Preliminary and Incomplete

Estimating the Effect of Welfare on the Education, Employment, Fertility and Marriage Decisions of Women

Michael P. Keane

New York University and University of Minnesota

and

Kenneth I. Wolpin

University of Pennsylvania

May, 2000

I. Introduction

There is an extensive literature in economics that seeks to determine the quantitative impact of welfare programs on female labor supply and the propensity of women to participate in the welfare system. A growing literature also examines the impact of welfare generosity on fertility and marriage - behaviors that influence welfare eligibility and the level of benefits. There are almost no studies that consider the relationship between welfare and schooling decisions.

Most of the studies adopt a static choice framework, albeit not always explicitly, to motivate their empirical specifications.[1] However, the behaviors that are presumably affected by the welfare system (fertility, marriage, work, school) clearly have both immediate and long-term consequences. If potential welfare recipients are (to any extent) forward-looking (rather than completely myopic), they will consider these long term consequences when making current decisions. In this paper, we structurally estimate a behavioral model of the joint schooling, work, fertility and marriage decisions that accounts for both the short-term benefits as well as the potential long-term consequences of welfare participation.

Two recent papers concerned with estimating the impact of AFDC on behavior have structurally implemented models that incorporate dynamic decision-making. Sanders (1993) specifies and estimates a model in which women sequentially make monthly decisions about whether or not to work and whether or not to receive welfare. A woman works full-time (40

[1]Moffitt (1983) is an exception in that he explicitly specifies and structurally estimates a static model of labor supply and welfare participation. Keane and Moffitt (1994) extend that framework to a consideration of multiple program participation. Explicit models of demographic behavior are rarer, although Rosenzweig (1995) and Keane and Wolpin (1999) are exceptions. Keane and Wolpin (1999) also consider the impact of welfare on schooling decisions. For recent surveys of the literature see Moffitt (1992, 1996).

hours per week) if she works and is not on welfare and part-time (10 hours per week) if she works and is on welfare. Marriage and fertility are taken as exogenous, although the events are history-dependent and occur probabilistically. Dynamics arise in the decision process for several reasons: (1) Because the hourly wage rate is a function of past work decisions, depending on cumulative work experience and work experience in the recent past, forgoing work reduces future wages; (2) Past work experience affects the disutility of working and; (3) Being on welfare in the past directly affects the utility associated with being on welfare currently.

Swann (1996) adds marriage to the choice set, but maintains the assumption that fertility is exogenous, and, unlike Sanders, assumes that working precludes welfare receipt. Decisions are made annually. Marital duration affects current utility and thus introduces additional dynamic considerations. Both of these papers assume perfect foresight concerning future welfare benefits.

We extend these efforts in a number of dimensions. We augment the choice set to include schooling and fertility in addition to work, marriage and welfare participation. Moreover, in addition to considering a larger choice set, the modeling framework with respect to each of these alternatives is richer. Specifically, with respect to the work alternative, employment may be either part- or full-time. The markets for part- and full-time employment are treated as distinct. In each period, with some probability a woman receives a part-time wage offer and, likewise, with some probability a full-time wage offer. With respect to the welfare alternative, in addition to stigma effects of participation, we also allow for effects of past welfare participation on labor market and marriage opportunities. Moreover, we explicitly account for imperfect foresight in the structure of the rules governing the level of welfare benefits and model welfare rules more completely than previously.

The marriage market is modeled in a search context. In each period a woman receives a marriage offer with some probability that depends on her current characteristics. The permanent earnings potential of the person she meets is drawn from a distribution that also depends on her characteristics. If the marriage offer is accepted, the husband's actual earnings evolves over time stochastically. The woman receives a fraction of the total of her earnings and her husband's earnings. If a woman is not married, there is some probability, determined by current characteristics, that she co-resides with her parents. In that case, she receives a fraction of her parents' income that also depends on her characteristics.

In modeling the fertility decision, it is assumed that a woman receives utility from children, but bears a time cost of rearing them that depends on their current age distribution. Sequential decisions about school attendance are governed by direct preferences and by the additional human capital, and thus wages, gained from schooling.

We implement the model using 15 years of data from the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY79), supplemented with state level welfare benefit rules that we have collected for each state over a 23 year period prior to the new welfare reform. Benefit levels changed considerably over the decision-making period of the women in the NLSY79 sample. We develop simplified representations of state- and year-specific welfare benefit formulas to estimate forecasting rules for the agents that they are assumed to use in the decision model. The model is estimated on five of the largest states represented in the NLSY79 (California, Michigan, New York, North Carolina and Ohio). As a method of assessing the adequacy of the model, based on the estimates of the model we forecast the behavior of women residing in Texas, a state with much lower welfare benefits than the five used in estimation. We use the model estimates to perform a number of policy experiments, including some that have been newly instituted under recent welfare reform, e.g., time limits and work fare.

II. Model

Choice Set:

We consider a woman who makes a joint decision at each age a of her lifetime consisting of the following set of discrete choices: whether or not to attend school, s_a, work part-time, h_a^p, or full-time, h_a^f, in the labor market, be married (if faced with the opportunity), m_a, become pregnant if the woman is of a fecund age, p_a, and receive government welfare if the woman is eligible, g_a. There are thus as many as 36 mutually exclusive choices that a woman makes at each age during her fecund life cycle stage and 18 choices during her infecund stage.[2] The fecund stage is assumed to begin at age 14 and to end at age 45; the decision period extends to age 62. Decisions are made at discrete six month intervals, i.e., semi-annually. In terms of the timing of births, a woman who becomes pregnant at age a has a birth at age a+1, with n_{a+1} representing the discrete birth outcome.[3]

Preferences:

The woman receives a utility flow at each age that depends on the following: her goods consumption, c_a, her stock of children, N_a, an aged indexed stock of children, N_a^*, the number of

---

[2] Being married and receiving welfare is not an option. Spending on AFDC-Unemployed Parent (AFDC-UP), which provides welfare for in-tact families in which one parent has a history of employment but is currently unemployed, is a very small proportion of total spending on AFDC.

[3] In keeping with the assumption that pregnancies can be perfectly timed, we only consider pregnancies that result in a live birth, i.e., we ignore pregnancies that result in miscarriages or abortions. We assume that a woman cannot become pregnant in two consecutive six month periods.

hours she spends at work, h_a (0, 500 or 1000 hours), her current school attendance status, her current marital status, her current pregnancy status, her current welfare status and whether or not she is currently living with parents, z_a, taken to be exogenous although probabilistic (see below):

$$(1) \quad U_a = U(c_a,\ N_a,\ N_a^*,\ h_a,\ s_a,\ m_a,\ p_a,\ g_a,\ z_a;\ e_a^j),$$

where e_a^j is a vector of serially independent preference shocks (j = h, s, m, p, g).

The stock of children at age a, N_a, evolves according to

$$(2) \quad N_a = N_{a-1} + p_{a-1},$$

where at a=1 (the initial fecund age), the woman has zero children.[4] The age-indexed stock of children is intended to reflect a time cost of children that depends on the age distribution of the current stock. In particular,

$$(3) \quad N_a^* = n_a + d_1 N_a^{1,6} + d_2 N_a^{7,13} + d_3 N_a^{14,18},$$

where N_a^{jk} is the stock of children between the ages of j and k, and the parameters are weights.

Budget Constraint:

We assume that the budget constraint is satisfied each period:

$$(4) \quad \begin{aligned} c_a = {} & y_a^o(1-m_a)(1-z_a) + [y_a^o + y_a^m]m_a t_a^m + [y_a^o + y_a^z t_a^z]z_a \\ & + g_a k_b b_a - [tc_1 I(S_a \geq 12) - tc_2 I(S_a \geq 16)]s_a, \end{aligned}$$

[4] By convention, lower case letters are flow variables and upper case letters are stock variables.

where y_a^o is the woman's own earnings at age a, y_a^m is the spouse's earnings if the woman is married, t_a^m is the share of household income the woman receives if she is married, y_a^z is her parents' income, a share, t_a^p, of which she receives if she co-resides with her parents, b_a is the amount of welfare benefits the woman is eligible to receive and k_b is a fraction that converts welfare dollars into a monetary equivalent consumption value, tc_1 is the tuition cost of college and tc_2 the additional cost of graduate school, S_a is the completed level of schooling at age a and $I(\cdot)$ is an indicator function equal to unity when the argument is true.[5] Notice that we assume income is pooled when married, but not when co-residing with parents.

Living with parents and being married are taken to be mutually exclusive states. In particular, a woman who chooses to be married, conditional on receiving a marriage offer (see below), cannot live with her parents while a woman who does not choose to be married lives with her parents according to a draw from an exogenous probability rule, p^z. We assume that the probability of co-residing with her parents, given the woman is unmarried, depends on her age and on her co-resident status in the previous period. The woman's share of her parents' income, which itself is stochastic and depends on parents' schooling, when co-resident depends on her age and whether she is attending post-secondary school (see Keane and Wolpin (forthcoming)) .

It is assumed that there is stochastic assortative mating. In each period a single woman draws an offer to marry with probability p^m, that depends on her age and welfare status. If the woman is currently married, with some probability that depends on her age and duration of

[5] Welfare recipients are restricted in what they may purchase with welfare benefits, e.g., food stamps cannot be used to purchase alcohol. As seen in (4), we also do not allow for an explicit goods cost to bearing and rearing children, which cannot be separately identified from, and is thus reflected in, the marginal utility of an additional child.

marriage, she receives an offer to continue the marriage. If she declines to continue, the woman must be single for one period (six months) before receiving a new marriage offer. While married the woman receives a fixed share of the sum of her earnings and her husband's earnings.

The woman's earnings is the product of her hourly wage rate and the number of hours she works, $y_a^o = 500 \cdot w_a^p h_a^p + 1000 \cdot w_a^p h_a^f$. The hourly wage rate is the product of the woman's human capital stock, $?_a$, and its per unit rental price, which is allowed to differ between part- and full-time jobs, r^j for j=p, f. Specifically, her ln hourly wage offer is

$$(5) \quad \ln w_a^j = r^j + ?(\cdot) + v_a^w, \quad j = p, f.$$

We adopt a fairly general formulation for the woman's human capital stock; $?_a$ is modeled as a function of completed schooling S_a, the stock of accumulated work hours up to age a, H_a, whether or not the woman worked part- or full-time in the previous period and current age. As with preference shocks, random shocks to a woman's human capital stock, v_a^w, are also assumed to be serially independent.

In each period a woman receives a part-time job offer with probability p^{wp} and a full-time job offer with probability p^{wf}. Each of these offer rates depends on the woman's previous period's work status and on her previous period's welfare status. They are thus not independent of past behaviors.

The husband's earnings depends on his human capital stock, $?_a^m$. Conditional on receiving a marriage offer, the potential husband's human capital is drawn stochastically. The human capital of the spouse that is drawn depends on a set of the woman's current characteristics, her schooling attainment, work experience, current marital status and her age. In addition, there is an iid random component to the draw of the husband's human capital that reflects a permanent characteristic of the husband unknown to women prior to meeting, μ^m. Women can therefore profitably search in the marriage market for husbands with more human capital, and can also directly affect the quality of their husbands by their own behaviors. After marriage, husband's earnings evolve with a fixed trend subject to a serially independent random shock, v_a^m. Specifically,

$$(6) \quad \ln y_a^m = \mu^m + ?^m_{0a}(\cdot) + v_a^m$$

where $?^m_{0a}$ is the deterministic component of the husband's human capital stock net of the permanent component.[6]

Welfare eligibility and the benefit amount for a woman residing in state s at calendar time t depends on the number of children residing with her and on her household income. For any given number of minor children (under the age of 18) residing in the household, the schedule of benefits can be represented by two line segments. The first line segment corresponds to the guarantee level; it is assumed (approximated) to be linearly increasing in the number of minor children and, in the case of a woman co-residing with her parents, linearly declining in parents' income, y_a^z. The second line segment is negatively sloped as a function of the woman's own earnings, y_a^o, plus parents' income if she is co-resident, and also linearly increasing in the number of minor children. The negative slopes reflect the benefit reduction (or tax) applied to income. A dollar of parents' income is, in terms of benefit reduction, treated as only a fraction,

---

[6] The human capital rental price is impounded in this term.. In addition, husband's labor supply is assumed to be an exogenous component of his earnings.

k_z, of a dollar.[7]

In general, benefits are equal to the guarantee level (for given numbers of children and parents' income if co-resident) up to some positive level of the woman's earnings (the two line segments intersect at positive earnings) due to a child care allowance for working mothers. Denoting this (state-specific) level of earnings, the disregard, as $y_{at}^{s1}(N_a^{18})$ and the level of earnings at which benefits become zero (where the second line segment intersects the x-axis) as $y_{at}^{s2}(N_a^{18})$, the benefit schedule for a woman with $N_a^{18} > 0$ children is given by

$$
\begin{aligned}
(7)\qquad b_t^s(N_{at}^{18}, y_{at}^o, y_{at}^z) &= b_{0t}^s + b_{1t}^s N_{at}^{18} - b_{4t}^s k_z y_{at}^z z_{at} && \text{for } y_{at}^o < y_{at}^{s1}(N_a^{18}), \\
&= b_{2t}^s + b_{3t}^s N_{at}^{18} - b_{4t}^s[(y_{at}^o - y_{at}^{s1}) + k_z y_{at}^z z_{at}] && \text{for } y_{at}^{s1}(N_a^{18}) < y_{at}^o < y_{at}^{s2}(N_a^{18}) \\
&= 0 && \text{otherwise}.
\end{aligned}
$$

We refer to $b_t^s(N_{at}^{18}, y_{at}^o, y_{at}^z)$ as the benefit rule and to the b_{kt}^s's as the benefit rule parameters. We exclude k_z from this set for reasons that will become clear.

The benefit rule parameters, and thus benefits themselves, change over time. Therefore, if women are at all forward-looking, they will incorporate their forecasts of the future values of the benefit rule parameters into their decision rules. We assume that benefit rule parameters evolve according to the following general vector autoregression (VAR) and that women use the VAR to form their forecasts of future benefit rules:

$$(8)\qquad \mathbf{b}_t^s = ?^s + ?^s \mathbf{b}_{t-1}^s + \mathbf{u}_t^s$$

[7] The exact treatment of parents' income is quite complicated, varying among and within states (at the local welfare agency level) and over time. Rather than attempting to model the rules explicitly, as an approximation we instead estimate the fraction of parents' income that is subject to tax as a parameter.

where $\mathbf{b}_t^s$ and $\mathbf{b}_{t-1}^s$ are 5×1 column vectors of the benefit rule parameters, **?**$^{\mathbf{s}}$ is a 5×1 column vector of regression constants, **?** $^{\mathbf{s}}$ is a 5×5 matrix of autoregressive parameters and $\mathbf{u}_t^s$ is a 5×1 column vector of iid innovations drawn from a stationary distribution with variance-covariance matrix **?** $^{\mathbf{s}}$. We call (8) the evolutionary rule (ER) and **?**$^{\mathbf{s}}$, **?** $^{\mathbf{s}}$, **?** $^{\mathbf{s}}$ the parameters of the ER. Evolutionary rules are specific to the woman's state of residence. It is assumed that a woman remains in the same location throughout her life.[8]

Objective Function:

The woman is assumed to maximize her expected present discounted value of remaining lifetime utility at each age. The maximized value (the value function) is given by

$$(9) \quad V_a(O_a) = \max E\left[\sum_{t=a}^{62} \delta^{t-a} U_t(\cdot) \mid O_a\right],$$

where the expectation is taken over the distribution of future preference shocks, labor market, marriage and parental co-residence opportunities, and the distribution of the future innovations of the benefit ER. The decision period is six months until age 45, the assumed age at which the women becomes infecund, and one year thereafter.[9] In (8), the state space O_a denotes the relevant factors known at age a that affect current or future utility or that affect the distributions of the future shocks and opportunities.

Decision Rules:

---

[8] Introducing migration in a forward-looking model such as this would greatly complicate the decision problem.

[9] Allowing for a longer decision period at ages past 45 reduces the computational burden of the model (see Wolpin (1992)).

The solution to the optimization problem is a set of age-specific decision rules that relate the optimal choice at any age, from among the feasible choices, to the elements of the state space at that age. Recasting the problem in a dynamic programming framework, the value function, $V_a(\Omega_a)$, can be written as the maximum over alternative-specific value functions, denoted as $V_a^j(\Omega_a)$, i.e., the expected discounted value of choice $j \in J$, that satisfy the Bellman equation, namely

$$V_a(\Omega_a) = \max_{j \in J}[V_a^J(\Omega_a)]$$

$$(10) \quad V_a^j(\Omega_a) = U_a^j + \delta E(V_{a+1}(\Omega_{a+1})|j \in J, \Omega_a) \text{ for } a < A,$$

$$= U_A^j \qquad \text{for } a = A.$$

A woman at each age a (permanently) residing in state s, and thus facing a benefit rule given by (6), with current state Ω_a (including realizations of the benefit rule parameters corresponding to the calendar time the woman is age a, preference shocks, own and husband's earnings shocks, parental income shocks, and labor market, marriage and parental co-residence opportunities), chooses the option with the greatest expected present discounted value of lifetime utility.

Solution Method:

The solution of the optimization problem is in general not analytic. In solving the model numerically, one can regard its solution as consisting of the values of $EV_{a+1}(\Omega_{a+1}|j \in J, \Omega_a)$ for all j and elements of Ω_a. We refer to this function as $Emax$ for convenience. As seen in (10), treating these functions as known scalars for each value of the state space transforms the dynamic optimization problem into the more familiar static multinomial choice structure. The solution method proceeds by backwards recursion beginning with the last decision period. Given the Emax functions, a sample of women belonging to the same cohort can be simulated by drawing at a=1 a set of preference shocks, own and husband's earnings shocks, parental income shocks, labor market, marriage and parental co-residence opportunities, and a set of benefit rule parameters that apply to all women residing in the same state, determining each woman's optimal choice, updating the state space, drawing a new set of shocks, opportunities and benefit parameters at a=2, etc.[10] We use such simulated samples as the basis of our estimation procedure described below.

III. Data

The 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY79) contains extensive information about schooling, employment, fertility, marriage, household composition, geographic location and welfare participation for a sample of over 6,000 women who were age 14-21 as of January 1, 1979. In addition to a nationally representative core sample, the NLSY contains oversamples of blacks and Hispanics. The analysis we perform makes use of the annual interviews between 1979 and 1991 for women from the core sample and from the black and Hispanic oversamples.

The NLSY79 collects much of the relevant information, births, marriages and divorces, periods of school attendance, job spells, and welfare receipt, as dated events. This mode of

[10] Because the size of the state space is very large, we adopt an approximation method to solve for the Emax functions. The Emax functions are calculated at a limited set of state points and their values are used to fit a polynomial approximation in the state variables consisting of linear, quadratic and interaction terms. See Keane and Wolpin (1994, forthcoming) for further details. As a further approximation, we use only the forecast variance in the benefit parameters, ignoring higher order moments, in calculating the Emax functions.

collection allows the researcher the freedom to choose a decision period essentially as small as one month, i.e., to define the choice variables on a month-by-month basis. Although the exact choice of the length of a period is arbitrary, we adopted as reasonable a decision period of six months. Periods are defined on a calendar year basis, beginning either on January1 or on July 1 of any given year. We begin the analysis with data on choices starting from the first six month calender period that the woman turned age 14 and ending in the second six month calendar period in 1990 (or, if the woman attrited before then, the last six-month period in which the data are available). The first calendar period observation, corresponding to that of the oldest NLSY79 sample members, occurs in the second half of 1971. There are fifteen other birth cohorts who turned age 14 in each six month period through January, 1979.

We restrict the sample to the six states in the U.S. that have the largest representations of NLSY79 respondents: California, Michigan, New York, North Carolina, Ohio and Texas. However, the estimation is performed using only the first five states. Texas is used as a holdout sample on which to perform out-of-sample validation tests of the model. The reason for this choice is that, as shown below, Texas is by far the least generous state in terms of welfare benefits and thus requires an extreme out-of-sample extrapolation.

Choice Set:

As noted, we consider the following choices: whether or not to (i) attend school (ii) work (part- or full-time), (iii) be married, (iv) become pregnant and (v) receive welfare (AFDC). The variables are defined as follows:

School Attendance:

The NLSY79 collects data that permits the calculation of a continuous monthly attendance record for each women beginning as of January, 1979. A woman was defined to be attending school if she reported being in school each month between January and April in the first six-month calendar period and each month between October and December in the second calendar period.[11] Given the sample design of the NLSY79, school attendance records that begin at age 14 exist only for the cohort that turned 14 in January, 1979.

School attendance prior to age 14 is not explicitly treated as a choice. However, completed schooling at any age, including at age 14 which we refer to as initial schooling, affects opportunities and thus choices. Given the sample design, we know initial schooling only for one of the cohorts. An estimation procedure has to deal with this serious missing initial conditions problem as well with the missing observations on schooling choices for many of the cohorts.

Employment Status:

At the time of the first interview, an employment history was collected back to January 1, 1978, which provided details about spells of employment with each employer including the beginning and ending dates (to the week) of employer attachments as well as gaps within employer-specific spells. Subsequent rounds collected the same information between interview

---

[11] Beginning with the 1981 interview, school attendance was collected on a monthly basis for the prior calendar year. In the two prior interviews, attendance was ascertained at the interview date and, if not attending, the date of last attendance was obtained. If a woman was attending (not attending) at the time of the1979 interview, which, in every case, took place during the first six months of 1979, and similarly in the first period of 1980 according to the above rule, then the individual was coded as attending (not attending) in both periods of 1979. If attendance differed between the two years, enrollment was considered missing in the second half of 1979. We do not use the data prior to 1979 because only the last spell of non-attendance, and then only for individuals not attending at the 1979 interview, can be determined. In addition, because reported attendance and completed schooling levels were often longitudinally inconsistent, the attendance data was hand-edited to create a consistent attendance-highest grade completed profile.

dates. Using this information together with data on usual hours worked at each employer, we calculated the number of hours worked in each six month period. A woman was considered working part-time in the period (500 hours) if she worked between 260 and 779 hours and full-time (1000 hours) if she worked at least 780 hours during the period. As with school attendance, employment data does not extend back to age 14 for many of the cohorts. Unlike schooling, we assume that initial work experience, that is, at age 14, is zero.

Marital Status:

The NLSY79 provides a complete event-dated marital history that is updated each interview. However, dates of separation are not reported. Therefore, for the years between 1979 and 1990, data on household composition was used to determine whether the woman was living with her spouse. But, because these data are collected only at the time of the interview, marital status is treated as missing during periods in which there were no interviews, in most cases for one six-month period per year. Marital event histories were used for the periods prior to 1979 even though it is uncertain from that data whether the spouse was present in the household.

Pregnancy Status:

Although pregnancy rosters are collected at each interview, conception dates are noisy and miscarriages and abortions are under-reported. We ignore pregnancies that do not lead to a live birth, dating the month of the conception as occurring nine months prior to the month of birth. Except for misreporting of births, there is no missing information on pregnancies back to age 14 for any of the cohort.

Welfare Receipt:

AFDC receipt is reported for each month within the calendar year preceding the interview year, i.e., from January 1978. The respondent checks off each month from January through December that a payment was received.[12] We define a woman as receiving welfare in a period if she reported receiving an AFDC payment in at least three of the six months of the period.[13] As with school attendance and employment, data are missing back to age 14 for most of the cohorts. It is assumed that none of the women received welfare prior to age 14, as is consistent with the fact that none had borne a child by that time.

Descriptive Statistics:

Table 1 provides the (marginals of the) sample choice distribution by full-year ages and by race aggregated over the five states used in the estimation. As seen, school attendance is essentially universal until age 16, drops about in half at age 18, the normal high school graduation age, and falls to under 10 percent after age 22. About 3 percent of the sample attends school at ages after 25. The implied school completion levels that result from these attendance patterns are, at age 24, 12.9 for whites, 12.7 for blacks and 12.2 for Hispanics. For white and Hispanic women, employment rates (either part- or full-time) increase rapidly through age 18 and then slowly thereafter, although they are higher for whites throughout by about 10-20 percentage points. Employment rates for black females rise more continuously, approximately doubling between age 18 and 25, and are comparable to that of Hispanics at ages after 25. Marriage rates

---

[12] This method of data collection has led to a serious seam problem. In the monthly data, there are many more transitions out of welfare between December of one year and the following January than there are between any two months within any calendar year. We attempt to account for this problem in the empirical specification we adopt.

[13] The use of almost any cutoff in establishing welfare participation would have only a small effect on the classification; most women who report receiving welfare in any one month during a six month period report receiving it in all six months.

rise continuously for whites and Hispanics, reaching about 60 percent by age 25 for whites and 50 percent for Hispanics. However, for blacks, marriage rates more or less reach a plateau at about age 22, at between 20 and 25 percent. With respect to fertility, it is more revealing to look at cumulative children ever born rather than at pregnancy rates within six-month periods. By age 20, white females in the sample on average had .28 live births, black females .47 live births and Hispanic females .40 live births. Teenage pregnancies that lead to a live birth are higher by 68 percent for blacks than for whites and by 43 percent for Hispanics than for whites. By age 27, the average number of live births are 1.06, 1.36 and 1.39, and by age 30, 1.54, 1.61 and 1.76. Viewed differently, the first age at which the sample women have had one child on average was 27 for whites, 24 for blacks and 24.5 for Hispanics. Welfare participation naturally increases with age given the eligibility requirement associated with having had at least one child. Race differences are large; at it peak, participation reaches 7 percent for whites, 28 percent for blacks and 17 percent for Hispanics.

In order to estimate the benefit schedules (7) and the evolutionary rules governing changes in benefit parameters (8), we collected information on the rules governing AFDC and Food Stamp eligibility and benefits in each of the 50 states for the period 1967-1990. The parameters of the benefit schedule are obtained by estimating (7) for each state separately in each year using the sum of the monthly benefits from AFDC and Food Stamps, with monthly benefit amounts expressed in 1987 New York equivalent dollars. Thus, for each state, s, we obtained an estimate of the benefit rule parameters separately for each year, $b_{t0}^{s}, b_{t1}^{s}, b_{t2}^{s}, b_{t3}^{s}, b_{t4}^{s}$. The approximation given by (7) fits the monthly benefit data quite well, with R-squared statistics for the first line segment mostly above .99 and for the second, mostly above .95.[14] Given the estimates of the benefit rule parameters, we then estimated (8), the evolutionary rule.

Table 2 transforms the benefit parameters obtained from the estimates of (7) into a more convenient set of benefit measures, namely the total monthly income of non-working women (with zero non-earned income) who have either one, two or three children and the total monthly income of women with two children who have part-time earnings of 500 dollars or full-time earnings of 1000 dollars.[15] Referring to table 2, among the six states, NY, CA and MI are considerably more generous than NC, OH and TX. Among the first group Michigan is the most generous, with average benefits over the 24 years for a woman with one child being 699 (1987 NY) dollars per month, and among the second group Texas is the least generous, with the same average benefits figure only 228 dollars. NY and CA were about equally generous on average over the period as were NC and OH. Benefit reduction rates, net of child-care allowances, are fairly high. For example, a woman who had two children and earned 500 dollars per-month while working part-time would have kept 80 per cent of her earnings if she resided in Texas and between 60 and 70 per cent if she resided in any of the other five states.[16] If instead she had earned 1,000 dollars per month working full-time, her retained portion would have been as high

[14] These regressions are available on request.

[15] See appendix table A.2 for summary statistics of the actual parameters themselves.

[16] Benefit reduction rates for AFDC and for Food Stamps are federally set. The reason they implicitly differ among states in the tables is due to our approximation and the fact that AFDC payments terminate at different income levels among the states while food stamp payments are still non-zero and the two programs have different benefit reduction rates. There is thus a kink in the schedule of total welfare payments with income that our approximation smooths over.

as 70 percent in Texas and as low as 51 percent in New York and Michigan.

Table 2 also presents evidence on the trends in welfare benefit levels. The general pattern is a steep decline in benefit amounts between 1975 and 1985 and relative constancy thereafter. For example, in Michigan monthly benefits fell from $1108 for a woman with no earnings and two children in 1975 to $704 in1985 and to $694 in 1995. For the same woman with $500 in monthly earnings, benefits would have fallen from $966 in 1975 to $405 in 1985, and then risen slightly to $495 in 1990.

Table 3 shows the choice distribution disaggregated by state (for each of the five states used in the estimation and for Texas) and race, averaged over all sample ages. It is interesting to contrast the simple cross-sectional ranking of states by welfare generosity and by behaviors. In alphabetical order, from table 2 the ranking by welfare generosity is (2, 1, 2, 4, 4, 5). The most naive behavioral model would imply that the ranking of states by welfare take-up rates should be the same as that by generosity. From table 3, that ranking is (2, 1, 5, 4, 3, 6) for whites and (2, 1, 4, 5, 3, 6) for blacks. Thus, only MI, CA, and TX are conformably ranked for both races.[17] In fact, the take-up rate in NY is one-quarter of CA's for whites and one-half for blacks, and is comparable to the take-up rate in NC. A similar comparison can be performed for the other behaviors. With respect to fertility, for example, a naive behavioral model would imply that fertility be positively related to welfare generosity. Table 3 shows that for teenage pregnancies, i.e., the average number of live births by the age of 19, the ranking of states is (3, 3, 6, 1, 2, 5) for whites and (3, 1, 3, 5, 5 , 2) for blacks, which does not conform very well to the naive prediction. Similarly, naive predictions with respect to marriage and employment are not generally validated

[17] We ignore Hispanics because they are only significantly represented in three states.

in the data.

IV. Estimation Method:

The (approximate) solution to the agents' maximization problem provides (polynomial approximations to) the Emax functions that appear on the right hand side of (10). The alternative-specific value functions, V_t^k for k=1,..,K, are known up to the random utility shocks, e_a^j (j = h, s, m, p, g), the wage offer shock of the woman and the earnings shock of the husband if the woman receives a marriage offer, v_a^j (j=w, m), the implicit shocks that determine whether a marriage offer is received and whether the woman will reside with her parents if she is not married, and the benefit parameter shocks in the evolutionary rule, $\mathbf{u}_t^s$. Thus, conditional on the deterministic part of the state space, the probability that an agent is observed to choose option k takes the form of an integral over the region of the space of a subset of the random shocks such that k is the preferred option. The specific subset depends on which option k is being considered. If option k corresponds to a work option, then the wage offer is observed by us, and the wage shock is not in the subset over which the integration occurs. In that case, the likelihood contribution for the observation also includes the density of the wage error. If the woman is married (living with parents), then the husband's (parents') income is observed by us, that shock is excluded from the integration and the likelihood contribution includes the husband's (parents') income density. The shocks in the benefit parameter evolutionary rules are always observed by us.

As noted, the choice set contains as many as 36 elements. It is well know that evaluation of choice probabilities is computationally burdensome when the number of alternatives is large. But in recent years, highly efficient smooth unbiased probability simulators, such as the GHK method (see, e.g., Keane (1993, 1994)), have been developed for these situations. Unfortunately, the GHK method, as well as other smooth unbiased simulators, rely on a structure in which there is a separate additive error associated with each alternative. Further, as discussed in Keane and Moffitt (1998), in estimation problems where the number of choices exceeds the number of error terms, the boundaries of the region of integration needed to evaluate a particular choice probability are generally intractably complex. Thus, given our model, the most practical method to simulate the probabilities of the observed choice set would be to use a kernel smoothed frequency simulator. These were proposed in McFadden (1987), and successfully applied to estimate a structural model with a large choice set in Keane and Moffitt (1998).[18]

But in the present context, this approach is not feasible because of severe problems created by unobserved state variables. Because, as we have noted, we do not have a complete history of employment, schooling or welfare take-up for most of the cohorts back to age 14, the state variables accounting for work experience, schooling and welfare dependence cannot be constructed. Parental co-residence is also observed only once a year as is marital status that takes into account spousal co-residence.

Further complicating the estimation problem, as also noted, the youth's initial schooling level at age 14 is observed only for one of the 16 cohorts. It has been well known since Heckman (1981) that unobserved initial conditions, and unobserved state variables more generally, pose formidable computational problems for estimation of dynamic discrete choice models. If some or all elements of the state space are unobserved, then to construct conditional choice probabilities one must integrate over the distribution of the unobserved elements. Even in much simpler

[18] Kernel smoothed frequency simulators are, of course, biased for positive values of the smoothing parameter, and consistency requires letting the smoothing parameter approach zero as sample size increases.

dynamic models than ours, such distributions are typically intractably complex.

In a previous paper (Keane and Wolpin (forthcoming)), we have developed an estimation algorithm that deals in a practical way with the problem of unobserved state variables. The algorithm is based on simulation of complete (age 14 to the terminal age) outcome histories for a set of artificial agents. An outcome history consists of the initial school level of the youth, S_{14}, along with simulated values in all subsequent periods for all of the outcome variables in the model (school attendance, part- or full-time work, marriage, pregnancy, welfare participation, the woman's wage offer, the husband's earnings, parents' income). The construction of an outcome history can be described compactly as follows:

At the current trial parameter value:

1) Draw the youth's initial schooling and parents' schooling from the joint distribution;

2) Draw the relevant set of random shocks necessary to compute the alternative-specific value functions at a=1;

3) Choose the alternative with the highest alternative-specific value function;

4) Update the state variables;

5) Repeat steps (2) – (4) for a=2, ... , A;

Repeat (1) - (5) N times to obtain simulated outcome histories for N artificial persons. Denote by $\tilde{O}^n$ the simulated outcome history for the nth such person, $\tilde{O}^n = (S^n_{14}, \tilde{O}^n_{a=1}, ..., \tilde{O}^n_{a=A})$, for n = 1,..., N.

In order to motivate the estimation algorithm, it is useful to ignore for now the complication that some of the outcomes are continuous variables. Let O^i denote the observed outcome history for person i, which may include missing elements. Then, an unbiased frequency simulator of the probability of the observed outcome history for person i, $P(O^{i})$, is just the fraction of the N simulated histories that are consistent with O^{i}. In this construction, missing elements of O^{i} are counted as consistent with any entry in the corresponding element of $\tilde{O}^{n}$. Note that the construction of this simulator relies only on unconditional simulations. It does not require evaluation of choice probabilities conditional on state variables. Thus, unobserved state variables do not create a problem for this procedure.

Unfortunately, this algorithm is not practical. Since the number of possible outcome histories is huge, consistency of a simulated history with an actual history is an extremely low probability event. Hence, simulated probabilities will typically be 0, as will thus be the likelihood, unless an impractically large simulation size is used (see Lerman and Manski 1981). In addition, the method breaks down if any outcome is continuous, e.g., the woman's wage offer, regardless of simulation size, because agreement of observed with simulated wages is a measure zero event.

We solve this problem by assuming, as is apt, that all observed quantities are measured with error. With measurement error there is a nonzero probability that any observed outcome history might be generated by any simulated outcome history. Denote by $P(O^{i} * \tilde{O}^{n})$ the probability that observed outcome history O^{i} is generated by simulated outcome history $\tilde{O}^{n}$. Then $P(O^{i} * \tilde{O}^{n})$ is the product of classification error rates on discrete outcomes and measurement error densities for wages and assets that are needed to make O^{i} and $\tilde{O}^{n}$ consistent. Observe that $P(O^{i} * \tilde{O}^{n}) > 0$ for any $\tilde{O}^{n}$, given suitable choice of error processes. The specific measurement error processes that we assume are described below. The key point here is that $P(O^{i} * \tilde{O}^{n})$ does not depend on the state variables at any age a, but only depends on the outcomes.

Using N simulated outcome histories we obtain the unbiased simulator

$$(11)\quad \hat{P}_N(O^i) = \frac{1}{N}\sum_{n=1}^{N} P(O^i|\tilde{O}^n).$$

Note that this simulator is analogous to a kernel-smoothed frequency simulator, in that $I(O^i = \tilde{O}^n)$ is replaced with an object that is strictly positive, but that is greater if $\tilde{O}^n$ is "closer" to O^i. However, the simulator in (11) is unbiased because the measurement error is assumed to be present in the true model.

It is straightforward to extend the estimation method to allow for unobserved heterogeneity. Assume that there are K types of women who differ in their permanent preferences for attending school, becoming pregnant and receiving welfare. In addition, women also differ in their human capital "endowment" at age 14 (the constant term in the wage functions) and in their potential husband's human capital stock. To handle unobserved heterogeneity (i.e. types) in this framework, define $p_{k|S_{14}}$ as the probability a person is type k given his initial school level, for k = 1,..., K, where K is the number of types. In this case, simulate N/K vectors $\tilde{O}_k^n$ for each type.[19] Then,

$$(12)\quad \hat{P}_N(O^i) = \frac{1}{N}\sum_{k=1}^{K}\sum_{n=1}^{N/K} P(O^i|\tilde{O}^n)\frac{p_{k|S_{14}}}{N/K}.$$

Observe that in (12), the conditional probabilities $P(O^i|\tilde{O}^n)$ are weighted by the ratio of the proportion of type k according to the model, $p_{k|S_{14}}$, to the proportion of type k in the simulator,

---

[19] Initial schooling is exogenous conditional on type. We also take the parents' schooling as an initial condition exogenous conditional on type.

N/K.

Note that this simulator is smooth in the model parameters if simulated outcome histories are held fixed and re-weighted as parameters are varied. Given an initial parameter vector $?_0$ and an updated vector $?^{)}$, the appropriate weights are the ratio of the likelihood of the simulated history under $?^{)}$ to that under $?_0$. Such weights have the form of importance sampling weights (i.e., the ratios of densities under the target and source distributions), and are smooth functions of the model parameters. Further, it is straightforward to simulate the likelihood of an artificial history using conventional methods because the state vector is fully observed at all points along the history. $P(O^{i}*\tilde{O}^{n})$ is simulated using a kernel smoothed frequence simulator and standard errors are obtained with the BHHH algorithm.

Lastly, it is necessary to describe the specific assumptions for the measurement error processes. First, we assume that discrete outcomes are subject to classification error. The structure we adopt is simply that there is some probability that the reported response category is the truth and some probability that it is not.[20] Second, we assume that the continuous variables are also subject to measurement error. In particular, we assume that the woman's wage offer error and the husband's income error are multiplicative and the parents' income error is additive. Both of these measurement errors are assumed to be serially independent and independent of each other.

---

[20]To ensure that the measurement error is unbiased, the probability that the reported value is the true value must be a linear function of the predicted sample proportion (see the appendix A for details). Obviously, measurement error cannot be distinguished from the other model parameters in a non-parametric setting. As in the model without measurement error, identification relies on a combination of functional form and distributional assumptions, and exclusionary restrictions.

Parameterizations:

The solution/estimation of the model requires the choice of explicit functional form and distributional assumptions. Because the solution of the model is numerical, functional forms need not be chosen for analytical convenience, but rather can be chosen for their correspondence to existing literature, their ease of interpretation and their ability to fit the data. Indeed, the exact specifications were not chosen *a priori*, but rather reflect an iterative specification search based on assessing the fit of the model to elemental aspects of the data.[21] The exact specifications are given in Appendix A.2.

V. Results (Very preliminary)

As a preliminary exercise, we have estimated the model on the subset of white women from 7 of the16 birth cohorts.[22] Parameter estimates are provided in appendix table A.3 (These parameter estimates should not be assumed to be convergent). Tables 4 and 5 show the within-sample fit for a number of the sample statistics. Overall, the fit to the choice distribution is quite good. As seen in table 4, the model captures well the age patterns in school enrollment, fertility, part- and full-time work and (accepted) wage rates, marriage, welfare participation and husband's income (given marriage). Transition patterns, as seen in table 5, fit the data reasonably well. However, the model systematically overstates the degree to which school enrollment is continuous; there are too few periods in which a woman is predicted to return to school and too

[21] Although this method of iterating between model specification and model fit clearly contaminates statistical measures of model fit, it would seem that such a strategy is unavoidable given the complexity of the behaviors that we model.

[22] Women residing in the same state but of different birth cohorts face different welfare realizations over their life cycles.

many periods of consecutive enrollment. In addition, the model understates the extent of consecutive periods of welfare participation. In the data, 17.5 percent of the periods a woman who is on welfare exits the next period, while the model predicts that to occur in 27.4 percent of the periods.[23]

The model's estimates also indicate considerable heterogeneity. The estimation assumes six distinct types with respect to preferences and opportunities. As table A.4 indicates, types 1 and 2 are highly educated and have the most educated parents, have a strong attachment to the work force and very rarely are on welfare. The opposite is true for types 5 and 6, with types 3 and 4 being between the extremes. With respect to welfare participation, at the extremes, type 6's have received welfare in almost 5 periods by the time they reach age 30 while type 1's have received welfare in .03 periods. In addition, by age 30 type 6's have .5 more children, have 4 fewer years of schooling, have worked about 7 years less at full-time jobs and are about 6 times more likely to have not ever been married.

The estimated model can be used to simulate a number of interesting counter-factual experiments. Given the preliminary nature of these results, we present only three such experiments as illustrations. In table 6, the first column presents predicted baseline statistics for the whole sample at age 30 and for the sample of women at age 30 for whom neither parent is a high school graduate.[24] The first experiment eliminates welfare entirely, the second reduces the

---

[23] It is possible that the mode of collection for the welfare participation data overstates the degree of permanence.

[24] The parents' schooling variable is defined as the maximum of the highest grade completed of either parent. If one of the parent's schooling is missing, the one that is not missing is used.

income of potential husband's by 20 percent and the third provides a 25 percent subsidy to attending college.

As seen in table 6, eliminating welfare reduces the number of children at age 30 by 7.3 percent overall and by 9.6 percent for the women whose parents are the least educated. Completed schooling increases by .3 years for both groups. The extent of part-time work up to age 30 falls by on average 2.5 periods overall and by 5.4 periods for the group of high school dropout parents, which is compensated for by an increase in the number of periods of full-time work of 2.3 periods and 6.2 periods respectively. As one would expect, the percent never married falls for both groups, by 6 percent for the first and by 7.3 percent for the second. As this counter factual simulation suggests, the impact of eliminating welfare on behaviors is greater among individuals more likely to receive welfare.

To summarize the other experiments briefly, the impact of reducing the income of potential husband, and thus the value of marriage, is to increase the proportion of women who have never been married by age 30, by double or more, to reduce the level of completed schooling, by about .5 years, and to increase the number of periods on welfare, by about 60 percent for both groups. It appears that an important investment component of schooling is to obtain marital offers from husbands with higher income. The reduced market opportunities, wage offers, that results, together with the reduced gains to marriage, increases the propensity to rely on welfare. The effect on overall fertility is small.

The effect of a 25 percent college tuition subsidy on completed schooling is substantial, increasing schooling by 1.2 years overall and by .9 years for the high school dropout parent group. The number of periods on welfare falls by about 15 percent for both groups and an increase in full-time work experience by 1 to 1.5 periods. As is consistent with the previous experiment, there is also a substantial reduction in the proportion of never married women, by about 50 percent. As in the previous experiment, there is only a slight effect on fertility.

Appendix B

Exact functional forms

In addition to the prior notation, it is useful to define hours spent not in leisure as $h_a^* = h_a + ?_1 N_a^* + ?_2 s_a + ?_3 g_a$, where the first term is hours worked, the second reflects the time cost of children, the third the time cost of attending school and the last the time cost of participating in the welfare program (meeting with case workers, etc.). Age at marriage is denoted as a_m. There are assumed to be k=1,...,K types. $I(\cdot)$ is the indicator function equal to one if the term inside the parentheses is true and zero otherwise.

Utility Function:

$$
\begin{aligned}
(\text{A.1}) \quad u_a = {} & c_a + \sum_{k=1}^{K} a_{1k} h_a^* + \sum_{k=1}^{K} a_{2k} s_a + \sum_{k=1}^{K} a_{3k} m_a + \sum_{k=1}^{K} a_{4k} g_a + \sum_{k=1}^{K} a_{5k} p_a \\
& + a_6 (h_a^*)^2 + a_7 h_a^* c_a + a_8 h_a^* c_a m_a + a_9 h_a^* c_a N_a + a_{10} h_a^p s_a + a_{11} h_a^f s_a \\
& + a_{12} h_a^p h_{a-1}^p + a_{13} h_a^f h_{a-1}^f + a_{14} h_a^p s_a I(S_a \geq 12) + a_{15} h_a^f s_a I(S_a \geq 12) \\
& + a_{16} s_a I(a<16) + a_{17} s_a I(a<18) + a_{18} s_a p_a + a_{19} s_a (1 - s_{a-1}) \\
& + a_{20} s_a (1 - s_{a-1}) I(S_a \geq 12) + a_{21} m_a I(a<21) + a_{22} m_a I(a<25) \\
& + a_{23} p_a I(a<18) + \sum_{j=0}^{3} a_{24+j} p_a a^{j+1} + a_{28} g_a g_{a-1} + a_{29} N_a + a_{30} N_a^2 \\
& + e_a^h h_a^* + e_a^s s_a + e_a^m m_a + e_a^p p_a + e_a^g g_a
\end{aligned}
$$

Woman's Human Capital Stock Function:

$$
\begin{aligned}
(\text{A.2}) \quad ?_a = {} & \sum_{k=1}^{K} ?_{0k} + ?_1 S_a + ?_2 S_a^2 + ?_3 H_a + ?_4 H_a^2 + ?_5 h_{a-1}^p + ?_6 h_{a-1}^f \\
& + ?_7 a + ?_8 I(a<16) + ?_9 I(a<18) + ?_{10} I(a<21)
\end{aligned}
$$

Husband's Human Capital Stock Function:

$$
(\text{A.3}) \quad ?_a^m = \overline{\mu_0^m} + f\, ?_{0k} + ?_1^m S_{a_m} + ?_2^m a_m + ?_3^m a_m^2 + ?_4^m (a - a_m)
$$

where $\overline{\mu_0^m}$ is the mean of the distribution of the husband's permanent component of human capital. μ_0^m is normal with variance $s^2_{\mu_0^m}$.

Marriage Offer Probability Function:

$$
(\text{A.4}) \quad 1 - p_a^m = \{1 + \exp[\sum_{k=1}^{K} p_{0k}^m + p_1^m a + p_2^m a^2 + p_3^m I(a>30) + (p_4^m + p_5^m (a - a_m)) m_{a-1} + p_6^m g_{a-1}]\}^{-1}
$$

Parental Co-Residence Probability Function:

$$\text{(A.5)}\quad 1 - p_a^z = \{1 + \exp[\sum_{k=1}^{K} p_{0k}^z + p_1^z a + p_2^z I(a<18) + p_3^z I(a<21) + p_4^z I(a<25) + p_5^z z_{a-1}]\}^{-1}$$

Job Offer Probability Functions:

$$\text{(A.6)}\quad \begin{aligned} 1 - p_a^{wp} &= \{1 + \exp[p_0^{wp} + p_1^{wp} h_a^p h_{a-1}^p + p_2^{wp} g_{a-1}]\}^{-1} \\ 1 - p_a^{wf} &= \{1 + \exp[p_0^{wf} + p_1^{wf} h_a^f h_{a-1}^f + p_2^{wf} g_{a-1} + p_3^{wf} I(a<21)]\}^{-1} \end{aligned}$$

Income Sharing Functions:

$$\text{(A.7)}\quad \begin{aligned} t_a^z &= t_0^z + t_1^z I(a<16) + t_2^z I(a<18) + t_3^z s_a I(S_a \geq 12) + t_4^z s_a I(S_a \geq 16) \\ t_a^m &= t_0^m \end{aligned}$$

Type probability functions:

$$\text{(A.8)}\quad \begin{aligned} p_k^T &= \frac{\exp[p_{0k}^T + p_{1k}^T S^P + p_{2k}^T S_{14} + p_{3k}^T I(S^P > 12)]}{1 + \sum_{k=1}^{K} \exp[p_{0k}^T + p_{1k}^T S^P + p_{2k}^T S_{14} + p_{3k}^T I(S^P > 12)]} \quad \text{for } k = 2,..,K; \\ \sum_{k=1}^{K} p_k^T &= 1 \end{aligned}$$

Joint Initial Schooling, Parent Schooling Probability Distribution Function:

$$\text{(A9)}\quad \begin{aligned} \Pr(S_{14}=j,\ S^P=k) &= \frac{\exp[?_{0j} + ?_1 S^P S_{14}]}{1 + \sum_j \exp[?_{0j} + ?_1 S^P S_{14}]} \Pr(S^P=k) \\ \Pr(S^P=k) &= \sum_j \Pr(S_{14}=j,\ S^P=k) \end{aligned}$$

where k=7,...,20 and j = 6, 7, 8, 9 if the youth first reaches age 14 as of July 1 of a given year and j = 6.5, 7.5, 8.5, 9.5 if the youth first reaches age 14 as of January1 of a given year.

Error Distribution:

$$\begin{pmatrix} e_a^h \\ e_a^s \\ e_a^m \\ e_a^p \\ e_a^g \\ e_a^w \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} s_h^2 & & & & & \\ s_{hs} & s_h^2 & & & & \\ s_{hm} & s_{sm} & s_m^2 & & & \\ s_{hp} & s_{sp} & s_{mp} & s_p^2 & & \\ s_{hg} & s_{sg} & s_{mg} & s_{pg} & s_g^2 & \\ s_{hw} & s_{sw} & s_{mw} & s_{pw} & s_{gw} & s_w^2 \end{pmatrix} \right)$$

Classification Error Rates:

Consider first the classification error process for school attendance:

$?_{0a}^{s} =$ probability that school attendance is correctly recorded at age a.

$?_{1a}^{s} =$ probability that school attendance is reported when person did not attend school.

$$?_{0a}^{s} = Es + (1 - Es)\, f(s_a = 1)$$

$$?_{1a}^{s} = (1 - ?_{0a}^{s})\; f(s_a = 1)/[1 - f(a_t = 1)]$$

where $f(s_a = 1) = \frac{1}{N}\sum_{i=1}^{N} I(s_a = 1)$ in the simulation and Es is a parameter to be estimated.

Similar classification error processes are assumed for all other decision variables, and for living with parents, initial schooling and parents' schooling. Following previous notation, the corresponding parameters are Ew^f, Ew^p, Em, Eg, Ep, Ez, ES_{14}, ES^P.

Measurement Error in Hourly Wages:

$$w_a^{f,observed} = w_a^f \exp\{e_a^{w,m}\}$$

$$w_a^{p,observed} = w_a^p \exp\{e_a^{w,m}\}$$

$$e_a^{w,m} \sim N(0, s_{w,m}^2)$$

Table 1: Choice Distribution by Age and Race: Five States Combined

	Attending School			Working (PT or FT)			Married			Becomes Pregnant			Receives AFDC		
Age	W	B	H	W	B	H	W	B	H	W	B	H	W	B	H
14	100	93.3	100	14.3	10.5	12.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	97.7	100	100	11.4	9.9	5.2	0.0	0.0	0.0	1.0	3.4	1.0	1.0	1.3	0.0
16	88.3	87.5	90.3	30.0	14.5	19.3	3.0	1.0	2.9	3.1	3.8	2.1	1.0	1.0	1.0
17	84.6	80.7	79.2	50.0	26.9	32.4	8.7	1.4	6.4	5.6	5.3	2.5	1.3	2.5	2.3
18	42.8	50.9	41.5	63.0	32.6	50.7	16.4	3.7	11.9	3.7	4.5	6.7	2.6	9.0	3.3
19	32.5	32.1	27.1	65.6	43.4	51.2	24.9	7.1	19.9	4.5	8.6	5.6	3.6	15.6	6.8
20	23.8	22.2	18.8	67.5	46.4	52.2	31.5	11.7	27.1	4.3	6.0	4.9	5.4	17.3	10.3
21	19.4	12.3	12.2	69.6	49.2	58.3	37.1	14.4	34.2	6.0	7.9	6.3	5.1	21.2	13.7
22	10.8	8.3	7.7	70.0	52.5	60.6	37.5	20.3	35.9	4.5	5.3	5.7	6.1	25.6	15.1
23	4.2	6.2	3.9	72.0	54.2	58.5	49.1	22.3	39.7	5.9	6.1	5.3	6.2	27.2	15.3
24	3.8	5.4	4.6	72.7	55.4	57.7	54.1	22.8	45.7	6.6	6.9	7.9	7.0	27.8	17.2
25	4.0	5.9	2.9	73.8	62.8	55.6	58.5	20.9	47.2	7.6	7.0	7.2	6.4	26.8	16.0
26-33	3.5	3.3	2.3	71.7	61.5	58.5	65.7	27.0	53.2	5.4	4.0	5.7	4.4	25.0	15.2

W=white, B=black, H=Hispanic

Table 2: Summary Statistics of Total Monthly Benefits By Numbers of Children and Earnings: 1967-1990

	Monthly Earnings					
	Zero		$500		$1000	
	One child	Two children	One child	Two children	One child	Two children
CA						
μ	608	754	417	591	154	279
σ	73	89	163	180	160	234
Min	469	596	226	378	0	2
Max	736	919	668	878	404	599
1970	471	599	471	599	404	599
1975	708	881	651	857	348	554
1980	616	757	405	560	156	311
1985	596	730	259	414	0	46
1990	594	728	301	475	0	113
MI						
μ	709	877	493	693	195	360
σ	165	186	229	256	209	269
Min	528	683	214	378	0	36
Max	988	1191	874	110	556	793
1970	723	899	655	871	366	582
1975	908	1108	729	966	391	629
1980	656	815	429	603	161	336
1985	551	704	237	405	0	58
1990	544	694	302	494	0	156

Table 2: cont.

NY						
μ	616	773	392	578	131	250
σ	115	138	189	220	142	236
Min	509	633	171	316	0	0
Max	872	1054	667	894	340	567
1970	608	786	533	749	248	464
1975	774	963	608	829	275	496
1980	545	678	325	474	67	215
1985	517	642	190	334	0	0
1990	523	649	233	399	0	37
NC						
μ	496	587	327	439	102	203
σ	64	77	125	143	110	142
Min	419	505	180	270	0	0
Max	618	738	523	676	275	418
1970	491	557	432	538	193	300
1975	617	738	523	676	253	406
1980	462	553	262	365	34	137
1985	454	543	200	296	0	71
1990	438	530	254	370	29	145

Table 2: cont.

OH						
μ	508	638	336	483	85	197
σ	53	73	136	161	99	157
Min	449	559	176	292	0	0
Max	611	779	523	710	250	418
1970	508	630	436	599	166	328
1975	610	778	523	710	206	392
1980	496	619	289	426	18	156
1985	458	570	187	306	0	30
1990	454	566	223	352	0	111
TX						
μ	379	483	265	380	69	169
σ	52	67	96	117	82	115
Min	301	375	147	227	0	20
Max	471	507	399	532	205	334
1970	444	559	372	523	122	273
1975	434	564	391	532	172	313
1980	334	435	202	298	3	99
1985	375	473	170	264	0	53
1990	343	442	186	289	9	113

Table 3: Choice Distribution by Race and State

	Whites						Blacks					
	CA	MI	NY	NC	OH	TX	CA	MI	NY	NC	OH	TX
Percent of Person-Periods												
Attending School	20.1	18.4	22.2	11.4	17.0	22.2	15.1	20.7	19.3	17.9	15.9	19.1
Working	66.4	61.8	68.0	75.0	65.1	72.5	51.3	44.1	44.6	60.8	50.7	63.4
Married	25.2	34.0	23.9	35.8	35.1	23.9	13.6	6.9	12.5	17.1	8.4	18.4
Becomes Pregnant	3.9	4.7	3.8	4.0	4.7	3.2	5.0	4.9	4.8	5.2	5.6	5.2
Receives Welfare	6.3	7.3	1.5	1.9	2.9	0.5	21.4	35.6	12.0	11.2	26.7	6.3
Average number of live births: by age 19	0.19	0.18	0.09	0.27	0.22	0.11	0.36	0.53	0.34	0.27	0.25	0.41
: by age 27	1.02	1.13	0.80	1.02	1.20	0.78	1.46	1.23	1.16	1.41	1.53	1.44
Number of Women	139	122	70	67	129	65	56	39	71	48	57	85
Number of Person-Periods	1482	3803	2112	2008	3933	1975	1729	1182	2088	1460	1811	2499
Mean Age	21.5	21.7	21.6	21.5	21.6	21.5	21.6	21.6	21.5	21.5	21.8	21.4

Table 3 cont.

	Hispanics					
	CA	MI	NY	NC	OH	TX
Percent of Person-Periods						
Attending School	16.4	-	19.6	-	-	19.2
Working	55.9	-	39.8	-	-	58.2
Married	26.5	-	18.6	-	-	29.4
Becomes Pregnant	5.7	-	4.6	-	-	5.1
Receives Welfare	12.0	-	13.0	-	-	4.2
Average number of live births: by age 19	0.26	-	0.24	-	-	0.30
: by age 27	1.42	-	1.29	-	-	1.28
Number of Women	224	3	65	0	6	135
Number of Person-Periods	6692	85	1889	0	214	4086
Mean Age	21.4	-	21.5	-	-	21.5

Table 4: Actual and Predicted Choices at Selected Ages

	Enrolled in School		Children Ever Born		Works Part-Time		Works Full-Time		Is Married	
Age	A	P	A	P	A	P	A	P	A	P
14	1.00	.999	0.00	0.00	a	.082	a	.002	.000	.000
15	.977	.997	0.01	0.02	.100	.163	.014	.002	.006	.000
16	.883	.944	0.02	0.05	.275	.305	.025	.015	.034	.036
17	.846	.825	0.06	.009	.410	.419	.090	.091	.087	.111
17.5	.677	.669	0.07	0.11	.423	.473	.145	.190	.141	.137
18	.428	.435	0.13	0.14	.381	.451	.249	.291	.164	.158
20	.237	.168	0.28	0.30	.270	.316	.405	.501	.315	.308
22	.108	.099	0.45	0.50	.199	.252	.501	.554	.375	.409
24	.038	.055	0.67	0.74	.147	.201	.580	.561	.541	.525
26	.034	.045	0.94	0.97	.191	.220	.555	.544	.605	.590
28	.024	.040	1.19	1.22	.186	.209	.512	.505	.678	.686
30	.042	.028	1.29	1.51	.188	.166	.515	.485	.675	.725

a. less than 50 observations
b. six-months

Table 4: Cont.

	Receives Welfare		Accepted Part-Time Wage		Accepted Full-Time Wage		Husband's Income	
Age	A	P	A	P	A	P	A	P
14	0.00	.000	a	2.97	a	4.13	a	a
15	.000	.004	a	3.15	a	3.79	a	a
16	.008	.010	4.08	3.85	a	3.68	a	5789
17	.013	.018	4.32	4.02	4.90	4.91	a	7405
17.5	.018	.024	4.59	4.05	5.19	5.04	a	8070
18	.026	.027	4.49	4.36	5.41	5.49	6941	8399
20	.054	.052	5.20	4.53	5.93	5.88	10305	9288
22	.061	.068	6.13	5.50	6.97	7.60	9890	10596
24	.070	.071	6.41	5.85	7.98	8.20	12309	11840
26	.068	.047	7.04	6.01	9.67	8.79	13743	13259
28	.039	.040	7.63	6.45	8.92	9.38	15201	14545
30	.032	.044	8.27	7.04	9.03	10.02	13242	16005

Table 5: Actual and Predicted Transition Rate
For Selected Variables

	Age=a+1			
	Actual		Predicted	
Age=a	0	1	0	1
Enrolled				
0	.973	.027	.990	.010
1	.227	.773	.125	.875
Work Part-Time				
0	.850	.150	.837	.163
1	.495	.505	.525	.475
Work Full-Time				
0	.863	.137	.855	.145
1	.145	.855	.176	.824
On Welfare				
0	.990	.010	.987	.013
1	.175	.825	.274	.726
Is Married				
0	.970	.030	.935	.065
1	.021	.979	.041	.959

Table 6: Counterfactual Experiments

	Baseline		Eliminate Welfare		Reduce Potential Husband's Income By 20 Percent		2 25% College Subsidy	
	All Women	Parents HS Dropouts	All Women	Parents HS Dropouts	All Women	Parents HS Dropouts	All Women	Parents HS Dropouts
Welfare Receipt (Age 30)								
Mean No. Periods	1.39	3.00			2.24	4.92	1.18	2.55
Percent								
0 Periods	78.0	59.5			71.3	50.7	80.4	62.7
1-4 Periods	12.0	19.1			13.0	14.8	11.1	18.4
5-10 Periods	5.3	11.1			6.9	14.1	4.7	9.9
11-20 Periods	3.9	8.3			6.9	15.9	3.2	7.6
21+	0.8	2.1			1.8	4.5	0.7	1.9
Children Ever Born (Age30)								
Mean	1.51	1.66	1.40	1.50	1.49	1.65	1.49	18.9
Percent 0	20.6	17.7	22.5	20.2	21.0	17.7	2.16	59.6
1-2	60.0	59.5	61.4	62.8	60.0	58.9	60.5	20.3
3-4	18.4	21.0	15.7	16.2	18.0	21.8	16.9	1.6
5+	1.0	1.7	0.5	0.9	1.0	1.6	11.0	
Highest Grade Completed (Age 30)								
Mean	13.2	12.0	13.5	12.3	12.5	11.5	14.4	12.9
Percent								
HS Dropout	14.3	27.7	10.4	21.5	24.8	45.0	8.5	16.8
HS Graduate	43.1	53.1	41.7	54.4	44.7	45.4	28.1	42.0
Some College	25.2	17.0	25.6	20.5	21.2	8.7	26.5	29.1
College Graduate	17.5	2.2	22.3	3.6	9.3	0.9	36.9	11.3

Table 6: Cont.

Employment (Age 30)								
Mean No. Periods Worked								
Part-Time	8.1	10.1	7.9	9.8	8.5	10.7	8.1	9.8
Full-Time	13.1	8.9	13.5	10.2	13.4	8.9	14.2	10.4
Percent Working								
Part-Time	16.6	23.6	14.1	18.2	19.3	24.1	14.9	20.0
Full-Time	48.5	29.6	50.8	35.8	51.1	32.2	58.6	39.7
Mean Wage Offer								
Part-Time	7.54	6.17	7.76	6.39	7.19	5.95	8.39	6.73
Full-Time	8.77	7.07	8.99	7.31	8.34	6.84	9.74	7.75
Mean Accepted Wage								
Part-Time	7.06	6.27	7.42	6.42	6.57	5.65	7.73	6.79
Full-Time	10.11	8.31	10.22	8.35	9.53	8.21	10.97	9.04
Marriage								
Percent								
Never Married	10.0	15.0	9.4	13.9	19.8	31.7	5.3	8.8
Married (Age 30)	72.5	66.5	73.4	67.3	63.1	53.0	72.1	65.5
Mean Age at First Marriage (if ever married)	22.2	22.7	22.3	22.7	23.3	24.0	22.6	22.7

Table A.1: Summary Statistics of Parameters of Benefits Rules by State: 1967-1990 (a,b)

		b_0	b_1	b_2	b_3	b_4
CA						
	μ	462	146	570	.65	179
	σ	59	18	115	.13	23
	Min	343	126	449	.46	152
	Max	554	182	785	.89	219
MI						
	μ	540	168	627	.67	200
	σ	144	23	197	.09	28
	Min	375	134	391	.53	146
	Max	786	205	955	.92	243
NY						
	μ	459	157	540	.67	186
	σ	95	30	142	.10	36
	Min	382	123	384	.51	144
	Max	690	199	767	.92	233

Table A.1: cont.

		b_0	b_1	b_2	b_3	b_4
NC						
	μ	405	91	458	.48	112
	σ	55	19	119	.11	20
	Min	332	53	296	.51	89
	Max	498	123	640	.92	153
OH						
	μ	379	129	476	.57	147
	σ	36	23	127	.10	26
	Min	337	107	309	.45	115
	Max	442	179	653	.88	187
TX						
	μ	274	104	362	.42	115
	σ	40	19	91	.10	24
	Min	206	55	233	.31	80
	Max	354	131	471	.72	151

a. 1987 NY dollars
b. Based on Monthly AFDC plus Food Stamp Benefits

Table A.2: Evolutionary Rules for Benefit Parameters[a]

	CA					MI				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}
$b_{0,t-1}$	.792 (.179)	.068 (.068)	.226 (.466)	-	.320 (.089)	-1.16 (.303)	-.230 (.072)	-2.36 (.395)	-	-.350 (.080)
$b_{1,t-1}$	.766 (.628)	.008 (.238)	2.61 (1.63)	-	.301 (.312)	3.65 (1.19)	.944 (.284)	3.60 (1.55)	-	.720 (.315)
$b_{2,t-1}$	-.281 (.183)	.049 (.069)	-.657 (.476)	-	-.237 (.091)	1.36 (.242)	.196 (.058)	2.43 (.316)	-	.250 (.064)
$b_{3,t-1}$	94.4 (43.4)	21.5 (16.7)	16.3 (11.3)	.625 (.160)	18.2 (21.7)	14.8 (94.6)	24.1 (22.7)	14.5 (.123)	.441 (.163)	-11.9 (25.0)
$b_{4,t-1}$	1.11 (.834)	.354 (.316)	5.45 (12.17)	-	1.73 (.415)	-1.54 (1.08)	-.232 (.258)	-1.36 (1.40)	-	.357 (.285)
Constant	10.7 (.114)	35.4 (43.4)	-41.0 (28.7)	-.235 (.105)	-.417 (56.8)	10.3 (.132)	72.5 (31.6)	.133 (.172)	-.361 (10.9)	31.2 (34.9)
R^2	.86	.79	.77	.46	.78	.95	.88	.96	.29	.91
P. Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Mean	462	146	570	.65	179	540	168	627	.67	200
RMSE	24.7	9.6	64.7	.089	12.4	36.7	8.8	49.7	.068	9.70

Table A.2: cont.

	NY					NC				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}
$b_{0,t-1}$	.781 (.346)	.-.141 (.047)	-.324 (.413)	-	-.087 (.070)	1.13 (.158)	.109 (.070)	1.76 (.409)	-	.259 (.090)
$b_{1,t-1}$	3.42 (2.54)	.509 (.341)	1.80 (3.03)	-	.311 (.512)	-1.37 (.657)	.361 (.290)	-4.08 (1.70)	-	-.796 (.374)
$B_{2,t-1}$	-.020 (.453)	.120 (.061)	.807 (.541)	-	.064 (.092)	-.035 (.139)	-.033 (.058)	-.129 (.339)	-	-.137 (.075)
$B_{3,t-1}$	-.112 (11.6)	.325 (15.5)	34.6 (13.9)	.476 (.173)	-31.5 (23.9)	162.3 (34.5)	53.1 (15.1)	73.5 (89.7)	.608 (.108)	43.7 (19.5)
$B_{4,t-1}$	-2.06 (1.35)	.167 (.182)	-.136 (1.61)	-	.610 (.273)	.459 (.845)	.318 (.374)	3.50 (2.19)	-	1.50 (.482)
Constant	-11.7 (16.1)	44.3 (21.7)	15.2 (19.2)	-.340 (.116)	4.89 (132.7)	110.4 (37.6)	20.9 (16.6)	-189.5 (97.3)	-.176 (.053)	3.71 (21.4)
R^2	.82	.97	.89	.30	.44	.96	.92	.94	.64	.91
P. Value	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Mean	459	157	540	.67	186	405	91	458	.48	112
RMSE	46.0	6.2	55.2	.078	9.61	12.6	5.5	32.2	.051	

Table A.2: cont.

	OH					TX				
	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}	b_{0t}	b_{1t}	b_{2t}	b_{3t}	b_{4t}
$b_{0,t-1}$	-.150 (.197)	-.035 (.112)	-	-	-.121 (.091)	.719 (.102)	-	-	-	-
$b_{1,t-1}$	-.136 (.319)	.605 (.182)	-	-	-.061 (.148)	-	.754 (.083)	-	-	-
$b_{2,t-1}$	.014 (.180)	.154 (.075)	.949 (.077)	.-	.164 (.061)	-	-	.855 (.057)	-	-
$b_{3,t-1}$	1.45 (38.1)	48.6 (21.9)	-	.573 (.126)	15.2 (17.7)	-	-	-	.720 (.079)	-
$b_{4,t-1}$	1.43 (.693)	-.349 (.395)	-	-	.378 (.321)	.-	.-	-	-	.804 (.057)
Constant	286.2 (60.6)	69.8 (34.7)	15.7 (38.5)	-.229 (.074)	74.3 (28.3)	76.9 (28.4)	27.1 (8.94)	44.4 (21.9)	-.103 (.035)	20.8 (6.80)
R^2	.85	.86	.09	.52	.92	.68	.57	.88	.69	.88
P. Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mean	379	129	476	.57	147	274	104	362	.42	115
RMSE	15.5	9.7	43.5	.059	8.5	23.8	11.0	32.3	.04	8.6

[a]benefit parameters in annual amounts

Table A.3 Parameter Estimates

Utility Function[a]

Hours		School		Marriage		Fertility		Welfare	
α_{11}	2.037	α_{21}	1.098	α_{31}	-12.194	α_{41}	.000	α_{51}	-1.348
		α_{22}	.905			α_{42}	.412	α_{52}	-1.348
		α_{23}	-.502			α_{43}	.271	α_{53}	-1.331
		α_{24}	-.702			α_{44}	.436	α_{54}	-1.362
		α_{25}	-1.452			α_{45}	1.193	α_{55}	-.582
		α_{26}	-1.652			α_{45}	1.062	α_{56}	-.242
α_{6}	-.0008	α_{7}	4.20	α_{8}	-.90	α_{4}	-.66		
α_{10}	-.75	α_{11}	-1.20	α_{12}	.20	α_{13}	.74		
α_{14}	.25	α_{15}	1.20	α_{16}	1.02	α_{17}	.53	α_{18}	-1.20
α_{19}	-6.585	α_{20}	.518	α_{21}	.54	α_{22}	1.98	α_{23}	-1.00
α_{24}	.601	α_{25}	-.133	α_{26}	.0074	α_{27}	-.00013		
α_{28}	.55	α_{29}	.18	α_{30}	-.045				

[a] Utility function parameters should be multiplied by 1000, except α_{11}, α_{6}, α_{7}, α_{8}, α_{9}.

Table A.3: Cont.

Wage Function

ψ_{01}	7.649	ψ_1	.0908	ψ_7	.0055
ψ_{02}	7.648	ψ_2	-.0075	ψ_8	-.1205
ψ_{03}	7.549	ψ_3	.0131	ψ_9	-.0532
ψ_{04}	7.550	ψ_4	-.00009	ψ_{10}	-.1609
ψ_{05}	7.449	ψ_5			
ψ_{06}	7.449	ψ_6	.040		

Husband Wage Function

Ψ_0^m	6.871	ψ_1^m	.030
ϕ	2.0	ψ_2^m	.0904
σ_μ^2	.385	ψ_3^m	-.0796
σ_ε^2	.20	ψ_4^m	.0475

Marriage Offer Pros

π_0	-1.864	π_4	4.029	π_1	.125	π_2	-.0025
π_3	-.678	π_5	.039	π_6	-.875		

Table A.3: Cont.

Parental Co-Residence

π_0	-.20	π_1	-.08	π_2	2.05	π_3	.60
π_4	-.30	π_5	4.0				

Job Offer Probabilities

π_0^p	2.192	π_1^p		π_2^p	.70		
π_0^f	.6972	π_1^f	1.394	π_2^f	.70	π_3^f	-.50

Income Sharing Functions

τ_0^z	-1.30	τ_1^z	-.20	τ_2^z	-.20	τ_3^z	.065	τ_4^z	.030
τ_0^m	.20								

Type Probabilities

π_{02}	.00	π_{12}	.00	π_{22}	.00	π_{32}	.00
π_{03}	5.78	π_{13}	-.40	π_{23}	-.80	π_{33}	-.90
π_{04}	5.78	π_{14}	-.40	π_{24}	-.80	π_{34}	-.90
π_{05}	10.25	π_{15}	-.80	π_{25}	-1.60	π_{35}	-.90
π_{06}	10.25	π_{16}	-.80	π_{26}	-1.60	π_{36}	-.90

Table A.3: Cont.

Joint Initial School, Parents School Distribution

δ_{01}	2.42	δ_1	.19
δ_{02}	2.73		
δ_{03}	3.24		
δ_{04}			

Table A4: Selected Characteristics by Type

	Type 1	2	3	4	5	6
Mean No. Periods On Welfare Age 30	.03	.06	.27	.18	3.39	4.97
Mean No. Children At Age 30	1.37	1.32	1.52	1.41	1.66	1.87
Mean Highest Grade Completed	15.3	14.8	12.6	12.6	11.7	11.5
Mean Work Experience at Age 30						
Part-Time	5.8	5.5	7.8	8.0	10.8	11.1
Full-Time	19.1	19.2	13.3	13.0	6.7	5.8
Proportion Never Married at Age 30	.033	.033	.081	.087	.191	.205
Mean Initial Schooling	8.4	8.4	8.2	8.2	7.9	8.0
Mean Parents' Schooling	14.9	14.9	12.3	12.2	10.9	11.0
Sample Proportions	.178	.173	.162	.174	.166	.148