,616	\$5,832	\$5,286	\$3,516
3. Median rent	\$5,736	\$7,440	\$6,264	\$4,956
4. Median wage of non-white low-educated women aged 33-39	\$8.00	\$8.60	\$7.90	\$6.50
5. Median dropout rate for 11th graders	5.7%	6.4%	5.7%	6.2%
6. Median relative expenditure per pupil of public K-12 schools	17.2%	13.2%	17.9%	12.0%

F. Point Estimates from the Full Empirical Model

Appendix Table 2: Parameter Estimates of the Full Model

Appendix Table 2a: Education Production Function

Variable	Point Estimate	Std. Error[1]
Intercept	−0.3978	(0.3057)
Child age (in 10 years)	0.1377	(0.2833)
Child age (in 10 years) squared	−0.1322	(0.1363)
Age of mother (in 10 years)	0.1439***	(0.0221)
Married	0.0930***	(0.0180)
Mother high school	0.1311***	(0.0223)
Mother more than high school	0.2929***	(0.0277)
Non-white	−0.3423***	(0.0202)
Boy	0.0474***	(0.0145)
Dropout rate for 11^{th} graders	−0.0770	(2.6862)
Expenditure per pupil	1.4896**	(0.6825)
Teacher salary	−0.0838	(0.1525)
Teacher/pupil ratio	0.0857	(2.6573)
Move	−0.4938***	(0.0249)
Mother part-time work	0.0391	(0.0418)
Mother full-time work	0.2000***	(0.0494)
Loading on 1^{st} order heterogeneity factor	−0.7050**	(0.2764)
Loading on 2^{nd} order heterogeneity factor	−2.0016***	(0.6638)
Loading on 3^{rd} order heterogeneity factor	1.0194**	(0.4162)
Dropout rate × $Log(P-p)$	−0.1040	(1.1643)
Expenditure per pupil × $Log(P-p)$	−0.3903	(0.3098)
Teacher salary × $Log(P-p)$	−0.0113	(0.0670)
Teacher/pupil ratio × $Log(P-p)$	−0.1619	(1.1161)
Part-time × $Log(P-p)$	−0.0366*	(0.0217)
Full-time × $Log(P-p)$	−0.1313***	(0.0294)
Dropout rate × child age	−0.7162	(1.2395)
Expenditure per pupil × child age	−0.6914**	(0.3103)
Teacher salary × child age	0.0559	(0.0604)
Teacher/pupil ratio × child age	−0.2095	(1.3468)
Part-time × child age	0.0078	(0.0097)
Full-time × child age	0.0376***	(0.0113)
Mother's AFQT score	0.3741***	(0.1034)
Mother's AFQT score × $Log(P-p)$	0.6503***	(0.0597)
1^{st} order baseline hazard ($Log(P-p)$)	−4.6045***	(0.1865)
2^{nd} order baseline hazard($Log(P-p)^2$)	11.6659***	(0.2435)
3^{rd} order baseline hazard($Log(P-p)^3$)	−9.2949***	(0.1657)
4^{th} order baseline hazard($Log(P-p)^4$)	−2.1418***	(0.0373)

Notes: 1. All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively. 2. P is the highest level of discretized scores and p is any given discretized score level.

Appendix Table 2b: Utlilty Function

Variable	Point Estimate	Std. Error[1]
Intercept in reserve consumption	4.0528***	(0.7589)
Factor loading of 1^{st} order heterogeneity in reserve function	0.1152	(1.8345)
Factor loading of 2^{nd} order heterogeneity in reserve function	1.8297	(1.8566)
Power of consumption (γ_1)	-0.0161	(0.0252)
Married (α_0)	0.1608	(0.1064)
Any work (α_1)	-0.0288	(0.0418)
Relative power of mother's AFQT score (α_3)	1.3731***	(0.3195)
Scale on child's score / mother's AFQT score (α_2)	3.0292**	(1.4091)
Reserve child's score / mother's AFQT score (γ_2)	1.2454	(1.1229)
Power of child's score / mother's AFQT score (γ_3)	0.4409***	(0.0846)
Intercept in part-time leisure	0.6785*	(0.3614)
1^{st} order discrete factor loading in part-time leisure	-1.6928***	(0.5975)
2^{nd} order discrete factor loading in part-time leisure	0.1596	(1.1411)
3^{rd} order discrete factor loading in part-time leisure	1.4738	(0.9715)
# of young children (0-5) in part-time leisure	0.1037***	(0.0288)
# of old children (6-17) in part-time leisure	0.0051	(0.0053)
Married in $f_{4\alpha_1}(\cdot)$	-0.0339**	(0.0170)
Intercept in full-time leisure	-2.7402***	(0.4416)
1^{st} order discrete factor loading in full-time leisure	-0.4420	(0.5857)
2^{nd} order discrete factor loading in full-time leisure	3.0961***	(0.4868)
3^{rd} order discrete factor loading in full-time leisure	-1.0778***	(0.0472)
# of young children (0-5) in full-time leisure	0.6547***	(0.1769)
# of old children (6-17) in full-time leisure	-0.0176	(0.1472)
Married in $f_{4\alpha_2}(\cdot)$	-0.2038	(0.2834)
Intercept in moving cost (ϕ_0)	-3.6875***	(1.2607)
Slope in moving cost (ϕ_1)	-1.6285***	(0.0313)
Dummy for moving across states (α_6)	-2.3849***	(0.7942)
Inverse of parameter b_1 Gumbel error on working choice	3.1332***	(0.7498)
Inverse of parameter b_2 Gumbel error on lcoation choice	2.5849***	(0.8584)
Dummy for census region - Midwest (α_{72})	-0.3557***	(0.1328)
Dummy for census region - South (α_{73})	-0.1643**	(0.0779)
Dummy for census region - West (α_{74})	-0.0909	(0.0741)
Welfare stigma (α_5)	-0.0645	(0.7864)

Notes: 1. All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively. 2. Northeast is the reference region.

Appendix Table 2c : Women's Wage Function

Variable	Point Estimate	Std. Error[1]
Intercept	-1.9087^{***}	(0.1385)
Women's local median wage rate	0.9095^{***}	(0.1242)
Women's local median wage rate squared	-0.2533^{***}	(0.0339)
Women's local median wage rate × heterogeneity factor	0.1823	(0.1610)
Mother high school	0.2653^{***}	(0.0430)
Mother more than high school	0.5434^{***}	(0.0528)
Mother non-white	0.2312^{***}	(0.0274)
Mother's AFQT score	0.8454^{***}	(0.1102)
Mother's AFQT score × $Log(G-g)$	0.3527^{***}	(0.1062)
Loading on 1^{st} order heterogeneity factor	0.5430^{**}	(0.2708)
Loading on 2^{nd} order heterogeneity factor	-1.6412^{***}	(0.2091)
1^{st} order baseline hazard $(Log(G-g))$	-1.1450^{***}	(0.3651)
2^{nd} order baseline hazard$(Log(G-g)^2)$	-0.2569	(0.6429)
3^{rd} order baseline hazard$(Log(G-g)^3)$	-0.1178	(0.2905)

Notes: 1. All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively. 2. G is the highest level of discretized wage rates; g is any given discretized wage rate level.

Appendix Table 2d : Men's Wage Function

Variable	Point Estimate	Std. Error[1]
Intercept	-1.6553^{***}	(0.4265)
Men's local median earnings	0.0747	(1.3766)
Men's local median earnings squared	-0.1307	(0.1239)
Men's local median earnings × heterogeneity factor	0.1115	(0.3495)
Father high school	0.3490	(1.9654)
Father more than high school	0.8249	(0.5452)
Father non-white	-0.5620	(1.3659)
Loading on 1^{st} order heterogeneity factor	-2.6780^{***}	(0.1265)
Loading on 2^{nd} order heterogeneity factor	0.4972	(0.3173)
1^{st} order baseline hazard $(Log(G'-g'))$	-1.3837	(1.8479)
2^{nd} order baseline hazard$(Log(G'-g')^2)$	0.9706^{***}	(0.2739)
3^{rd} order baseline hazard$(Log(G'-g')^3)$	0.2663	(1.3551)

Notes: 1. All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively. 2. G' is the highest level of discretized wage rates; g' is any given discretized wage rate level.

Appendix Table 2e: Parameters Defining Prob. for Heterogeneity Points

Variable	Point Estimate	Std. Error[1]
Probability parameter at 0	1.4695^{***}	(0.1436)
Probability parameter at 1/3	-2.0207^{***}	(0.2178)
Probability parameter at 2/3	3.6988^{***}	(0.2231)

Notes: All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively.

Appendix Table 2f: Probability distribution of Heterogeneity Types

Heterogeneity Factor	Probability
0	0.1737
1/3	0.2772
2/3	0.3571
1	0.1920

Appendix Table 2g: Conditional Logit for Initially Observed Location

Variable	Point Estimate	Std. Error[1]
Dropout rate × Grandma high school grad	−0.1233	(0.1404)
Expenditure per pupil × Grandma high school grad	0.1260	(0.2998)
Mother wage rate × Grandma high school grad	0.0634	(1.8427)
AFDC max payment × Grandma high school grad	0.0931	(1.0590)
Tax liabilities × Grandma high school grad	0.2125	(0.4962)
Dropout rate × Mother non-white	0.5399***	(0.0548)
Expenditure per pupil × Mother non-white	0.4806***	(0.1579)
Mother wage rate × Mother non-white	0.0860	(0.7077)
AFDC max payment × Mother non-white	0.0732	(1.1507)
Tax liabilities × Mother non-white	0.0723	(1.3142)
Dropout rate × Mother high school grad	0.0718	(1.3848)
Expenditure per pupil × Mother high school grad	0.0825	(0.9152)
Mother wage rate × Mother high school grad	0.0779	(0.9661)
AFDC max payment × Mother high school grad	0.0757	(0.9512)
Tax liabilities × Mother high school grad	0.1128	(0.4743)
Dropout rate × Mother age (in 10 yrs)	0.0806	(0.1777)
Expenditure per pupil × Mother age (in 10 yrs)	0.0631	(0.3996)
Mother wage rate × Mother age (in 10 yrs)	0.1015	(1.3762)
AFDC max payment × Mother age (in 10 yrs)	0.0906	(0.2277)
Tax liabilities × Mother age (in 10 yrs)	0.0861	(1.0519)
Dropout rate × 1^{st} order heterogeneity factor	0.0636	(0.8371)
Dropout rate × 2^{nd} order heterogeneity factor	0.0649	(1.0079)
Dropout rate × 3^{rd} order heterogeneity factor	0.0657	(1.0641)
Expenditure per pupil × 1^{st} order heterogeneity factor	−2.7238***	(0.2500)
Expenditure per pupil × 2^{nd} order heterogeneity factor	−3.4382***	(0.0835)
Expenditure per pupil × 3^{rd} order heterogeneity factor	−1.9753***	(0.0727)
Mother wage rate × 1^{st} order heterogeneity factor	−0.4466***	(0.1003)
Mother wage rate × 2^{nd} order heterogeneity factor	0.0385***	(0.0034)
Mother wage rate × 3^{rd} order heterogeneity factor	−0.0001*	(0.0000)
AFDC max payment × 1^{st} order heterogeneity factor	−0.0347***	(0.0044)
AFDC max payment × 2^{nd} order heterogeneity factor	0.0474	(0.0384)
AFDC max payment × 3^{rd} order heterogeneity factor	0.4861***	(0.0453)
Tax liabilities × 1^{st} order heterogeneity factor	−0.7334***	(0.0285)
Tax liabilities × 2^{nd} order heterogeneity factor	−0.2973	(0.1991)
Tax liabilities × 3^{rd} order heterogeneity factor	−0.1987	(0.1612)
Any move	−0.3013	(0.2495)
Move across states	0.2934	(0.3922)
Move across regions	−0.3153**	(0.1558)

Notes: 1.All standard errors are robust standard errors; ***, **, and * indicate statistical significance at 99%, 95%, and 90% respectively. 2. “Mother wage rate” represents the local average wage rate for women of the same age, education attainment, and race. 3. “Tax liabilities” are calculated for a family of the corresponding size, marital status, and income equal to national median value.

Notes

[1] See review articles by Moffitt (1992) and Brueckner (2000).

[2] The PIAT is individually administered during household interviews. A body of literature has studied limitations in the use of standardized tests as measures of educational outputs (see Koretz (2002)). The PIAT is less likely to be subject to some problems encountered by standardized achievement tests, such as cheating, coaching, and reallocating achievements.

[3] Housing costs are deflated by national consumer price indices

[4] Van der Klaauw (1996) concludes that higher wage rates for women are associated with their lower marriage rates and ignoring the endogeneity of marital status decisions is likely to cause attenuated estimates of own and husband's wage effects on female labor supply. A number of studies also have investigated the interactions between welfare participation and marriage decisions (Moffitt, 1992; Fitzgerald, 1991; Bitler *et al.*, 2004).

References

Arcidiacono P. S. (2005) 'Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?' *Econometrica* 73: 1477–1524.

Bayer P., Ferreira F. and McMillan R. (2007) 'A Unified Framework for Measuring Preferences for Schools and Neighborhoods', *Journal of Political Economy*, in press.

Berry W., Fording R. and Hanson R. (2000) 'An Annual Cost of Living Index for the American states, 1960-2000 (revised)', *Journal of Politics* 60(2): 550–567.

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

Bitler M. P., Gelbach J. B., Hoynes H. W. and Zavodny M. (2004) 'The impact of welfare reform on marriage and divorce', *Demography* 41(2): 213–236.

Blank R. (1988) 'The Effect of Welfare and Wage Levels on the Location Decisions of Female-headed Households', *Journal of Urban Economics* 24: 186–211.

Blau F. D. and Grossberg A. J. (1992) 'Maternal Labor Supply and Children's Cognitive Development', *Review of Economics and Statistics* 74(2): 474–481.

Brueckner J. K. (2000) 'Welfare Reform and the Race to the Bottom: Theory and Evidence', *Southern Economic Journal* 66: 505–525.

Burke V. (1997) 'Brief legislative history of Title IV-A of the Social Security Act', Washington, D.C., Unpublished manuscript.

Feenberg D. and Coutts E. (1993) ‘An Introduction to the TAXSIM Model’, *Journal of Policy Analysis and Management* 12(1): 189–194.

Fitzgerald J. (1991) ‘Welfare durations and the marriage market: Evidence from Survey of Income and Program Participation’, *Journal of Human Resources* 26(3): 545–561.

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

Gramlich E. M. and Laren D. S. (1984) ‘Migration and income Redistribution Responsibilities’, *Journal of Human Resources* 19: 489–511.

Hanushek E. A. (2004) ‘Disruption versus Tiebout improvement: The costs and benefits of switching schools’, *Journal of Public Economics* 88(9-10): 1721–1746.

Hanushek E. A., Kain J. F. and Rivkin S. G. (1998) ‘Conclusions and Controversies about the Effectiveness of School Resources’, *Economic Policy Review* 4(1): 11–28.

Heckman J. (1981) ‘The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-discrete Stochastic Process’ in Manski C. and McFadden D. (eds.) *Structural Analysis of Discrete Data with Econometric Applications.* London: MIT Press: 114–179.

Keane, M. and Moffitt, R. (1998) ‘A Structural Model of Multiple Welfare Program Participation and Labor Supply’, *International Economic Review* 39(3): 553–589.

Kennan J. and Walker J. R. (2003) ‘The Effect of Expected Income on Individual Migration Decisions’, NBER Working Paper 9585.

Kennan J. and Walker J. R. (2001) ‘Geographical Wage Differentials, Welfare Benefits and Migration’, Unpublished manuscript.

Koretz D. (2002), ‘Limitations in the Use of Achievement Tests as Measures of Educators’ Productivity’, *Journal of Human Resources* 37(4): 752–777.

Krueger A. B. (1998) ‘Reassessing the View That American Schools Are Broken’, *Economic Policy Review* 4(1): 29–43.

Levine P. B. and Zimmerman D. J. (1999) ‘An empirical analysis of the welfare magnet debate using the NLSY’, Journal of Population Economics, 391–409.

Liu H. (2004) ‘Limiting the Number of Alternative Choices in Random Utility Models with Unobserved Heterogeneity’, Unpublished manuscript, East Carolina University.

Liu H., Mroz T. and van der Klaauw W. (2006) ‘Maternal Employment, Migration, and Child Development’, Unpublished manuscript.

Markwardt F. (1989) *Peabody Individual Achievement Test-Revised*, American Guidance Services, Circle Pines, MN.

McFadden D. (1978) ‘Modeling the Choice of Residential Location’, in Karlqvist A., Lundqvist L., Snichars F. and Weibull J. (eds.) *Spatial Interaction Theory and Planning Models*. Amsterdam: North Holland: 75–96.

Meyer B. D. (2000) ‘Do the Poor Move to Receive Higher Welfare Benefits?’ Unpublished manuscript.

Moffitt R. (1992) ‘Incentive Effects of the U.S. Welfare System: A Review’, *Journal of Economic Literature* 30: 1–61.

Moffitt R. (1983) ‘An Economic Model of Welfare Stigma’, *American Economic Review* 75(3): 1023–1035.

Van der Klaauw W. (1996) ‘Female Labour Supply and Marital Status Decisions: A Life-cycle Model’, *Review of Economic Studies* 63: 199–235.

Walker J. R. (1994) ‘Migration Among Low-Income Households: Helping the Witch Doctors Reach Consensus’, Discussion Paper 94-1031, Institute for Research on Poverty, University of Wisconsin-Madison.