Life-Cycle Fertility and Human Capital Accumulation *

George-Levi Gayle
Tepper School of Business, Carnegie Mellon University

Robert A. Miller
Tepper School of Business, Carnegie Mellon University

July 2012

Abstract

This paper analyzes the impact of policies expenditures on offspring and child care costs investigating on life cycle fertility and female labor supply. We investigate subsidizing time spent on subsidizing expenditure on offspring, child care, both through a wage and also by directly child care, paying women who bear children a wage, and retraining them when they reenter the labor force after time spent out to raise children. To analyze these policies we formulate and estimate a dynamic model of labor supply and fertility. The model accounts for maternal time spent raising offspring, plus the effect of time spent on current and summed discounted expenditures on them, We estimate the model with the PSID data, and solve for the policy functions with the estimated parameters perturbed by policy innovations. Generally speaking all these policies have a positive impact on fertility on almost all socioeconomic groups but retraining has the most pronounced increases on the birth rate.

1. INTRODUCTION

Both female labor supply and fertility behavior are topical issues of public interest. For example, the worldwide declining rates of fertility, especially amongst educated women, has consequences for intergenerational wealth transfers, along with the demand for public infrastructure and privately produced goods. And the persistence of the gender gap in U.S. wages, after a long period of shrinking, may have implications for employment discrimination laws, and is a topic of continuing research for labor economists. Sociologists, demographers and economists recognize that female labor supply and fertility behavior are intertwined. So in principle public policies affecting fertility should also affect female labor supply, and

*Gayle acknowledges support from the Andrew Mellon Research Fellowship, while Miller was supported by National Science Foundation Award SES0721098. We thank Elizabeth Powers for her comments, and we have benefited from presentations at the Universities of Essex, Illinois (Urbana-Champaign), Kansas, Pittsburgh Wisconsin and Cowles Foundation, Yale.

vice versa. But quantifying the effects of such policies and their implementation is quite challenging.

Social scientists have drawn upon all the usual tools in their attempts to predict how public policies affect fertility and female labor supply.[1] Public opinion data, such as survey responses to hypothetical counterfactuals and ideal family size provide a first pass at how populations might react to policy innovations (European Commission,1990; Goldstein, Lutz and Testa, 2003). Time series analysis, for example over the post war period, have been used to estimate the role of substitution and wealth effects of increasing female wages on labor supply and fertility (Butz and Ward, 1979, 1980; Buttner and Lutz,1990). Cross sectional studies, such as between OECD countries compare the effects of different policies across countries to address these issues (Billari and Kohler, 2004; Kogel, 2004). Event studies, say related to the adoption of new programs have also been analyzed. (Milligan, 2005; Laroque and Salanie, 2008; Cohen, Dejejia and Romanov, 2010).

Our work joins a handful of studies that recognize the dynamic interactions between female labor supply and fertility by modeling and estimating the sequential determination of these joint events with panel data (Hotz and Miller,1988; Francesconi, 2002; Keane and Wolpin 2010; Adda, Dustman and Stevens, 2011). The latter two also conduct counterfactual policy simulations. Keane and Wolpin investigate changes to the welfare system, while Adda et al. simulate the effects of increasing child allowances. We conduct counterfactual simulations on four policies: Paying for expenditure on offspring; providing child care; paying women a wage to bear children; retraining mothers who quit the labor force when they reenter it.

To analyze these policies we formulate and estimate a dynamic model of labor supply and fertility. The model accounts for maternal time spent raising offspring and the effect of time spent on current and summed discounted expenditures on them. We estimate the model with the PSID data, and solve for the policy functions with the estimated parameters perturbed by the counterfactual policy innovations.

Summarizing our results, all the policies we investigate increase total fertility rates (TFR) on almost all socioeconomic groups but do not affect labor force participation much. Retraining has the most pronounced increases in the birth rate, particularly amongst highly educated women. To amplify most of our 18 stratified groups have estimated TFR below replacement rate (say 2.1) under the current regime but if human capital lost from temporarily withdrawing from the labor force could be restored, than the TFR of all but one group (least educated unmarried white women) would rise to replacement rate.

As a practical matter, our model predicts that a large proportion of human capital from working experience is acquired within one working year. Therefore our model predicts that retrospectively paying women the difference between their wages in their first two years at work after returning to work following an absence from work to give birth, would go a long way to raising the fertility rate of the most educated workers. More generally, subsidizing this labor market outcome raises substantially fertility rates without affecting participation rates very much. This serves to emphasize a point on the first slide: that public policy on these issues must account for both the fertility and female labor market responses.

The next section provides the theoretical underpinnings to our empirical investigations,

[1] For example see the recent survey by Gauthier (2007).

by laying out a life cycle model of labor supply and fertility. Then in Section 3 we briefly summarize the sample of households used in our empirical work, which is drawn from the Panel Study of Income Dynamics (PSID). Section 4 explains our estimation strategy, while Section 5 reports our structural estimates. Then in Section 6 we conduct several policy simulations and summarize our findings. In Section 7 we conclude; all proofs and estimation details are contained in an Appendix.

2. A FRAMEWORK

In this model two kinds of human capital are accumulated, offspring and labor market experience. The benefits from bearing an additional child depend on the number and ages of its older siblings, while the time costs of raising the child are spread over several years. The value of past working experience is impounded in the current wage rate, and in addition leisure is not additively separable over time. Our model factors these considerations into a dynamic optimization problem of female labor supply and fertility behavior.

The model is set in discrete time, and measures the woman's age beyond adolescence with periods denoted by $t \in \{0, 1, \ldots, T\}$. The birth of a child at period t, a choice variable, is denoted by the indicator variable $b_{nt} \in \{0, 1\}$. There are two continuous choice variables, consumption x_{nt}, and hours worked in the labor force, denoted by $h_{nt} \in [0, 1]$. To capture nonlinearities in leisure and returns to labor market experience, we define the four discrete choice indicator variables that capture joint labor force participation and fertility choices as:

$$\begin{aligned} d_{1nt} &\equiv I\{h_{nt} = 0\} I\{b_{nt} = 0\}, \quad d_{2nt} \equiv I\{h_{nt} > 0\} I\{b_{nt} = 0\} \\ d_{3nt} &\equiv I\{h_{nt} = 0\} I\{b_{nt} = 1\}, \quad d_{4nt} \equiv I\{h_{nt} > 0\} I\{b_{nt} = 1\} \end{aligned}$$

where, for example, $I\{h_{nt} = 0\}$ is the indicator function for n staying out of the workforce in period t. Note that the choices are mutually exclusive and exhaustive, implying $d_{jnt} \in \{0, 1\}$ for $j \in \{1, \ldots, 4\}$ with $\sum_{j=1}^{4} d_{jnt} = 1$.

2.1. Preferences

Births contribute directly to household utility. We assume that the spacing of births is related to preferences by the household over the age distribution of its children, as captured by interactions in the birth dates of successive children. More specifically, let γ_0 denote the additional lifetime expected utility a household receives for its first child, let $\gamma_0 + \gamma_k$ denote the utility from having a second child when the first born is k years old, let $\gamma_0 + \gamma_k + \gamma_j$ denote the utility from having a third child when the first two are aged k and j years old, and so on. Thus the deterministic benefits from offspring to the n^{th} household in period t are:

$$u_{nt}^{(b)} \equiv b_{nt}(\gamma_0 + \sum_{k=1}^{\rho_b} \gamma_k b_{n,t-k} + \gamma_b \sum_{k=\rho_b+1}^{T} b_{n,t-k}) \tag{2.1}$$

Thus siblings k years apart are complementary in lifetime utility if $\gamma_k > 0$.

Apart from having utility for children, household utility also comes from its consumption of market goods, denoted x_{nt}, leisure, denoted l_{nt}, and some random disturbances. Our formulation incorporates both fixed and variable utility costs associated with working. We assume the utility loss the n^{th} female from working in period t are:

$$u_{nt}^{(l)} \equiv (d_{2nt} + d_{4nt})\, z_{nt}' B_0 + z_{nt}' B_1 l_{nt} + \sum_{s=0}^{\rho_l} \delta_s l_{nt} l_{n,t-s} \tag{2.2}$$

where z_{nt} is a vector that includes such variables as age, formal education, regional location, ethnicity and race. Thus B_0 is a parameter vector characterizing the fixed-costs of participating in the work force, and B_1 shows the effect of exogenous time-varying characteristics on the marginal utility of leisure. Preferences are increasing in leisure if:

$$z_{nt}' B_{11} l_{nt} + 2\delta_0 l_{nt} + \sum_{s=1}^{\rho_l} \delta_s l_{n,t-s} > 0$$

and concave if $\delta_0 < 0$. The parameters δ_s for $s = 1, ..., \rho$ capture intertemporal non-separabilities in preferences with respect to leisure choices. A value of $\delta_s < 0$ for $s = 1, ..., \rho$ means that leisure s periods ago increases the marginal utility of leisure, and results in less work and child care time today. Equivalently, a finding of $\delta_s < 0$ implies that current and past leisure time are substitutes where as $\delta_s > 0$ implies that current and past leisure time are complements.

The third component in utility is derived from current consumption. We denote by:

$$u_{nt}^{(x)} \equiv \alpha^{-1} x_{nt}^{\alpha} \exp(z_{nt}' B_2 + \epsilon_{0nt}) \tag{2.3}$$

the current utility from consumption of x_{nt} by household n in period t, and we assume ε_{0nt} is identically and independently distributed across (n, t). We also allow for idiosyncratic factors to affect the utility from making the four distinct economic choices by assuming there is a choice specific distrubance ϵ_{knt} that is identically and independently distributed across (k, n, t) as a Type 1 Extreme Value random variable.

Letting $\beta \in (0, 1)$ denote the subjective discount factor over time, we define realized lifetime utility as:

$$\sum_{t=0}^{T} \beta^t \left\{ u_{nt}^{(b)} + u_{nt}^{(l)} + u_{nt}^{(x)} + \sum_{k=1}^{4} d_{knt} \epsilon_{knt} \right\} \tag{2.4}$$

2.2. Costs and Constraints

Raising children requires market expenditure and parental time. We assume that the discounted cost of expenditures from raising a child is π, a parameter that varies with household demographics, and that a k year old requires nurturing time of ϕ_k up until age ρ_c, and a constant input per period denoted by ϕ from then on.[2] Letting c_{nt} denote the amount of time the n^{th} household spends nurturing children in the household, our assumption about nurturing implies:

$$c_{nt} = \sum_{s=0}^{t} \phi_s b_{n,t-s} \tag{2.5}$$

[2] Thus offspring are differentiated by market inputs but not by the input of their mother's time.

where $\phi_s = \phi$ for all $s > \rho_c$.[3]

Leisure in period t, denoted l_{nt}, is defined as the balance of time not spent at work or nurturing children. It follows that the time allocated between nurturing children, market work and leisure must obey the constraint:

$$1 = h_{nt} + l_{nt} + c_{nt} \tag{2.6}$$

where h_{nt} denotes the proportion of time worked in period t as a fraction of the total time available in the period.

Female labor market experience for the n^{th} household in our sample is embodied in the wage rate, denoted w_{nt}, and depends on labor market experience and the demographic variables z_{nt}. Let τ_{nt} denote the calendar year when the n^{th} female is t years old, and let $\omega(\tau)$ denote the wage of one efficiency unit of labor. Following the literature real wages are the product of $\omega(\tau)$ and an index capturing the number of efficiency units embodied in a worker; we assume the mapping from experience to the current wage rate in year τ_{nt} is given by:

$$w_{nt} = \omega(\tau_{nt}) \mu_n \exp\left[z'_{nt} B_3 + \sum_{s=1}^{\rho_w} (\delta_{1s} h_{n,t-s} + \delta_{2s} d_{2nt} + \delta_{2s} d_{4nt}) \right] \tag{2.7}$$

for some positive integer ρ. Thus Equation (2.7) shows that, in addition to the demographic variables, the current wage depends on past participation and past hours up to ρ periods ago.

Aside from the real wage wage $\omega(\tau)$, aggregate effects are transmitted through interest rates. We denote by $\lambda(\tau_{nt})$ the value of a consumption unit discounted back t periods, in other words the price of consuming in period τ_{nt} denominated in $(\tau_{nt} - \tau_{n0})$ consumption units, a notational convention we adopt so that the model can reflect our emphasis on the lifecycle rather than on aggregate factors. Valued at calendar date τ_{n0}, net transfers to household n at age t are then:

$$\lambda(\tau_{nt}) (x_{nt} + z'_{nt} \pi b_{nt} - w_{nt} h_{nt}) \tag{2.8}$$

2.3. Optimization

We finesse questions about how efficiently markets and government interventions together allocate resources in this economy by modeling behavior as the solution to a social planner's problem. For appropriately defined interest rates $\lambda(\tau)$ and real wage rates $\omega(\tau)$, shadow prices that reflect aggregate conditions and market clearance in general equilibrium, the planner's objective function is formed by summing the weighted expected value of utility defined by Equation (2.4) over the lifetime of the woman and subtracting the discounted

[3] This specification of maternal time inputs is broadly consistent with those considered in the literature. For example using data from time diaries, Hill and Stafford (1980) found that maternal time devoted to child care declines as the children age. Equation (2.5) implies that the child care process exhibits constant returns to scale in the number of existing children. The evidence on the importance of such scale economies is mixed; Lazear and Michael(1980) find evidence of large scale economies while Espenshade(1984) finds them to be small.

sum of expected net transfers each period defined by (2.8). Denoting by η_n^{-1} the social weight attached to individual n, Pareto optimal allocations are found by maximizing:

$$E_0\left[\sum_{t=0}^{T}\beta^t\left(u_{nt}^{(b)}+u_{nt}^{(l)}+u_{nt}^{(x)}+\sum_{k=1}^{4}d_{knt}\epsilon_{knt}\right)-\eta_n\lambda\left(\tau_{nt}\right)\left(x_{nt}+\pi b_{nt}-w_{nt}h_{nt}\right)\right] \tag{2.9}$$

with respect to $\{x_{nt},h_{nt},b_{nt}\}_{t=0}^{T}$, sequences of random variables that are successively measurable with respect to the information available at periods $t\in\{0,1,2,\ldots,T\}$, subject to the individual household time constraints (2.6) and childcare demands (2.5).[4]

Setting $\rho\equiv\max\{\rho_b,\rho_l+M,\rho_w\}$, the vector of state variables for the optimization problem are:

$$H_{nt}\equiv\left(t,z_{nt}',M_{nt},b_{n,t-\rho},...,b_{n,t-1},h_{n,t-\rho},...,h_{n,t-1},\epsilon_{0nt},\ldots,\epsilon_{4nt}\right)$$

Aside from demographics, H_{nt} captures the dependence of the current household state on lagged labor supply and birth choices. We denote the optimal choices solving (2.9) by $\{x_{nt}^o,h_{nt}^o,b_{nt}^o\}_{t=0}^{T}$, write d_{ntk}^o for the value of d_{knt} implied by $\left(h_{nt}^o,b_{nt}^o\right)$, and also set $h_{knt}\equiv h_k\left(H_{nt}\right)\equiv h_{nt}^o$ for each k where $h_{1nt}=h_{3nt}=0$.

As shown in Sections 4 and 6 our estimators and policy functions are based on the four discrete choices defined by birth and participation combinations, along with the first order conditions for the continuous choices for consumption and hours worked conditional on participation. Substituting (2.3) into (2.9) and differentiating with respect to x_{nt} yields the (logarithm of the) Frisch consumption demand functions:

$$\log x_{nt}^o=(\alpha-1)^{-1}\left(\log\eta_n+\log\lambda\left(\tau_{nt}\right)-z_{nt}'B_2-\epsilon_{0nt}\right) \tag{2.10}$$

Since the utility for x_{nt}^o is additively separable, its choice does not depend on the discrete choices d_{nt} or the disturbance vector $\left(\epsilon_{1nt},\ldots,\epsilon_{4nt}\right)$, so the remaining parts of the solution to the planning problem are determined separately.

We now define the deterministic components of current utility from leisure and births when any discrete choice $j\in\{1,\ldots,4\}$ is paired with $h_k\left(H_{nt}\right)$. Substituting the optimal

[4] There is a growing empirical literature that tests for deviations from Pareto optimal allocations, also described as efficient risk sharing, using panel data on individuals and households. See, for example, Altug and Miller (1990), Cochrane (1991), Mace (1991), Altonji, Hayashi and Kotilikoff (1995), Townsend (1994), Miller and Sieg (1997), and Mazzocco and Saini (forthcoming) Taken together, this body of work shows that, depending on how the population for an agent is defined (such as village or caste, family or dynasty), the restrictions imposed by Pareto optimal allocations are quite hard to reject with panel data, unless one assumes very limited forms of population heterogeneity, and also that preferences are strongly additive, two assumptions that are widely regarded by microeconomists as being implausible. As a practical matter there is little agreement amongst economists as precisely what departure from Pareto optimality should be adopted when estimating models of individual and household behavior off panels.

choice of hours worked into the implied utility from leisure:

$$
\begin{aligned}
u_1(H_{nt}) &\equiv \left(1-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)\left[z_{nt}'B_1+\delta_0\left(1-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)\right] \\
&\quad +\sum_{s=1}^{\rho_l}\delta_s\left(1-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)\left(1-h_{n,t-s}-\sum_{r=s+1}^{t}\phi_r b_{n,t-r}\right) \\
u_2(H_{nt}) &\equiv u_1(H_{nt})-h_2(H_{nt})\left[z_{nt}'B_1+2\delta_0\left(1-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)-\delta_0 h_2(H_{nt})\right] \\
&\quad -h_2(H_{nt})\sum_{s=1}^{\rho_l}\delta_s\left(1-h_{n,t-s}-\sum_{r=s+1}^{t}\phi_r b_{n,t-r}\right)+\eta_n\lambda(\tau_{nt})w_{nt}h_2(H_{nt}) \\
u_3(H_{nt}) &\equiv u_1(H_{nt})-\phi_0\left[z_{nt}'B_1+2\delta_0\left(1-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)-\delta_0\phi_0\right] \\
&\quad -\phi_0\left[\sum_{s=1}^{\rho_l}\delta_s\left(1-h_{n,t-s}-\sum_{r=s+1}^{t}\phi_r b_{n,t-r}\right)\right]-\eta_n\lambda(\tau_{nt})z_{nt}'\pi \\
u_4(H_{nt}) &\equiv u_3(H_{nt})-h_4(H_{nt})\left[z_{nt}'B_1+2\delta_0\left(1-\phi_0-\sum_{r=1}^{t}\phi_r b_{n,t-r}\right)-\delta_0 h_4(H_{nt})\right] \\
&\quad -h_4(H_{nt})\sum_{s=1}^{\rho_l}\delta_s\left(1-h_{n,t-s}-\sum_{r=s+1}^{t}\phi_r b_{n,t-r}\right)+\eta_n\lambda(\tau_{nt})w_{nt}h_4(H_{nt})
\end{aligned}
$$

Substituting in the optimal hours choices when the woman participates, we define the current t period expected value function for the leisure and birth choices as:

$$V(H_{nt})\equiv\max_{\{d_{ns}\}_{s=t}^{T}}E\left\{\sum_{s=t}^{T}\sum_{k=1}^{4}d_{kns}\beta^{s-t}\left[u_k(H_{ns})+\epsilon_{kns}\right]|H_{nt}|\right\}$$

Defining the conditional value function for each discrete choice as:

$$V_k(H_{nt})\equiv u_k(H_{nt})+E\left[\beta V(H_{n,t+1})\,|d_{ntk}=1,H_{nt}\right]$$

Bellman's principle implies that for all $j\in\{1,\ldots,4\}$ if $d_{knt}^{o}=1$ then:

$$V_k(H_{nt})+\epsilon_{knt}\geq V_j(H_{nt})+\epsilon_{jnt} \tag{2.11}$$

Finally, when the woman participates in the workforce, meaning $d_{2ns}+d_{4ns}=1$, then hours of work (which depend on whether there is a birth or not), h_{2nt} or h_{4nt},satisfy the first order condition:

$$z_{nt}'B_1+\lambda(\tau_{nt})w_{nt}+\frac{\partial}{\partial h_{2nt}}E\left[\beta V(H_{n,t+1})\,|d_{knt}=1,H_{nt}\right]=\sum_{s=1}^{\rho}\delta_s\left(1-\sum_{r=1}^{t-r}\rho_k b_{n,t-r-s}\right) \tag{2.12}$$

The estimation framework is directly based on our specification of wages (2.7), the Frisch demands for consumption (2.10), differences in the conditional valutaion functions (2.11) and the Euler equation that determines the number of hours work by women participating in the workforce (2.12).

3. Data

The data for this study are taken from the Family-Individual File, Childbirth and Adoption History File and the Marriage History File of the Michigan Panel Study of Income Dynamics (PSID). The variables used in the empirical study are h_{nt}, the annual fraction of hours work by individual n at date t; $\widetilde{w}_{nt}$, her reported real average hourly earnings at t; x_{nt}, real household food consumption expenditures; FAM_{nt}, the number of household members; $YKID_{nt}$, the number of children less than six years of age; $OKID_{nt}$, the number of children of ages between six and fourteen; AGE_{nt}, the age of the individual at date t; EDU_{nt}, the years of completed education of the individual at time t; $HIGH.SCH_{nt}$, completion of high school dummy; $BLACK$ and $HISPANIC$ race dummies for blacks and Hispanics, respectively; NE_{nt}, NC_{nt},SO_{nt}, which are region dummies for northeast, northcentral, and south, respectively, and MAR_{nt}, denoting whether a woman is married or not. The construction of our sample and the definition of the variables is described in greater detail in Appendix 3.

Table 1 contains summary statistics of our main variables. The sample has aged, household size has declined, and the decline is most pronounced amongst young children. The steep decline in household size over the two decades, and the aging evident in the sample, relative to aggregate trends in the US, largely reflects the sampling mechanism of the PSID. Thus we cannot infer any aggregate trend in fertility from this table. Household income has increased somewhat, but household consumption of food has declined. However, both food consumption and income per capita has increased over the sample period. More striking is the rise in female income, which greatly outstrips increases in household income. This is due to both higher wages and greater hours. Because schooling has not increased over the sample period, the number of years of formal education is not a factor in explaining aggregate trends in female wages and labor supply, or any changes that might have occurred in fertility.

4. Estimation

Our estimation strategy essentially follows Altug and Miller (1998) by extending their framework of female labor supply and human capital accumulation to incorporate choices about fertility. First we estimate the wage equation, and in the process recover the individual fixed effects from the wage equation. Then we estimate the social weights of the social planner's problem from the Frisch demand for consumption. Both that determine the conditional choice probability (CCP) mappings, which are estimated nonparameterically as a mapping of the fixed effects are arguments that along with the state variables and demographic characteristics that determine them. The structural parameters are estimated from equations that exploit the finite dependence properties of this model, and our standard errors account for the sequential estimation method.

4.1. Wages

We assume that the reported wage rate, denoted $\widetilde{w}_{nt}$ (for the n^{th} household in period t) measures the woman's marginal product in the market sector with error, so that:

$$\widetilde{w}_{nt} = g(\widetilde{A}_{nt}) \exp(\widetilde{\epsilon}_{nt}) \tag{4.1}$$

where the multiplicative error term in equation (4.1) is conditionally independent over people, the covariates in the wage equation and the labor supply decision. Taking logarithms on both sides of Equation (4.1), and then differencing, yields:

$$\triangle\widetilde{\epsilon}_{nt} = \triangle\ln(\widetilde{w}_{nt}) - \sum_{s=1}^{\nu}(\delta_{1s}\triangle h_{n,t-s} + \delta_{2s}\triangle d_{n,t-s}) - \triangle z'_{nt}B_3 - \triangle\omega_t \tag{4.2}$$

which we estimated with a linear instrumental variables estimator.

4.2. Consumption preferences

In our model, the effects of differences in wealth across households on their fertility and labor supply decisions is determined a single parameter, their weight in the social planner's problem. The inverse of their social weight is their marginal utility of wealth, and it can be estimated with household data on consumption. Taking logarithms of (2.10) and then first differencing yields:

$$(1-\alpha)^{-1} \triangle \epsilon_{0nt} = \triangle\ln(x_{nt}) - (1-\alpha)^{-1} \triangle z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t) \tag{4.3}$$

The assumptions in Section 2 imply that the unobserved variable ε_{5nt} is independent of individual specific characteristics, implying:

$$E(\left[\triangle\ln(x_{nt}) - (1-\alpha)^{-1} \triangle z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t)\,|z_{nt}\right] = 0$$

which can be exploited using a linear instrumental variable procedures similar to the estimated wage function.

4.3. Individual-specific effects

We assume the fixed effects η_n and μ_n are mappings of the household's permanent characteristics z_n, denoted by $\mu(z_n)$ and $\eta(z_n)$ respectively. They can be estimated nonparamatrically as level effects off the wage equation and the first order condition for consumption. Let:

$$\phi_{1nt} \equiv \ln(\widetilde{w}_{nt}) - \sum\nolimits_{s=1}^{\nu}(\delta_{1s}h_{n,t-s} + \delta_{2s}d_{n,t-s}) - z'_{nt}B_3 - \omega_t = \mu(z_n) + \widetilde{\epsilon}_{nt}$$

$$\phi_{2nt} \equiv -\left[\ln(x_{nt}) - (1-\alpha)^{-1}z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t)\right] = \eta(z_n) + \epsilon_{0nt} \tag{4.4}$$

By assumption both $\widetilde{\epsilon}_{nt}$ and ϵ_{0nt} are orthogonal to z_n, from which it follows that $\mu(z_n) = E[\phi_{1nt}\,|z_n]$ and $\eta(z_n) = E[\phi_{2nt}\,|z_n]$. We estimate $\mu(z_n)$ and $\eta(z_n)$ with Kernel regressions off the cross section, using consistent estimates of the wage and consumer preference parameters obtained in the previous stages of the estimation.

4.4. Labor force participation and fertility

Our estimation equations for labor force participation, hours worked and fertility behavior are based on the finite dependence property of the model, which provides a computationally convenient expression for the conditional valuation functions described in the next section, and the logit form of the conditional choice probabilities in the valuation functions. Finite dependence arises in this model because it is feasible for women to avoid pregnancy and not work each period, and from her perspective in the model, she would no longer care about her work or birth history if she has not worked for at least the previous ρ periods and all her children were at least ρ years old.

To demonstrate the finite dependence property, we now define four choice paths $\rho + 2$ periods into the future that a woman might take starting at period t, and the history of state variables they generate, denoted by $H^{(s)}_{knt}$ for $k \in \{1, \ldots, 4\}$ and $s \in \{1, \ldots, \rho + 2\}$, and defined as:

$$
\begin{aligned}
H^{(s)}_{1nt} &\equiv \left(z'_{nt+s}; h_{n,t-\rho+s}, ..., h_{n,t-1}, 0, ..., 0; b_{n,t-\overline{\rho}+s}, ..., b_{n,t-1}, 0, 1, 0, ..., 0\right)' \\
H^{(s)}_{2nt} &\equiv (z'_{nt+s}; h_{n,t-\rho+s}, ..., h_{n,t-1}, h_{2nt}, 0, ..., 0; b_{n,t-\overline{\rho}+s}, ..., b_{n,t-1}, 0, 1, 0, ..., 0)' \\
H^{(s)}_{3nt} &\equiv (z'_{nt+s}; h_{n,t-\rho+s}, ..., h_{n,t-1}, 0, ..., 0; b_{n,t-\overline{\rho}+s}, ..., b_{n,t-1}, 1, 0, ..., 0)' \\
H^{(s)}_{4nt} &\equiv (z'_{nt+s}; h_{n,t-\rho+s}, ..., h_{n,t-1}, h_{4nt}, 0, ..., 0; b_{n,t-\overline{\rho}+s}, ..., b_{n,t-1}, 1, 0, ..., 0)'
\end{aligned}
$$

Note that all four histories evolve by choosing $d_{1n,t+s} = 1$ for $s > 1$, namely not participating in the labor force and not giving birth. $H^{(s)}_{1nt}$ denotes the state variables for the problem at period $t + s$ when a woman makes choices $d_{1n,t+1} = 1$ and $d_{3n,t+1} = 1$ in periods t and $t + 1$. $H^{(s)}_{2nt}$ only differs from $H^{(s)}_{1nt}$ by setting $d_{2nt} = 1$ (and $h_{nt} = h_{2nt}$). Both $H^{(s)}_{3nt}$ and $H^{(s)}_{4nt}$ set $d_{1n,t+s} = 1$ for all $s > t$; $H^{(s)}_{3nt}$ sets $d_{3nt} = 1$ while $H^{(s)}_{4nt}$ sets $d_{4n,t+1} = 1$. By construction it follws that for all $k \in \{1, 2, 3, 4\}$:

$$H^{(\rho+2)}_{knt} = (z'_{n,t+\rho+2}, 0, ..., 0; 0, ..., 0)' \equiv H_n$$

showing that it is feasible to reach a point $\rho + 2$ periods hence, where differences in two choices in periods t and $t + 1$ (equalizing family size) followed by a sequence of the same choice (not to work) obliterate any future consequences of choices prior to period t.

To show how finite dependence is exploited in the representation of the conditional value function and hence in estimation, define $l^{(s)}_{knt} \equiv l_{n,t-s}$ for all $s \in \{-1, \ldots, -\rho\}$, and for $s \in \{0, 1, \ldots, \rho + 1\}$ let $l^{(s)}_{knt}$ as the amount of leisure consumed in period $t + s$ when this choice path indicated by the state variables $H^{(s)}_{knt}$ is followed. For example:

$$l^{(s)}_{ntk} = \left(1 - \sum_{k=1}^{t} \rho_k b_{n,t-k} - \rho_k\right)$$

The leisure component of utility acruing over the periods $t+1$ though S from setting $d_{n,t+j,1} = 1$ each period $t + s$ is thus:

$$\eta_n \sum_{s=1}^{\rho-1} \beta^s \left[z'_{n,t+j} B_1 l^{(s)}_{ntj} + \sum_{r=1}^{\rho} \delta_s l^{(s)}_{ntj} l^{(s-r)}_{ntj} \right]$$

The following Lemma now provides a characterization of the conditional valuation functions.

Lemma 4.1. *Define for $k \in \{1, \ldots, 4\}$:*

$$W_k\left(H_{nt}\right)=\sum_{s=1}^{T-t}\beta^{s-t}\left\{z'_{n,t+s}B_1 l^{(s)}_{ntk}+\sum_{r=1}^{\rho}\delta_s l^{(s)}_{ntj} l^{(s-r)}_{ntj}\right\}-\sigma\sum_{s=1}^{\overline{\rho}}\beta^{s-t}\ln p_1\left(H^{(s)}_{knt}\right)+\beta^{\rho+2-t}V\left(H_n\right)$$

Then:

$$V_k\left(H_{nt}\right)=\begin{cases} u_k\left(H_{nt}\right)+W_k\left(H_{nt}\right) & \text{for } k\in\{3,4\} \\ u_k\left(H_{nt}\right)-z'_{n,t+1}B_0\pi+W_k\left(H_{nt}\right) & \text{for } k\in\{1,2\}\end{cases}$$

and for $j \in \{1, \ldots, 4\}$:

$$V_j\left(H_{nt}\right)-V_k\left(H_{nt}\right)=\sigma\ln p_j\left(H_{nt}\right)-\sigma\ln p_k\left(H_{nt}\right) \tag{4.5}$$

The log odds ratio scales the difference in conditional value functions by a variance parameter because of wages. Similarly the correction factor on the the choice probabilities that offset the difference between the conditional valuation functions and the (unconditional) value function. Equating the right side of both. Differentiating

5. Results

This section reports on the results of the structural estimation. Tables II through VII contain estimates of estimates of the parameters determining wages, consumption preferences, the participation cost, the child nurturing time, plus the utility from leisure and offspring.

5.1. Wages

Our estimates of the wage equation, displayed in Table III, are comparable to those reported in Miller and Sanders (1997) for the National Longitudinal Survey for Youth (NLSY), Altug and Miller (1998) also using the PSID, and others. All the coefficients are significant. Working an extra hour increases the wage rate up to four years hence, although in diminishing amounts. The effect is nonlinear, and this is captured by the participation variables. Age has a quadratic effect, eventually leading to declining productivity, and additional education mitigates the onset of the decline. We note that the linear terms on age are not identified.

The estimate quantitative magnitudes of past experience are also plausible. Recent working experience is more valuable than more distant experience: at 2000 hours per year, the wage elasticity of hours lagged once is about 0.18, but the wage elasticity of hours lagged twice is only 0.03. Also the further back the work experience is, the less the timing matters; an extra hour worked one year in the past has about twice the effect on current wages as an extra hour worked two years in the past, but the difference between the wage effects of an extra hour worked three and four years in the past, respectively, is less than 40%.

Another measure of the effect of past labor supply on wages: consider the total change in wages for a woman who has not worked up to date $t-\rho$ and then works the sample average of hours for those women who work, denoted $\overline{h}_t$. Then this measure is given by $\sum_{s=1}^{4}[\delta_{1s}\overline{h}_{t-s}+\delta_{2s}]=0.12$. Much of this long-term effect is due to hours worked in the past year. Specifically, the growth in wages between $t-1$ and t for a woman who does not participate from $t-\rho$ to $t-2$, but works the sample average at $t-1$ is $\delta_{11}\overline{h}_{,t-s}+\delta_{21}=0.08$.

On the other hand, women who worked less than 1000 hours the previous year do not receive this increase in wages, this may be capturing the effect of discouragement normally found in the standard job search model. It should be noted that we do not explicitly model this type of search cost in our model, however, we can pick up the lower bound of this effect. This means that not everybody gets the benefit from past job experience, there is a threshold number of hours of about 1500 for this positive effect to kick in. This will impact fertility behavior even more than if there were positive benefit from all levels of past hours, since a mother could reduce her hours and still continue to enjoy the benefit of higher future wages. We will come back to this point in the empirical findings section when we will have estimates of the fraction of time a mother spends nurturing her new born.

The estimated change in aggregate wages over our sample period is displayed in Figure I, along with its 99% confidence interval. The most striking feature of that plot is that although the magnitude of the changes fluctuate over the sample period, the signs are always positive. This shows that over time the aggregate females wage has been increasing. This is not a surprising finding, given the fact the wage gap between males and females having been closing over time. However it does raise an interesting issue as to whether the attachment of females to the labor force, in term of their persistence in labor participation, is having an aggregate effect. For example, suppose by more females working more hours and participating on a more consistent level equivalent to men, then the employers in the aggregate are willing to pay females higher wages closer to males. This higher wages, some would argue, would then cause females to work more and have less children. Our approach can also disentangled such a result by controlling for aggregate shock, and then seeing the relative importance of the wage effect.

5.2. Preferences over Consumption and Wealth Effects

The estimates of the consumption equation are based on the main sample of females for the years 1968 to 1992. Consumption for a given year in our study is measured by taking 0.25 of the value of the different components for year $t-1$ and 0.75 of it for year t. This is explained in more detail in the data appendix. The elements of z_{nt} used in this stage of the estimation are defined as FAM_{nt}, $YKID_{nt}$, $OKID_{nt}$, AGE_{nt}^2, NC_{nt} and SO_{nt}. The estimates in Table 4 show that consumption increases with family size and children consume less than adults, since the coefficients on children between the ages of zero and fourteen are negative and smaller in absolute magnitude than the coefficient on total household size. Furthermore, the behavior of consumption over the life-cycle is concave since the coefficient on age squared is negative. All the other coefficients are significant. The agregate shocks components are estimated very precisely. In fact, there is also significant variation over time as the test statistic for the null hypothesis that $(1-\alpha)^{-1}\Delta\ln(\lambda_t) = (1-\alpha)^{-1}\Delta\ln(\lambda_{t-1})$ for $t = 1969, ..., 1992$ is 395. Under the null hypothesis, it would be distributed as a χ^2 with 23 degrees of freedom, implying rejection of the null at 99% significance levels.

5.3. Fixed Cost of Participation

Table VI contains estimates of the fixed cost of participation. First the constant term is negative, which means that participation in the labor force has a fixed utility cost instead of a benefit, which is what standard economic theory would predict. Age reduces this cost of participation in the labor, but this reduction is at a decreasing rate as the parameter estimate on the AGE^2 is negative. Education decreases the cost associated with age. There is a positive sign on the estimates of AGE$\times$EDUC which implies that a more educated female has a lower cost of participation for a given age than a less educated female. To understand the overall effect of age and education on the fixed cost of participation, we investigate what is the shape of this function conditional on education. Married women have a lower cost of participation while blacks have a higher cost of participation for a given age and education level. Again these results are not surprising since the standard literature has documented similar results(see for example Altug and Miller (1998))..

5.4. Nurturing Cost

Table VII contains the results from the estimation of the fraction of time spent nurturing a child. These estimates seems quite small and the only significant cost is that of older children. These are similar results to those found by Hotz and Miller (1984) which found that these parameters follow off a geometric rate. This is very important in our model, since with the nonlinearity observed in the estimates of the wage equation, this implies that if a female reduces her time in the labor force to have a child, then they would not benefit from the increases in wages as a result of human capital accumulation in terms of their previous labor supply. So holding all other things constant, this would make having children less desirable for a female who is on a high wage trajectory. This combined with the estimates of the risk aversion parameter means that females would like to smooth more there consumption, hence working more in earlier years and delaying child-bearing to later years. This would mean that working females would have less children than nonworking female.

5.5. Utility Cost of Leisure

Table VIII contains the estimates for the utility cost of leisure. Leisure have the expected sign, the direct effect of age and age square are both insignificant. However, for a given age education increases the value of leisure, this effect is working in the opposite direction from these for participation. This mean that for given age a female with higher education participate more but conditional on participation the would take more leisure. Our estimates suggest that leisure is intertemporally nonseparable. Past leisure are substitute with for current leisure These is opposite to what is found in Altug and Miller (1998), among other, about the separability of leisure when one does not control for children. Another, surprising results we found is that the sign on marriage in our results is negative. At first glance, this would imply that married females love leisure less. One explanation for this effect could be simple the fact that married females are working more than before and is still having children. Since we do not allow at the moment for the utility of birth or the time cost of raring a child to depend on such demographics, as marital status, then the only way they found then be having children and still working is if they as a group love leisure less. Another explanation may be due to the welfare system. In the era of our sample, a subsistence income (AFDC) is available to unmarried mothers, but (basically) only conditional on them not working. Married females do not face a similar trade-off. Since welfare participation among female heads is quite common in this era (roughly around one-third), this is definitely an important enough phenomenon to account for this results. In short, the "leisure" time of female heads is highly subsidized, and they may well have similar preferences as wives.[5] This is some thing that we will explore further.

5.6. Birth Effects

We concurred with the classical literature that children are good and not bad, since we find a positive net utility all birth except for the child. It should be noted that this is the next benefit all this suggest is that women prefer two children to one. The parameter on the timing of births for example, would imply that the optimal space of a two-child family would be 3 to 4 years apart. So, having children too close or too far apart is less desirable. Turning to the cost of a child, we find that both sets of estimates give similar results. We find that having at least a high school education significantly increases that cost. After controlling for education, we find that Blacks have a significantly lower cost than White. The fact that education significantly increases the cost of having a birth coincides with our earlier hypothesis, and can help explain the unanimous empirical finding that number of children is negatively related to level of education.

6. Policy simulations

There are many ways in which public policy over the last century has affected the costs and benefits of having children. From child labor laws to the public provision of schooling, from the subsidizing of health care to local taxes that support amenities such as swimming pools, as well as sporting and other events for children, raising children depends on social infrastructure that is often taken for granted in modern developed societies. Over the last several decades, greater attention has been paid to jointly determining fertility and female labor supply. Part of the concern about the falling rates of fertility are related to the long-term viability of the social security system in many developed countries, especially in Western Europe.

This section considers a variety of policies that subsidize fertility to investigate how responsive women are to changes in the incentives they factor in between market work and raising a family. Our study shows that different policies not only have different aggregate or average effects on fertility and female labor supply, but also have very significant compositional effects, or incidence across this heterogeneous population. We hasten to add that our contribution is positive, not normative, seeking to provide quantitative analysis against which different policy options can be evaluated.

6.1. Overview of the simulations

We substituted the parameters obtained from our estimation procedures into the utility function, the equation characterizing the returns to experience, and the child care cost equation and solved the decision-maker's problem. We conducted simulations for a wide

[5] We would like to thank Elizabeth Powers for pointing out this very insightful possibility to us.

range of female types in the population, but they are not exhaustive. We stratified the population, breaking down the groups according to a three-way classification scheme, by race, marriage and education, and considered an individual whose unobserved fixed effects correspond to the estimated means of the distributions. Three racial types were considered, namely Black, White and Hispanic (respectively abbreviated B, H and M in Tables IX and X below). Marriage was a dichotomous variable partitioning women by marital status at age 25, where M denotes she was married at age 25 or before, and U if not. We considered three educational groups, those who completed some years at college (denoted by the inequality sign >), those who completed some years at high school but not college (denoted by HS), and those with less education than that (denoted by a < sign). Thus our simulations apply to women in the 18 categories whose marginal utilities' of wealth, and whose endowed marginal product of labor (controlling for schooling and experience), correspond to the estimated sample means.

The models we simulated are slightly less complex than the estimation framework itself in three ways. The first simplification was to limit the choice set. Rather than assuming that workers made a discrete choice about whether to participate in the labor force or not with a continuous hours choice, we discretized the labor supply choice set facing workers, limiting them to 10 equally spaced choices in the $[0, 1]$ interval. Second, we linearized the value of marginal consumption around the marginal utility of consumption achieved in the current regime. Thus in the objective function (2.9), U_{3ntk} is replaced with $\widetilde{U}_{3ntk} \equiv \eta_n^{-1}$. Third, we investigated an economy where there are no aggregate shocks. As a practical matter, the quantitative significance of aggregate demographic shocks (such as the baby boom in the U.S., the AIDS crisis in Botswana and other countries, the effects on fertility of immigration both legal and illegal into U.S. and parts of Western Europe) is difficult to overstate, and we think that excluding them is the main reason why our results should be treated cautiously.

The model was solved for each group under five policy regimes. The benchmark regime, labelled Estimation, is the current one, which may be compared with the conditional sample means from the data set. In the first two alternative regimes we analyze the subsidy to having children does not vary with the recipient, although the value a mother places on the scheme depends on her wealth and wage rate. In the regime labelled Expenses, the state pays all the estimated monetary costs associated with raising children, removing the wedge in the marginal utility of wealth between households that have children and those that do not. Under the Day-care policy, maternal time is replaced with publicly funded child care centers. In the other two regimes the payment mothers receive depends on her wages and hours she worked before taking time off to have a child. The Wages policy would pay the mother the wages she would have received if she had decided against having her child. If the Retraining policy is adopted, mothers are given retraining upon reentering the workforce that fully restore the human capital from lost workforce experience.

In our model there are three costs associated with child-care: the lifetime discounted cost of market inputs used up raising a child, the direct time cost in terms of the required for nurturing, and the human capital accumulation cost stemming from the experience acquired from working that is not used when women quit the labor force to have children.

6.2. Solving the Model

The Type 1 extreme value also implies that for each $j \in \{1, \ldots, 4\}$

$$V_j\left(z_{nt}^*\right) = u_{ntj}^{(b)} + u_{ntj}^{(l)}\left(h_{nt}^{(j)}\right) - \lambda\left(\tau_{nt}\right)\left(\pi b_{ntj} - w_{nt} h_{nt}^{(j)}\right) + \beta \log\left\{\sum_{k=1}^{4} \exp V_k\left(z_{n,t+1}^{(j)}\right) + \gamma\right\}$$

Upon defining p_{knt} as the conditional choice rate in period t, we obtain the probability of making choice k by the n^{th} female in period t as:We also use the fact that an interior solution for those participating in the labor force requires $\partial V_{1nt} \backslash \partial h_{nt} = 0$ or $\partial V_{3nt} \backslash \partial h_{nt} = 0$. Differentiating with respect to hours we have:Thus if $\mathbb{I}_{knt}^{o} = 1$ for $k = \{1, 3\}$, then h_{nt}^{o} solves:

$$\lambda\left(\tau_{nt}\right) w_{nt} - z_{nt}' B_1\left(1 - c_{nt}^{(j)} - h_{nt}^{(j)}\right) - \sum_{s=0}^{\rho} \delta_s l_{nt-s} = \sum_{k=3}^{4} p_k\left(z_{nt}^*\right) \frac{\partial V_k\left(z_{n,t+1}^{(j)}\right)}{\partial h_{nt}^{(j)}} \tag{6.1}$$

The left side of Equation (6.1) gives the current benefits and costs of spending a marginal hour working, comprising a utility cost in terms of leisure foregone, and the value of the extra goods and services produced. The right side shows the expected future benefits. Marginally adjusting current hours worked directly affects future productivity as well as the benefits of future leisure. Moreover, supposing the probability of working next period increases next period from this adjustment, the net benefits of working next period should be applied to the increase. This is captured in the second expression on the right side of Equation (6.1).

We first simulated the prediction of the model for females in each of the categories described above over the 25 years of a partial life cycle starting at age 20, for use as a bench mark case. This requires us to solve 18 valuation functions for the optimization problem each type solved, obtain the optimal decision rules, and thus compute the probabilities of observing any given decision, as a mapping of the state variables, which in this case are the vector of lagged labor supplies and a vector for the ages of the offspring. An appendix describes the algorithm in detail. Briefly, we combined the use of both policy function iteration (using Newton steps) with value function iteration (using the contraction operator on the value function). Convergence to the solution of the infinite horizon problem occurred relatively quickly, typically within seven iterations.

The labor force participation rate and expected fertility rate over this period (essentially the TFR) for each type is reported in the second column of Tables IX and X under the heading of Estimation. A sense of how representative our groups are is found by comparing the simulated results for our estimated model with their corresponding sample means in the first column, headed Actual. Note that the numbers are not very close, although many of the inequalities within each column are preserved. This is attributable to two factors. The first is estimation error. The second is that the sample means do not condition on the values of the unobservables, which enter in a highly nonlinear way into the participation and fertility choices. To separate out these separate influences, we will nonparametrically estimate the same set of statistics for that person in the group with the estimated mean fixed effects, which simply weights the data used to obtain the averages in the first column by how close each observation is to the mean estimated fixed effect vector.

Table IX shows most of the types have fertility rates below the replacement rate of 2. For example, the TFR of all the college educated groups are all below the replacement rate.

College educated white females bear the least number of children (1.1 for the group as a whole and 1.2 at the mean fixed effects), and black married females with less than high school education the most (2.1 for the overall group and 2.4 at the mean fixed effects).

In most, but not all groups, those married by 25 bear more children than those who had not married by then. Table X shows that, with the notable exception of college educated whites, unmarried women are more likely to participate in the labor force. At 0.93, the labor force participation rate for a married college educated white female with the mean fixed effects exceeds all other groups, closely followed by unmarried college educated black women (at 0.91). Across education achievement and marital status but within race categories, blacks exhibit the biggest range in labor force participation rates. The exact derivation is presented in more details in Appendix 1.

6.3. Child-care Support

There are many ways to subsidize fertility by having the state pay for the discounted lifetime cost of children. For example, it could be achieved though tax credits at upper income levels and child support payments for those who do not receive enough taxable income. In this framework this is equivalent to imposing the constraints $\pi_0 = 0$ and $\pi_1 = 0$ in the expression for child care costs:

$$\pi(z_{nt}) = \pi_0 + z'_{nt}\pi_1$$

The total fertility and labor force participation rates that are induced by this subsidy are shown in the third columns of Tables IX and X. Paying the market goods inputs for raising children has a substitution and wealth effects. In a static model, the substitution effect induces women to have more children and reduce their own consumption of leisure and other goods, while the wealth effect induces them to increase their consumption of leisure and children. The results of the dynamic simulations lend support to this intuition. In 16 of the 18 groups labor force participation declines, and in all but one instance fertility rises, 6 groups (compared to 4) now settling above the replacement rate. The 3 types whose fertility behavior is most sensitive to this policy shift are the married non-college educated black female and the unmarried lowest educated black female. By way of contrast the biggest reduction in labor force participation rate is amongst unmarried high school educated whites.

6.4. Day-care

Rather than pay for market inputs directly, another public policy for subsidizing fertility is to expand the availability of child care services for the mothers of infants and preschool age children, by financially supporting centers, or reimbursing mothers who place their children in them. In our framework a policy that eliminates the maternal time inputs altogether would set $\rho_i = 0$ for $i \in \{1, \ldots, 5\}$. This increases the amount of time mothers of young children have for leisure and work. In a static model of fertility and labor supply, fertility increase in response to a reduction in one of its factor inputs, maternal time. Furthermore, the time freed up from looking after children is distributed between extra leisure, and working for more goods and services over and above those used up by the additional children. Consequently, one predicts that both fertility and labor supply would increase, the latter less than the amount of time released from child care.

The fourth column shows the labor force participation and fertility outcomes from solving the optimization problem under the Day-care policy. As expected all the group exhibit higher fertility rates, 12 now at or above the replacement rate of 2.0, with married high school educated white females registering the biggest increase (from a TFR of 1.52 to 2.30). Comparing the effects on TFR across different groups, we see that switching from subsidizing market inputs to replacing maternal time inputs has a far greater impact on females with some college education than those who did not complete high school. Indeed in just one group, married blacks who did not complete high school, TFR would actually fall from 2.63 to 2.41 if subsidizing market inputs were replaced with subsidizing maternal inputs. This finding demonstrates that the type of subsidy to child care helps determine not just the aggregate level of births, but also their composition within different types of households.

The change in labor force participation rates are more ambiguous, in fact puzzling. Since returns from experience on the job is likely to strengthen attachment to the labor force beyond that predicted by the static model, we are further investigating this counter-intuitive result.

6.5. Paid Maternity Leave

Paying females wages when they take maternity leave is a third way of promoting higher fertility. A distinguishing feature of this policy is that women with high wages receive greater payment than those receiving lower wages. (Note that if the payment is a fixed allowance, then the analysis of Expenses policy applies.) In contrast to the two previous schemes, (each of which has only one degree of freedom, the proportion of costs or time covered), this scheme has two, what percentage of her market wage a mother is paid while on maternity leave, and the maximum eligibility period per child. Under the Wages policy, mothers are paid the wage they would have received if they had not given birth, and the maximum eligibility period is the amount of time they would have withdrawn from the workforce in the absence of the subsidy. These variables are for the most part negatively correlated, and therefore affect the total payment in offsetting directions.

In particular, suppose the woman gives birth at period t, let $h^o_{n,t+s}(b_{nt}=0)$ denote the woman's labor supply s periods after the birth had she not left the workforce to give birth, let $w^o_{n,t+s}(b_{nt}=0)$ denote her wage rate had she not given birth, and let τ^0_n denote the number of periods she would have taken off if there were no provisions for paid maternity leave. Then in this policy regime the wage payment she receives upon having a child is:

$$\sum\nolimits_{s=0}^{\tau_n} \lambda_{n,t+s} w^o_{n,t+s}(b_{nt}=0)\, h^o_{n,t+s}(b_{nt}=0)$$

In a static framework, paid maternity leave induces women to reduce their labor supply and have larger families. In our dynamic framework paying wages does not fully compensate a mother for taking maternity leave, because job market experience acquired before giving birth depreciates over the time spent out of the labor force. Consequently, females who decide to have a child because of the paid maternity leave may simply exit the labor force permanently if their market capital has depleted sufficiently quickly. This scenario certainly arises when, in the absence of the paid leave policy, women essentially choose between having a career and having a family.

Our preliminary simulation results are displayed in the fifth columns of Tables IX and X. They show that in 13 out of the 18 cases the labor supply participation falls, because of the substitution effect into child rearing activities, and the compounding effect of human capital depletion. Although total fertility rates increase in all categories, this policy is not as effective as directly paying for the time inputs; in every category fertility rates under subsidized Daycare exceed those in attained when there is paid maternity leave as mandated in Wages.

6.6. Retraining

In our framework mothers lose human capital from temporarily withdrawing from the labor force. The last counter factual regime we consider does not make any payments to mothers, but offers partial compensation by putting women returning to work from maternity leave on an equal footing with those who chose not to have children. The policy scheme simulated in Retraining restores them to the wage trajectory they would have been on if they not withdrawn from the workforce to have children. In our framework the labor force experience over the previous ν periods helps determine the current wage. Thus, if the female in Model 4 reenters τ_n^4 periods after she has her birth, the natural logarithm of her wages increases by:

$$\sum\nolimits_{s=0}^{\min\left\{\nu,\tau_n^4\right\}} \left[\delta_{1s} h_{n,t-s}^o \left(b_{nt}=0\right) + \delta_{2s} d_{n,t-s}^o \left(b_{nt}=0\right)\right]$$

The last columns of Tables IX and X display the results, which in some ways are the most dramatic. The total fertility rate of every group except the unmarried white females with less than high school education rises above the replacement rate, and for one group, married black females with high school education, reaches 3.

7. Conclusion

This paper develops a dynamic model of female labor supply and fertility behavior and estimates its structural parameters. Previous empirical research on female labor supply had shown that current labor supply choices affect future wages and utility through intertemporal nonseparabilites in the production function (such as through learning by doing or staying in practice), and in utility (for example, through the household production function and also possibly due to the intertemporal nature of utility from leisure). In addition, there are a small number of studies of fertility behavior that suggest the timing of later births is partly determined by economic factors. Our study nests both kinds of dynamic interactions within a unified structural model.

Our estimates reaffirm the importance of dynamic factors in labor supply and fertility choices. Wages increase with experience up to four years in the past, recent experience counting the most. Leisure taken in different periods are substitutes. Estimated preferences peg optimal birth gestation at about two years.

From a policymaker's perspective: Restoring human capital from work experience is the biggest factor in raising TFR, even though this does not directly subsidize childbearing and fertility inputs. Paying for daycare, expenses incurred raising children, or women a working wage while they have children, all increase fertility rates. None of the policies have much effect on labor supply, which is largely determined by the human capital considerations.

References

[1] Ahn, Namkee(1995), *Measuring the Value of Children by Sex and Age Using a Dynamic Programming Model*, Review of Economic Studies 62, 361-79

[2] Altug, Sumru and Robert A. Miller (1998), *The Effect of Work Experience on Female Wages and Labour Supply*, Review of Economic Studies, 45-85.

[3] Altug, Sumru and Robert Miller(1990), *Household Choices in Equilibrium*, Econometrica, Vol. 58. no. 3., 543-570.

[4] Angrist, J. and W. N. Evans(1998), *Children and Their Parents' Labour Supply: Evidence from Exogenous Variation in Family Size*, American Economic Review 88, 450-77.

[5] Arroyo, Cristino R. and Junsen Zhang (1997), *Dynamic Microeconomics Models of Fertility Choice: A Survey.*, Journal of Population Economics, Vol.10, 23-65.

[6] Becker, G. (1965), *A Theory of Allocation of Time, Economic, Journal, 75, 496-517.*

[7] Becker, Gary S., Kevin M. Murphy and Robert Tamura (1990), *Human Capital, Fertility, and Economic Growth*, Journal of Political Economy, Vol. 98, no.2, S12-S37.

[8] Del Boca, Daneila (2002), *The Effect of Child Care and Part-Time Opportunities on Participation and Fertility Decisions in Italy,* Journal of Population Economics, Vol.

[9] Butz, William P. and Michael P. Ward (1980), *Completed Fertility and its Timing,* Journal of Political Economy, Vol.88, no. 5, 917-940.

[10] Butz, William P. and Michael P. Ward (1979), *The Emergence of Countercyclical U.S. Fertility*, American Economic Review, Vol. 69, no. 3 , 318-328.

[11] Card D. (1990), *Labour Supply With a Minimum Threshold*, Carnegie-Rochester Series on Public Policy, 33, 137-168.

[12] Eckstein anfd K. Wolpin (1989), *Dynamic labour Force Participation of married women and Endogenous Work Experience*, Review of Economic Studies, Vol. 56, 1989, 375-590.

[13] Ejrnaes, Mette and Astrid Kunze (2002), *Wage Dips and Drops around First Birth*, Working Paper, University of Copenhagen.

[14] Francesconi, M. (2002), *A Joint Dynamic Model of Fertility and Work of Married Women*, Journal of Labor Economics, 20, 336-380

[15] Heckman, James J. , V. Joseph Hotz and James R. Walker, *New Evidence on the Timing and Spacing of Births*, American Economic Review, Vol. 75, no. 2 (1985) 179-184.

[16] Heckman, James J. and James R. Walker (1990), *The relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data*, Econometrica, Vol. 58, no. 6, 1411-1441.

[17] Hotz, V. Joseph, Jacob Alex Klerman and Robert J. Willis (1997), *The Economics of Fertility in Developed Countries, in Handbook of Population and family Economics*, ed. by M. Rosenzweig and O Stark. North Holland .

[18] Hotz, V. Joseph and Robert A. Miller (1993), *Conditional Choice Probabilities and the Estimation of Dynamic Models of Discrete Choice*, Review of Economic Studies, 60, 497-429.

[19] Hotz, V. Joseph and Robert A. Miller (1988), *An Empirical Analysis of Life Cycle Fertility and Female Labour Supply*, Econmetrica, Vol. 56. no. 1, 91-118.

[20] Hotz, V. Joseph, Robert A. Miller , S. Sanders and J Smith (1994), *A Simulation Estimator for Dynamic Models of Discrete Choice*, Review of Economic Studies, 61 , 265-289.

[21] Krammer, Walter and Klaus Newusser (1984), *The Emergence of a Countercyclical U.S. Fertility: Note,* American Economic Review, Vol. 74, no. 1, 201-202.

[22] Lewis, Frank D. (1983), *Fertility and Saving in the United States: 1830-1900*, Journal of Political Economy, Vol. 91, no. 5, 825-840.

[23] Lundlolm and Ohlsson(1998), *Wages, Taxes and Publicly Day Care,* Journal of Population Economics, Vol. no.

[24] Mace, B. (1991), *Full Insurance in the Presence of Aggregate Uncertainty*, Journal of Political Economy, vol. 89, 1059-1085.

[25] Macunovich, Diane J. (1998), *Fertility and the Easterlin Hypothesis: An assessment of the literature*, Journal of Population Economics, Vol. 11, 53-111.

[26] Merrigan, Philip and Yvan St.-Pierre (1998), *An Econometric and Neoclassical Analysis of the Timing and Spacing of Births in Canada from 1950 to 1990*, Journal of Population Economics, Vol. 11 , 29-51.

[27] Miller , Robert A. (1997), *Estimating Models of Dynamic Optimization with Microeconomics Data*, in Pesaran M. and Schmidt, P. (eds) Handbook of Applied Econometrics 2. Microeconometrics (London: Basil Blackwell), 247-299.

[28] Miller, Robert A. (1984), *Job Matching and occupational Choice*, Journal of Political Economy, 92, , 1086-1120.

[29] Miller. Robert A. and Sanders, S. (1997), *Human Capital development and Welfare Participation*, Carnegie-Rochester Conference Series on Public Policy, 46, 237-253.

[30] Newman, John L. (1983), *Economic Analyses of the Spacing of Births*, American Economic Review, Vol. 73, no. 2, 33-37.

[31] Olsen, Randall J. (1983), *Mortality Rates, Mortality Events, and the Number of Births*, American Economic review, Vol. 73, no. 2, 29-32.

[32] Powell ,James L., James H. Stock, Thomas M. Stoker(1989),*Semiparametric Estimation of Index Coefficients,* Econometrica, Vol. 57, No. 6., pp. 1403.

[33] Schultz, Paul T. (1994), *Human Capital, Family Planning, and Their Effects on Population Growth*, American Economic Review, Vol. 84, no. 2, 255-260.

[34] Van Der Klaauw, Wilbert (1996), *Female Labour Supply and Marital Status Decisions: A :Life Cycle Model.*, Review of Economic Studies, Vol. 63, no. 2, 199-235.

[35] Waldfogel, Higuchi and Abe (1999), *Family Leave Policies and Women's Retention After Childbirth: Evidence from the United States, Britain and Japan,* Journal of Population Economics, Vol.

[36] Walker, J. (1996), *Parental Benefits, Employment and Fertility Dynamics,* Research in Population Economics, Vol. 8

[37] Willis, Robert J. (1973), *A New Approach to the Economic Theory of Fertility Behavior*, Journal of Political Economy, Vol. 81, no. 2, S14-S64.

[38] Wolpin, Kenneth I. (1984), *An Estimable Dynamic Stochastic Model of fertility and Child Mortality*, Journal of Political Economy, Vol. 92, no. 5, 852-874

8. Appendix 1

In this appendix we defines a class of conditional choice probability (CCP) estimators, to which the estimator used in Section 8.3 belongs, and show consistency and asymptotic normality of these estimators.

8.0.1. A class of CCP estimators.

The estimators (Θ_3^N, Γ^N) defined by equation (**??**) and (**??**) are examples of CCP estimators, in which the individual-specific effects $\mu_n^N \eta_n^N$, time-specific effects $\omega_t^N \lambda_t^N$ and the conditional choice probabilities p_{knt}^N for $k = 0, .., 3$ and $p_{0nt}^{(s,k,N)}$ for $s = 1, ..., \overline{\rho}$ enter as incidental parameters. This estimator falls within a class of CCP estimators that can be described as follows.

Let $D_n(\Theta, \mu_n, p_n)$ be a $q \times 1$ vector function such that $\Theta_0 \equiv (\Theta_{30}, \Gamma_0)$ is the unique root of $E[D_n(\Theta, \mu_n, p_n)]$. For each, $n \in \{1, 2, ...\}$ and $\Theta \in \Xi$, let μ_n^N be a kernel or traditional estimator which converges uniformly to μ_n, let $p_n^N(\mu_n)$ be a kernel estimator which converges uniformly to $p_n(\mu_n)$. We define Θ^N as any solution to

$$\frac{1}{N} \sum_{n=1}^{N} D_n(\Theta^N, \mu_n^N, p_n^N(\mu_n^N)) = 0 \tag{8.1}$$

The proof of proposition 1 below shows that Θ^N is asymptotically normal, but is not centered on zero. While an asymptotically unbiased estimator could be calculated following the procedure in Hotz and Miller (1993) by forming a linear combination of the estimators which are based on different bandwidths for the incidental parameters, the limited empirical evidence suggests that the asymptotic bias is unimportant.[6]

Proposition 1: Θ^N converges to Θ_0 and $\sqrt{N}(\Theta^N - \Theta_0)$ is asymptotic normal with mean $-E(v_n)/2$ and covariance matrix $(D_0')^{-1} S_0 D_0$, where v_n D_0 and S_0 are defined by equations (8.10), (8.17) and (8.18).

Proof.

For ease of notation, we assume that μ_n^N and $p_n^N(\mu_n^N)$ take the form of nonparametric kernel estimators weighted or unweighted probability density functions of the form

$$\mu_n^N = \sum\nolimits_{m=1, m \neq n}^{N} \phi_m \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] \tag{8.2}$$

and

$$p_n^N(\mu_n^N) = \sum\nolimits_{m=1, m \neq n}^{N} d_m \delta^{-q} J[\delta_N^{-1}(k(z_m, \mu_m^N) - k(z_n, \mu_n^N))] \tag{8.3}$$

where $k(z_m, \mu_m^N)$ is mapping that defines the distance between the observations. The proof that Θ^N converges in probability to Θ_0 is standard, relying on the uniform convergence of the incidental parameters to their true values, so that the approximating sample moments obtained by substituting the incidental parameter estimates for their respective true values only affect the resulting structural parameter estimates by an $o_p(1)$ term.[7]

[6] For evidence on the magnitude of this asymptotic bias, see the Monte Carlo simulations in Powell, Stock and Stoker (1989) and the fertility application in Hotz and Miller (1993).

[7] See Hotz and Miller(1993) for a consistency proof of a very similar semiparametric estimator.

To establish the mean, covariance, and bias, we first consider an other estimator denoted by $\widetilde{\Theta}^N$, and show that this has the same asymptotic distributional properties as Θ^N. For ease of notation, let $D_n \equiv D_n(\Theta_0, \mu_n, p_n)$, $p_n \equiv p_n(\mu_n)$

$$D_{0n} \equiv \left[\frac{\partial D_n(\Theta_0, \mu_n, p_n)}{\partial \Theta}\right]$$

$$D_{1n} \equiv \left[\frac{\partial D_n(\Theta_0, \mu_n, p_n)}{\partial \mu_n} + \frac{\partial D_n(\Theta_0, \mu_n, p_n)}{\partial p_n} . \frac{p_n(\mu_n)}{\partial \mu_n}\right]$$

and

$$D_{2n} \equiv \left[\frac{\partial D_n(\Theta_0, \mu_n, p_n)}{\partial p_n}\right]$$

The estimator $\widetilde{\Theta}^N$ satisfies the equation

$$-N^{-1}\sum\nolimits_{n=1}^{N}[D_n + D_{0n}(\widetilde{\Theta}^N - \Theta_0)] \tag{8.4}$$

$$= N^{-1}\sum\nolimits_{n=1}^{N}[D_{1n}(\mu_n^N - \mu_n) + D_{2n}(p_n^N(\mu_n^N) - p_n(\mu_n))] \tag{8.5}$$

Define the quantities

$$v_{1mn}^N \equiv D_{1n}[\phi_m \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] - \mu_n] + D_{1m}[\phi_n \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] - \mu_m] \tag{8.6}$$

$$v_{2mn}^N \equiv D_{2n}[d_m \delta^{-q} J[\delta_N^{-1}(k(z_m, \mu_m^N) - k(z_n, \mu_n^N))] - p_n] \tag{8.7}$$

$$+D_{2m}[d_n \delta^{-q} J[\delta_N^{-1}(k(z_m, \mu_m^N) - k(z_n, \mu_n^N))] - p_m] \tag{8.8}$$

$$v_{mn}^N \equiv v_{1mn}^N + v_{2mn}^N \tag{8.9}$$

$$v_n = f(x_n)[D_{1n}(\mu_n + \phi_n) + D_{2n}(p_n + d_n)] - D_{1n}\mu_n - D_{2n}p_n \tag{8.10}$$

where $f(x_n)$ is the density of x_n.

Expanding the first expression on the right-side of 8.4 using the definition of the non-parametric estimator for μ_n yields

$$N^{-1}\sum\nolimits_{n=1}^{N}[D_{1n}(\mu_n^N - \mu_n)$$

$$= N^{-1}\sum\nolimits_{n=1}^{N} D_{1n}[\sum\nolimits_{m=1,m\neq n}^{N} \phi_m \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] - \mu_n] \tag{8.11}$$

$$= N^{-1}\sum\nolimits_{n=1}^{N}\sum\nolimits_{m=1,m\neq n}^{N} D_{1n}[\phi_m \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] - \mu_n]$$

$$= N^{-1}(N-1)^{-1}\sum\nolimits_{n=1}^{N-1}\sum\nolimits_{m=n+1}^{N} v_{1mn}^N \tag{8.12}$$

Similarly, the second expression on the right side of 8.4 may be written as

$$N^{-1}\sum\nolimits_{n=1}^{N} D_{2n}[p_n^N(\mu_n^N) - p_n(\mu_n)] = N^{-1}(N-1)^{-1}\sum\nolimits_{n=1}^{N-1}\sum\nolimits_{m=n+1}^{N} v_{2mn}^N \tag{8.13}$$

Following Hotz and Miller (1993), it is straight forward to show that $E[\left\|v_{imn}^N\right\|^2] = o(N)$ for $i = 1, 2$. Then appealing to lemma 3.1 of Powell, Stock and Stoker (1989), p.1410

$$\begin{aligned} N^{-1}(N-1)^{-1} \sum_{n=1}^{N-1} \sum_{m=n+1}^{N} v_{mn}^N &= \frac{E[v_{mn}^N]}{2} + (N-1)^{-1} \sum_{n=1}^{N-1} \{E[v_{mn}^N \mid n] \\ &\quad - E[v_{mn}^N]\} + o_p(1) \end{aligned} \tag{8.14}$$

The right-side of 8.14 depends on N. To derive the asymptotic distribution of $\widetilde{\Theta}^N$. Lemma 1 derives the appropriate limit for the right side of 8.14 as

$$\begin{aligned} & N^{-\frac{1}{2}} \frac{E[v_{mn}^N]}{2} + N^{-\frac{1}{2}} \sum_{n=1}^{N} \{E[v_{mn}^N \quad | \quad n] - E[v_{mn}^N]\} \\ & \qquad = N^{-\frac{1}{2}} \frac{E[v_n]}{2} + N^{-\frac{1}{2}} \sum_{n=1}^{N} \{v_n - E(v_n)\} + o_p(1) \end{aligned} \tag{8.15}$$

The conditions that define $\widetilde{\Theta}^N$ can now be written as

$$\begin{aligned} -N^{-\frac{1}{2}} \sum_{n=1}^{N} [D_n + D_{0n}(\widetilde{\Theta}^N - \Theta_0)] &= N^{-\frac{1}{2}} \frac{E[v_n]}{2} + N^{-\frac{1}{2}} \sum_{n=1}^{N} \{v_n \\ &\quad - E(v_n)\} + o_p(1) \end{aligned} \tag{8.16}$$

The Central Limit Theorem implies that the right-side of 8.16 converges in distribution to a normal random variable with mean $-\frac{E[v_n]}{2}$. Hence, $\sqrt{N}(\widetilde{\Theta}^N - \Theta_0)$ converges to a normal random variable with mean $-\frac{E[v_n]}{2}$ and covariance $(D_0')^{-1} S_0 D_0^{-1}$ where

$$D_0 \equiv E[D_{0n}] \tag{8.17}$$

and

$$S_0 \equiv E[(D_n + v_n - E(v_n))(D_n + v_n - E(v_n))'] \tag{8.18}$$

We complete the proof of this proposition with lemma 2 provided below, which implies that Θ^N and $\widetilde{\Theta}^N$ have the same asymptotic distribution, that is, $\sqrt{N}(\widetilde{\Theta}^N - \Theta^N)$ is $o_{p(1)}$ Q.E.D.

Lemma 1: $N^{-\frac{1}{2}} \frac{E[v_{mn}^N]}{2} + N^{-\frac{1}{2}} \sum_{n=1}^{N} \{E[v_{mn}^N \mid n] - E[v_{mn}^N]\}$
$= N^{-\frac{1}{2}} \frac{E[v_n]}{2} + N^{-\frac{1}{2}} \sum_{n=1}^{N} \{v_n - E(v_n)\} + o_p(1)$

Proof.
Consider v_{1mn}^N which has the form

$$\begin{aligned} v_{1mn}^N \equiv\ & D_{1n}\phi_m \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] \\ & - D_{1n}\mu_n + D_{1m}\phi_n \delta^{-q} J[\delta_N^{-1}(x_m - x_n)] - D_{1m}\mu_m \end{aligned} \tag{8.19}$$

Taking the first on the right-side of 8.19

$$\begin{aligned}
E[D_{1n}\phi_m\delta^{-q}J[\delta^{-1}(x_m-x_n)] \quad | \quad x_n] & \\
&= D_{1n}\int \mu(x)\delta^{-q}J[\delta^{-1}(x-x_n)]f(x)dx \\
&= D_{1n}\int \mu(x_n+\delta u)J(u)f(x_n+\delta u)du \\
&= \int D_{1n}\{\mu(x_n)f(x_n)+\mu(x_n+\delta u)f(x_n+\delta u) \\
&\quad -\mu(x_n)f(x_n)\}J(u)du \\
&= D_{1n}\mu(x_n)f(x_n)+D_{1n}t_n(\delta)
\end{aligned}$$

where $t_n(\delta) \equiv \int[\mu(x_n+\delta u)f(x_n+\delta u)-\mu(x_n)f(x_n)]J(u)du$. Furthermore,

$$\begin{aligned}
E[t(\delta)^2] &= E\left\{\phi_n^2\left[\int[D_1(x_n+\delta u)f(x_n+\delta u)-D_1(x_n)f(x_n)]J(u)du\right]^2\right\} \\
&= E\left\{\phi_n^2\left[\int_{x_n}^{x_n+\delta u}\frac{\partial(D_1 f)(x)}{\partial x}J(u)du\right]^2\right\} \\
&\leq E\left[\phi_n^2\int\delta^2u^2\left\|\frac{\partial(D_1 f)(x)}{\partial x}\right\|J(u)du\right] \\
&= E\left[\phi_n^2\delta^2\left\|\frac{\partial(D_1 f)(x)}{\partial x}\right\|\sigma_u^2\right] \\
&= o_{p(1).}
\end{aligned}$$

Thus, $t_n(\delta)$ has a negligible effect because its variance asymptotes to zero and it has a mean of zero. As a consequence,

$$\begin{aligned}
N^{-\frac{1}{2}}\sum_{n=1}^N\{E[D_{1n}\phi_m\delta^{-q}J[\delta^{-1}(x_m-x_n)] \quad | \quad x_n] & \\
-E[D_{1n}\phi_m\delta^{-q}J[\delta^{-1}(x_m-x_n)]]\} & \\
&= N^{-\frac{1}{2}}\sum_{n=1}^N\{D_{1n}\mu_n f(x_n) \\
&\quad -E[D_{1n}\mu_n f(x_n)]\}+o_{p(1).}
\end{aligned}$$

Similarly, considering the third term in 8.4

$$\begin{aligned}
N^{-\frac{1}{2}}\sum_{n=1}^N\{E[D_{1m}\phi_n\delta^{-q}J[\delta^{-1}(x_m-x_n)] \quad | \quad x_n] & \\
-E[D_{1m}\phi_n\delta^{-q}J[\delta^{-1}(x_m-x_n)]]\} & \\
&= N^{-\frac{1}{2}}\sum_{n=1}^N\{D_{1n}f(x_n)\phi_n \\
&\quad -E[D_{1n}f(x_n)\phi_n]\}+o_{p(1).}
\end{aligned}$$

It now follows that

$$\begin{aligned}
N^{-\frac{1}{2}}\sum_{n=1}^N\{E[v_{1mn}^N \quad | \quad n]-E[v_{1mn}^N]\} & \\
&= N^{-\frac{1}{2}}\sum_{n=1}^N\{D_{1n}f(x_n)(\mu_n+\phi_n)-D_{1n}\mu_n \\
&\quad -E[D_{1n}f(x_n)(\mu_n+\phi_n)+D_{1n}\mu_n]+o_{p(1).} \qquad (8.20)
\end{aligned}$$

By a similar argument

$$
\begin{aligned}
& N^{-\frac{1}{2}} \sum\nolimits_{n=1}^{N} \{E[v_{2mn}^{N} \quad | \quad n] - E[v_{2mn}^{N}]\} \\
& \qquad = \quad N^{-\frac{1}{2}} \sum\nolimits_{n=1}^{N} \{D_{2n} f(x_n)(p_n + d_n) - D_{2n} p_n \\
& \qquad \qquad -E[D_{2n} f(x_n)(p_n + d_n) + D_{2n} p_n] + o_{p(1).} \qquad (8.21)
\end{aligned}
$$

Q.E.D.

Lemma 2: $\sqrt{N}(\widetilde{\Theta}^N - \Theta^N)$ is $o_{p(1)}$.
Proof.
Expanding the right-side of 8.1 about the true structural parameters, Θ_0 and the true incidental parameters, we obtain

$$
\begin{aligned}
& -\frac{1}{N} \sum\nolimits_{n=1}^{N} [D_n + \widetilde{D}_{0n}(\Theta^N - \Theta_0)] \\
= \quad & \frac{1}{N} \sum\nolimits_{n=1}^{N} [\widetilde{D}_{1n}(\mu_n^N - \mu_n) + \widetilde{D}_{2n}(p_n^N(\mu_n^N) - p_n(\mu_n))] \qquad (8.22)
\end{aligned}
$$

where $\sim$ indicates that the appropriate partial derivatives are evaluated at points on the line segment joining (Θ_0, μ_n, p_n) and $(\Theta^N, \mu_n^N, p_n^N)$. Subtracting 8.22 from 8.4 gives

$$
\begin{aligned}
& -\frac{1}{N} \sum\nolimits_{n=1}^{N} [\widetilde{D}_{0n}(\Theta_0 - \Theta^N) - D_{0n}(\Theta_0 - \widetilde{\Theta}^N)] \\
= \quad & \frac{1}{N} \sum\nolimits_{n=1}^{N} [(\widetilde{D}_{1n} - D_{1n})(\mu_n^N - \mu_n) \\
& +(\widetilde{D}_{2n} - D_{2n})(p_n^N(\mu_n^N) - p_n(\mu_n))] \qquad (8.23)
\end{aligned}
$$

consider the following asymptotic expansion

$$
\begin{aligned}
& \frac{1}{N} \sum\nolimits_{n=1}^{N} \{\widetilde{D}_{0n}(\Theta_0 - \Theta^N) - D_{0n}(\Theta_0 - \widetilde{\Theta}^N)\} \\
= \quad & \frac{1}{N} \sum\nolimits_{n=1}^{N} \{D_{0n}(\Theta_0 - \Theta^N) - D_{0n}(\Theta_0 - \widetilde{\Theta}^N) + (\widetilde{D}_{0n} - D_{0n})(\Theta_0 - \Theta^N)\} \\
= \quad & \frac{1}{N} \sum\nolimits_{n=1}^{N} D_{0n}(\widetilde{\Theta}^N - \Theta^N) + o_{p(1)}(\Theta_0 - \Theta^N) \\
= \quad & \{E[D_{0n}] + o_p(1)\}(\widetilde{\Theta}^N - \Theta^N) + o_{p(1)}(\Theta_0 - \Theta^N) \\
= \quad & \{E[D_{0n}] + o_p(1)\}(\widetilde{\Theta}^N - \Theta^N) + o_{p(1)} \qquad (8.24)
\end{aligned}
$$

Considering the second expression in 8.23

$$
\frac{1}{N} \sum\nolimits_{n=1}^{N} [(\widetilde{D}_{1n} - D_{1n})(\mu_n^N - \mu_n) = o_p(1) \frac{1}{N} \sum\nolimits_{n=1}^{N} (\mu_n^N - \mu_n) \qquad (8.25)
$$

where the right of 8.25 follows from the fact that $\widetilde{D}_{1n}$ converges in probability to D_{1n} uniformly in n. Similar U-statistic arguments to that used to justify the asymptotic normality of $\sqrt{N}(\widetilde{\Theta}^N - \Theta_0)$, show that $N^{-\frac{1}{2}} \sum_{n=1}^{N} (\mu_n^N - \mu_n)$ converges in distribution to a normal random variable which is $o_p(1)$. Therefore, 8.25 is $o_p(N^{\frac{1}{2}})$. Finally the third expression in 8.25 can be written as

$$\begin{aligned}
& \frac{1}{N}\sum\nolimits_{n=1}^{N}[(\widetilde{D}_{2n}-D_{2n})(p_n^N(\mu_n^N)-p_n(\mu_n))] \\
= \;& o_p(1)\frac{1}{N}\sum\nolimits_{n=1}^{N}(p_n^N(\mu_n^N)-p_n(\mu_n)) \\
= \;& o_p(1)\frac{1}{N}\sum\nolimits_{n=1}^{N}(p_n^N(\mu_n^N)-p_n(\mu_n^N)) \\
& +o_p(1)\frac{1}{N}\sum\nolimits_{n=1}^{N}(p_n(\mu_n^N)-p_n(\mu_n))
\end{aligned} \tag{8.26}$$

This mean that

$$\frac{1}{\sqrt{N}}\sum\nolimits_{n=1}^{N}(p_n^N(\mu_n^N)-p_n(\mu_n^N))$$

and

$$\frac{1}{\sqrt{N}}\sum\nolimits_{n=1}^{N}(p_n(\mu_n^N)-p_n(\mu_n))$$

are asymptotically normal. Then using the results obtained for 8.24, 8.25 and 8.26 in 8.22. We thus establish that

$$0=\{E[D_{0n}]+o_p(1)\}\sqrt{N}(\widetilde{\Theta}^N-\Theta^N)+op(1) \tag{8.27}$$

Noting that $E[D_{0n}]$ is nonsingular, 8.27 implies that $\sqrt{N}(\widetilde{\Theta}^N-\Theta^N)$ is $op(1)$ as claimed. Q.E.D.

9. Appendix 2

In part B of this appendix, we describe in more detail the construction of our sample and the construction of the variables used in our study. We used data from the Family-Individual File , Childbirth and Adoption History File, and the Marriage History File of the Michigan Panel Study of Income Dynamics (PSID). The Family- Individual File contains a separate record for each member of all households included in the survey in a given year. The Childbirth and Adoption History File contains information collected in 1985-1992 waves of PSID regarding histories of childbirth and adoption. The file contains details about childbirth and adoption events of eligible people living in a PSID family at the time of the interview in any wave from 1985 through 1992. Each set of records for a specified individual contains all known cumulative data about the timing and circumstances of his or her childbirth and adoption experience up to and including 1992, or those waves during that period when the individual was in a responding family unit. If an individual has never had any children, one record indicates that report. Note that “eligible” here means individuals of childbearing age in responding families. Similarly, the 1985-1992 Marriage History file contains retrospective histories of marriages for individuals of marriage-eligible age living in a PSID family between 1985 and 1992. Each set of records for a specified individual contains all known cumulative data about the timing and circumstances of his or her marriages up to and including 1992, or those waves during that period when the individual was in a responding family unit.

Our sample selection started from the Childbirth and Adoption history file, which contains 24,762 individuals. We initially selected women by setting "sex of individual" variable equal to two. Out of an initial sample of 24,762 individuals included in the Childbirth and Adoption file, this initial selection produced a sample of 12,784 female. We then drop any individual who was in the survey for four years or less, this selection criteria eliminated a further 1,946 individuals from our sample. We then drop all individuals who were older than 45 in 1967, this eliminated an additional 1,531 individuals. We then drop all individuals that were less than 14-years-old in 1991, this eliminated an additional 385 individuals.

The corresponding number of observations for the interviewing year 1968 through 1992 are given by 5,429,5,608, 5,793,5,970, 6,197, 6,346, 6,510, 6,696, 6,876, 7,094, 7,236, 7,320, 7,393, 7,455, 7,551, 7,634, 7,680, 7,761, 7,712, 7,666, 7,618, 7,574, 7,532, 7,378 and 7,233, respectively.

Since individuals who had become non-respondents as of 1992, either because they and their families were last to the study or they were mover-out non-respondents in years prior to the 1992 interviewing year, are not in the twenty-five Family-Individuals Respondents File, the number of observations increases with the interviewing years.

There were coding errors which occurred for the different measures of consumption in the PSID from which we construct our consumption measure. In particular, our measure of food consumption expenditures for a given year is obtained by summing the values of annual food expenditures for meals at home, annual food expenditures for eating out, and the value of food stamps received for the year. We measured consumption expenditures for year t by taking 0.25 of the value of this variable for the year $t-1$ and 0.75 of its value for the year t. The second step was taken to account for the fact that the survey questions used to elicit information about household food consumption is asked sometime in the first half of the year, while the response is dated in the previous year.

The variables used in the construction of the measure for total expenditures are also subject to the problem of truncation from above in the way they are coded in the 1983 PSID data tapes. The truncation value for the value of food stamps received for that year is \$999.00, while the relevant value for this variable in the subsequent years and for the value of food consumed at home and eating out is \$9,999.00. Taken by itself, the truncation of different consumption variables resulted in a loss of 467 person-years. We also use variables describing various demographic characteristics of the women in our sample. The dates of birth of the women were obtained from the Child Birth and Adoption file. The age variable resulted in a loss of 162 individuals.

The race of the individual or the region where they are currently residing were obtained from the Family portion of the data record. We defined the region variable to be the geographical region in which the household resided at the time of the annual interview. This variable is not coded consistently across the years. For 1968 and 1969, the values 1 to 4 denote the regions Northeast, Northcentral, South and West. For 1970 and 1971, the values 5 and 6 denote the regions Alaska and Hawaii, and foreign country, respectively. After 1971 a value of 9 indicates missing data but no person years were lost due to missing data for these variables.

We used the family variable "Race of The Household Head" to measure the race variable in our study. For the interviewing years 1968-1970, the values 1 to 3 denote White, black, and Puerto Rican or Mexican, respectively. 7 denotes other (including Oriental and Philippino), and 9 denotes missing data. For 1971 and 1972, the third category is redefined as Spanish-American or Cuban and between 1973-1984, just Spanish American. After 1984, the variable was coded in such a way that 1-6 correspond to the categories White, Black, American Indian, Aleutian or Eskimo, and Asian or Pacific Islander, respectively. A value of 7 denotes the other category, a value of 9 denotes missing. We used all available information for all the years to assign the race of the individual for years in the sample when that information was available.

We used a combination of individual and family level variables to construct our measure of educational attainment. This was because the variable for the individual does not contain data for the head of the household or wife, this we obtained from the family level files.

The marital status of a women in our subsample was determined by using the marriage history file. The number of individuals in the household and the total number of children within that household were also determined from the family level variables of the same name. In 1968, a code for missing data (equal to 99) was allowed for the first variable, but in the other years, missing data were assigned. The second variable was truncated above the value of 9 for the interviewing years 1968 and 1971. After 1975, this variable denotes the actual number of Children within the family unit.

We constructed some additional variables. The variable showing the value of home-ownership was constructed by multiplying the value of a household's home by an indicator variable determining home ownership. A similar procedure was followed to generate value of rent paid and rental value of free housing for a household. Mortgage payment and Principal of Mortgage outstanding were obtained from the family variables of the same names. Finally, household income was measured from the PSID variable total family money income, which included taxable income of head and wife total transfer of head and wife, taxable income of others in the family units and their total transfer payments.

We used two different deflators to convert such nominal quantities as average hourly earnings, household income, and so on to real. First, we defined the (spot) price of food consumption to be the numeraire good at t in the theoretical section. We accordingly measured real food consumption expenditures and real wages as the ratio of the nominal consumption expenditures and wages and the annual Chain-type price deflator for food consumption expenditures published in table t.12 of the National Income and Products. On the other hand, we deflated variables such as the nominal value of home ownership or nominal family income by the Chain-type price deflator for total personal consumption expenditures.

TABLE I

Main Sample Summary Statistics

(Standard Deviation is parenthesis)

Variable	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
Age	18.57	19.25	19.77	20.31	20.82	21.36	21.91	22.41	22.91	23.31	24.31	25.31	26.31
	(12.59)	(12.67)	(12.81)	(13.01)	(13.17)	(13.33)	(13.51)	(13.72)	(13.93)	(14.19)	(14.19)	(14.19)	(14.19)
Education	11.31	11.30	11.27	11.25	11.22	11.20	11.18	11.14	11.13	11.13	11.13	11.13	11.13
	(2.31)	(2.31)	(2.30)	(2.30)	(2.29)	(2.28)	(2.28)	(2.25)	(2.25)	(2.22)	(2.22)	(2.22)	(2.22)
Family Size	5.78	5.67	5.49	5.36	5.23	4.98	4.81	4.65	4.52	4.42	4.31	4.15	4.05
	(2.62)	(2.66)	(2.66)	(2.63)	(2.61)	(2.45)	(2.38)	(2.29)	(2.22)	(2.15)	(2.05)	(1.95)	(1.90)
Young Kids	1.61	1.54	1.45	1.38	1.31	1.26	1.20	1.14	1.08	1.02	0.97	0.93	0.88
	(1.46)	(1.43)	(1.40)	(1.37)	(1.34)	(1.32)	(1.309)	(1.28)	(1.25)	(1.23)	(1.20)	(1.18)	(1.15)
Old Kids	0.50	0.49	0.48	0.47	0.46	0.44	0.42	0.39	0.37	0.36	0.35	0.35	0.35
	(1.13)	(1.12)	(1.09)	(1.08)	(1.04)	(0.99)	(0.94)	(0.90)	(0.85)	(0.81)	(0.79)	(0.75)	(0.74)
Birth	4.88	5.08	5.42	5.17	4.98	4.91	4.79	4.95	5.03	5.22	4.92	5.47	5.49
	(21.55)	(21.97)	(22.64)	(22.14)	(21.74)	(21.61)	(21.36)	(21.68)	(21.86)	(22.25)	(21.64)	(22.75)	(22.78)
Age at 1st Birth	28.82	28.81	28.88	28.86	28.90	29.00	29.05	29.15	29.19	29.22	29.22	29.22	29.22
	(11.31)	(11.31)	(11.33)	(11.34)	(11.36)	(11.39)	(11.41)	(11.43)	(11.44)	(11.44)	(11.44)	(11.44)	(11.44)
Martial status	23.11	24.51	25.65	27.01	28.59	30.13	31.33	31.95	32.66	33.73	34.35	35.80	36.73
	(42.16)	(43.02)	(43.68)	(44.40)	(45.19)	(45.88)	(46.39)	(46.63)	(46.90)	(47.28)	(47.49)	(47.95)	(48.21)
Black	43.06	42.93	42.76	42.70	42.26	42.19	41.91	41.69	41.65	41.47	41.47	41.48	41.48
	(49.52)	(49.50)	(49.48)	(49.47)	(49.40)	(49.39)	(49.35)	(49.31)	(49.30)	(49.27)	(49.27)	(49.28)	(49.27)
Hispanic	0.43	0.43	0.43	0.43	0.43	0.43	0.44	0.44	0.44	0.44	0.44	0.44	0.44
	(2.03)	(2.04)	(2.03)	(2.03)	(2.04)	(2.03)	(2.05)	(2.06)	(2.05)	(2.05)	(2.05)	(2.05)	(2.05)
North East	16.65	16.44	16.63	16.60	16.53	16.61	16.43	16.18	16.04	15.97	15.86	15.37	15.13
	(37.26)	(37.07)	(37.24)	(37.21)	(37.15)	(37.22)	(37.06)	(36.83)	(36.70)	(36.63)	(36.53)	(36.07)	(35.84)
North Central	24.58	24.78	24.50	24.55	24.75	24.703	24.26	24.20	24.06	23.82	23.68	23.45	23.56
	(43.06)	(43.18)	(43.02)	(43.04)	(43.16)	(43.13)	(42.87)	(42.83)	(42.75)	(42.60)	(42.51)	(42.38)	(42.44)
South	44.72	44.61	44.44	44.62	44.85	44.73	45.29	45.26	45.25	45.47	45.33	45.95	46.122
	(49.73)	(49.71)	(49.70)	(49.71)	(49.74)	(49.731)	(49.78)	(49.78)	(49.78)	(49.79)	(49.79)	(49.84)	(49.85)
L.F. Participation	25.96	29.51	32.60	34.13	35.99	36.42	38.87	40.43	42.09	42.92	45.62	50.17	52.62
	(43.85)	(45.61)	(46.88)	(47.42)	(48.00)	(48.12)	(48.75)	(49.08)	(49.37)	(49.50)	(49.81)	(50.00)	(49.94
Hours Worked	299	339	375	394	415	436	480	494	515	542	586	656	693
	(641)	(670)	(697)	(706)	(725)	(739)	(768)	(764)	(778)	(790)	(815)	(846)	(857)
Hourly earnings	7.03	7.31	8.70	7.87	8.26	8.40	8.46	8.43	8.05	8.80	8.87	8.66	8.91
	(5.93)	(5.56)	(27.50)	(6.98)	(7.83)	(7.22)	(7.29)	(9.26)	(6.67)	(10.39)	(10.59)	(8.00)	(8.01
Food Consumption	7890	7577	7446	7464	7253	7704	7871	7323	7220	6665	6465	6546	63353
	(3432)	(3312)	(3226)	(3321)	(3258)	(3271)	(3606)	(3303)	(3358)	(3202)	(3264)	(3270)	(3213)
Observations	4691	4859	5040	5206	5407	5580	5758	5946	6142	6356	6356	6357	6357

Note: L.F. is an abbreviation for labor force. Education is years of completed education. Young kids is the number children less than 6 years old and Old kids is the number of children between the age of 6 and 14. Food consumption and hourly earnings are measured in 1996 dollars. Birth, race, martial status and regions is measured in percentage.

TABLE I (Continued)

Main Sample Summary Statistics

(Standard Deviation is parenthesis)

Variable	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Age	27.31	28.30	29.30	30.21	31.15	32.13	33.10	34.09	35.06	36.03	37.01	38.03
	(14.19)	(14.19)	(14.19)	(14.18)	(14.18)	(14.19)	(14.20)	(14.21)	(14.21)	(14.22)	(14.19)	(14.19)
Education	11.1312	11.13	11.13	11.13	11.13	11.136	11.13	11.14	11.14	11.15	11.15	11.16
	(2.22)	(2.22)	(2.22)	(2.22)	(2.22)	(2.22)	(2.22)	(2.23)	(2.23)	(2.23)	(2.23)	(2.23)
Family Size	3.97	3.87	3.79	3.72	3.63	3.56	3.48	3.42	3.37	3.30	3.25	3.17
	(1.90)	(1.83)	(1.80)	(1.72)	(1.67)	(1.65)	(1.61)	(1.58)	(1.58)	(1.57)	(1.57)	(1.54)
Young Kids	0.84	0.79	0.74	0.68	0.64	0.58	0.53	0.49	0.44	0.39	0.34	0.29
	(1.12)	(1.09)	(1.05)	(1.01)	(0.97)	(0.93)	(0.89)	(0.84)	(0.79)	(0.74)	(0.68)	(0.62)
Old Kids	0.35	0.36	0.37	0.37	0.38	0.39	0.41	0.41	0.41	0.42	0.41	0.41
	(0.73)	(0.73)	(0.73)	(0.74)	(0.75)	(0.75)	(0.76)	(0.77)	(0.77)	(0.77)	(0.76)	(0.77)
Birth	4.92	5.09	5.01	4.95	4.75	4.98	4.99	4.78	4.62	4.86	4.99	1.24
	(21.64)	(21.99)	(21.83)	(21.70)	(21.28)	(21.76)	(21.78)	(21.35)	(20.99)	(21.5)	(21.78)	(11.08)
Age at 1st Birth	29.22	29.22	29.21	29.21	29.20	29.21	29.21	29.21	29.24	29.19	29.17	29.21
	(11.44)	(11.44)	(11.43)	(11.43)	(11.42)	(11.42)	(11.42)	(11.41)	(11.40)	(11.39)	(11.38)	(11.39)
Martial status	37.28	38.05	39.11	40.11	40.91	41.31	41.68	42.35	43.04	43.55	43.68	43.75
	(48.36)	(48.56)	(48.80	(49.01)	(49.17)	(49.24)	(49.31)	(49.41)	(49.51)	(49.58)	(49.60)	(49.61)
Black	41.48	41.48	41.48	41.28	41.22	41.22	40.97	40.79	40.61	40.46	40.31	40.05
	(49.27)	(49.27)	(49.27)	(49.23)	(49.22)	(49.22)	(49.18)	(49.14)	(49.11)	(49.08)	(49.05)	(49.01)
Hispanic	0.44	0.44	0.44	0.44	0.44	0.44	0.43	0.43	0.43	0.43	0.43	0.44
	(2.05)	(2.05)	(2.05)	(2.05)	(2.05)	(2.05)	(2.04)	(2.03)	(2.04)	(2.04)	(2.04)	(2.04)
North East	14.96	14.77	14.48	14.38	14.18	14.09	13.92	13.69	13.50	13.08	12.75	12.54
	(35.67)	(35.48)	(35.20)	(35.09)	(34.89)	(34.79)	(34.62)	(34.38)	(34.18)	(33.72)	(33.35)	(33.12)
North Central	23.38	23.27	23.23	22.99	22.79	22.48	22.39	22.09	21.68	21.65	21.60	21.84
	(42.33)	(42.26)	(42.23)	(42.08)	(41.95)	(41.75)	(41.69)	(41.48)	(41.21)	(41.19)	(41.15)	(41.32)
South	45.95	45.95	46.02	46.25	45.94	45.67	45.64	45.57	45.43	45.60	45.10	45.53
	(49.84)	(49.84)	(49.84)	(49.86)	(49.83)	(49.81)	(49.81)	(49.80)	(49.79)	(49.81)	(49.76)	(49.80)
L.F. Particaption	54.47	55.01	55.85	57.74	61.75	62.71	64.57	67.01	68.32	70.76	72.13	71.47
	(49.80)	(49.75)	(49.65)	(49.39)	(48.60)	(48.36)	(47.83)	(47.02)	(46.52)	(45.48)	(44.83)	(45.15)
Hours Worked	725	744	761	790	879	907	937	980	1017	1041	1077	1085
	(866)	(875)	(883)	(889)	(915)	(929)	(927)	(938)	(947)	(937)	(946)	(962)
Hourly earnings	8.73	8.65	8.67	10.78	9.88	9.93	9.58	9.97	10.11	10.42	10.33	10.57
	(7.65)	(7.57)	(7.46)	(56.83)	(38.91)	(24.31)	(9.46)	(9.69)	(8.66)	(15.53)	(11.77)	(9.33)
Food Consumption	6162	5886	5753	5875	5963	5846	5741	5766	5631	5518	5435	5512
	(3118)	(2951)	(2966)	(3123)	(3707)	(3128)	(3105)	(2933)	(2835)	(3010)	(2834)	(3488)
Observations	6357	6357	6357	6298	6246	6159	6067	5971	5885	5815	5725	5627

Note: L.F. is an abbreviation for labor force. Note: Education is years of completed education. Young kids is the number children less than 6 years old and Old kids is the number of children between the age of 6 and 14. Food consumption and hourly earnings are measured in 1996 dollars. Birth, race, martial status and regions is measured in percentage.

TABLE II
Wage Equation

$$\ln(w_{nt}) = \ln(\omega_t) + \ln(\mu_n) + z'_{nt}B_3 + \sum_{s=1}^{4}(\delta_{1s}h_{n,t-s} + \delta_{2s}d_{2n,t-s} + \delta_{2s}d_{4n,t-s})$$

Variable	Parameter	Estimate†
Lags of hours worked		
$\triangle h_{n,t-1}$	δ_{11}	14.1011 (0.2337)
$\triangle h_{n,t-2}$	δ_{12}	10.9974 (0.2471)
$\triangle h_{n,t-3}$	δ_{13}	8.8360 (0.2437)
$\triangle h_{n,t-4}$	δ_{14}	5.4729 (0.2227)
Lags of participation		
$\triangle(d_{2n,t-1} + d_{4n,t-1})$	δ_{21}	-6.8664 $(4.01E-02)$
$\triangle(d_{2n,t-2} + d_{4n,t-2})$	δ_{22}	-4.4241 $(4.46E-02)$
$\triangle(d_{2n,t-3} + d_{4n,t-3})$	δ_{23}	-2.8986 $(4.44E-02)$
$\triangle(d_{2n,t-4} + d_{4n,t-4})$	δ_{24}	-1.6065 $(3.92E-02)$
Socioeconomic Variables		
$\triangle AGE^2_{nt}$	B_{31}	-0.0114 $(3.0E-04)$
$\triangle(AGE_{nt} \times EDU_{nt})$	B_{32}	0.0161 $(3.1E-03)$

† Estimated standard error in parenthesis. Note: EDU is the years of completed education.

TABLE III
Log Frisch Consumption Demand

$\ln(x_{nt}) = (1-\alpha)^{-1}[z'_{nt}B_2 - \ln(\eta_n) - \ln(\lambda_t) + \epsilon_{0nt}]$

Variable	Parameter	Estimate†
Risk Aversion	α	0.98 $(1.9E-02$
Socioeconomic variables		
$\triangle FAM_{nt}$	$(1-\alpha)^{-1}B_{21}$	$3.19E-02$ $(3.0E-04)$
$\triangle YKID_{nt}$	$(1-\alpha)^{-1}B_{22}$	$-3.33E-02$ $(1.6E-03)$
$\triangle OKID_{nt}$	$(1-\alpha)^{-1}B_{23}$	$-1.12E-02$ $(1.2E-03)$
$\triangle AGE^2_{nt}$	$(1-\alpha)^{-1}B_{24}$	$-1.0E-04$ (0.0000)
Region Dummies		
$\triangle NC_{nt}$	$(1-\alpha)^{-1}B_{25}$	$-3.7E-03$ $(3.3E-03)$
$\triangle SO_{nt}$	$(1-\alpha)^{-1}B_{26}$	$-1.19E-02$ $(3.2E-03)$

† Estimated standard error in parenthesis. Note: EDU is the years of completed education,FAM is the number of individual in the household, YKid is the number of kids less than 6 years old, Okid is the number of children between age 6 and 14, while NC and SO are regional dummies for the North Central and South respectively.

TABLE IV
Fixed Utility from Labor Force Participation
$(d_{nt2} + d_{nt4})\, z'_{nt} B_0$

Variable	Parameter	Estimate[†]
$CONSTANT$	B_{00}	-5.23 (0.66)
AGE_{nt}	B_{01}	0.26 $(4.6E-2)$
AGE^2_{nt}	B_{02}	$-4.8E-3$ $(7.46E-4)$
$AGE_{nt} \times EDUC_n$	B_{03}	$2.0E-3$ $(4.25E-4)$
$MART.STATUS_{nt}$	B_{04}	0.12 $(4.9E-2)$
$BLACK_{nt}$	B_{05}	$-5.9E-2$ $(5.2E-2)$

† Estimated standard error in parenthesis. Note: EDU is the years of completed education and MART. STATUS is a dummy variable equal 1 if the female is married and zero otherwise.

TABLE V
Nurturing Time Cost

$$c_{nt} = \sum_{s=0}^{5} \phi_s b_{n,t-s} + \phi_6 \sum_{s=6}^{t} b_{n,t-s}$$

Variable	Parameter	Estimate†
b_{nt}	ϕ_0	$4.0E-5$ $(2.28E-1)$
$b_{n,t-1}$	ϕ_1	$3.8E-4$ $(1.65E-1)$
$b_{n,t-2}$	ϕ_2	$1.1E-4$ $(9.0E-2)$
$b_{n,t-3}$	ϕ_3	$7.2E-5$ $(6.1E-2)$
$b_{n,t-4}$	ϕ_4	$2.7E-3$ $(1.0E-1)$
$b_{n,t-5}$	ϕ_5	$1.7E-2$ $(5.5E-2)$
$\sum_{s=6}^{t} b_{n,t-s}$	ϕ_6	$1.8E-2$ $(7.0E-03)$

† Estimated standard error in parenthesis

TABLE VI
Utility of Leisure

$$z'_{nt}B_{11}l_{nt}+\sum_{s=0}^{4}\delta_s l_{n,t-s}l_{nt}$$

Variable	Parameter	Estimate†
l_{nt}	B_{110}	1.96 (0.22)
$AGE_{nt}\times l_{nt}$	B_{111}	$-2.1E-3$ $(1.13E-2)$
$AGE^2_{nt}\times l_{nt}$	B_{113}	$-5.1E-6$ $(1.8E-4)$
$AGE_{nt}\times EDUC_n\times l_{nt}$	B_{113}	$2.3E-4$ $(1.0E-4)$
$MART.STATUS_{nt}\times l_{nt}$	B_{114}	$-3.6E-2$ $(1.2E-2)$
$BLACK_{nt}\times l_{nt}$	B_{115}	$5.4E-2$ $(1.2E-2)$
l^2_{nt}	δ_0	-0.14 $(5.8E-2)$
$l_{nt}l_{n,t-1}$	δ_1	-0.33 $(6.6E-2)$
$l_{nt}l_{n,t-2}$	δ_2	-0.17 $(4.5E-2)$
$l_{nt}l_{n,t-3}$	δ_3	-0.24 $(4.5E-2)$
$l_{nt}l_{n,t-4}$	δ_4	$-1.5E-2$ $(4.3E-2)$
$variance$	σ^{-1}	$1.0e-2$ $(7.8e-3)$

† Estimated standard error in parenthesis. Note: EDU is the years of completed education and MART. STATUS is a dummy variable equal 1 if the female is married and zero otherwise.

TABLE VII
Utility from Offspring and Monetary Costs

$$\gamma_0 b_{nt} + \sum_{k=1}^{5} \gamma_k b_{nt} b_{n,t-k} + \gamma_6 \sum_{k=6}^{t} b_{nt} b_{n,t-k}$$

$$z'_{nt}\pi_1$$

Variable	Parameter	Estimate†
b_{nt}	γ_0	-1.17 (0.30)
$b_{nt}b_{n,t-1}$	γ_1	-0.23 (0.08)
$b_{nt}b_{n,t-2}$	γ_2	0.79 (0.12)
$b_{nt}b_{n,t-3}$	γ_3	0.41 (0.08)
$b_{nt}b_{n,t-4}$	γ_4	-0.14 $(7.5E-2)$
$b_{nt}b_{n,t-5}$	γ_5	-0.22 (0.08)
$\sum_{k=6}^{t} b_{nt}b_{n,t-k}$	γ_6	-0.41 (0.05)
$NO\ HIGH\ SCH_{nt}$	π_1	$-5.1E-3$ (0.13)
$BLACK_{nt}$	π_2	$1.6E-2$ $(8.4E-2)$

† Estimated standard error in parenthesis. Note: NO HIGH SCH is dummy variable equal one if the female has not graduated high school and zero otherwise.

TABLE VIII
Predicted Completed Fertility Outcomes under Different Policies

Marital	Education	Actual	Estimation	Expenses	Daycare	Wages	Retraining
Black							
M	<	2.12	2.45	2.63	2.41	2.57	2.69
	HS	1.93	2.03	2.60	2.8	2.19	3.00
	>	1.35	1.68	1.71	2.3	1.66	2.50
U	<	2.15	2.35	2.56	2.58	2.41	2.57
	HS	1.82	1.97	2.04	2.1	1.98	2.05
	>	1.23	1.17	1.26	1.85	1.37	2.24
Hispanic							
M	<	2.08	2.19	2.23	2.31	2.25	2.02
	HS	1.83	1.79	1.89	2.03	1.87	2.35
	>	1.55	1.46	1.50	1.87	1.49	2.03
U	<	2.00	2.15	2.23	2.26	2.23	2.31
	HS	1.78	1.87	1.96	2.12	1.89	2.38
	>	1.46	1.56	1.67	2.00	1.72	2.30
White							
M	<	1.78	2.04	2.12	2.16	2.09	2.07
	HS	1.34	1.52	1.63	2.30	1.67	2.45
	>	1.12	1.23	1.32	1.97	1.24	2.03
U	<	1.47	1.56	1.54	1.78	1.58	1.87
	HS	1.25	1.31	1.56	1.90	1.67	2.08
	>	1.11	1.24	1.39	1.78	1.48	2.03

TABLE VIII(Percentage Change)
Predicted Completed Fertility Outcomes under Different Policies

					Percentage Change of Estmation			
	Marital	Education	Actual	Estimation	Expenses	Daycare	Wages	Retraining
Black								
	M	<	2.12	2.45	7.35	-1.63	4.90	9.80
		HS	1.93	2.03	28.08	37.93	7.88	47.78
		>	1.35	1.68	1.79	36.90	-1.19	48.81
	U	<	2.15	2.35	8.94	9.79	2.55	9.36
		HS	1.82	1.97	3.55	6.60	0.51	4.06
		>	1.23	1.17	7.69	58.12	17.09	91.45
Hispanic								
	M	<	2.08	2.19	1.83	5.48	2.74	-7.76
		HS	1.83	1.79	5.59	13.41	4.47	31.28
		>	1.55	1.46	2.74	28.08	2.05	39.04
	U	<	2	2.15	3.72	5.12	3.72	7.44
		HS	1.78	1.87	4.81	13.37	1.07	27.27
		>	1.46	1.56	7.05	28.21	10.26	47.44
White								
	M	<	1.78	2.04	3.92	5.88	2.45	1.47
		HS	1.34	1.52	7.24	51.32	9.87	61.18
		>	1.12	1.23	7.32	60.16	0.81	65.04
	U	<	1.47	1.56	-1.28	14.10	1.28	19.87
		HS	1.25	1.31	19.08	45.04	27.48	58.78
		>	1.11	1.24	12.10	43.55	19.35	63.71

TABLE IX

Predicted Annual Labor Force Participation Rates under Different Policies

Marital	Education	Actual	Estimation	Expenses	Daycare	Wages	Retraining
Black							
M	<	0.570	0.452	0.436	0.421	0.476	0.423
	HS	0.673	0.772	0.722	0.732	0.724	0.723
	>	0.781	0.729	0.745	.742	0.732	0.732
U	<	0.678	0.616	0.606	0.627	0.601	0.591
	HS	0.723	0.763	0.751	0.749	0.761	0.763
	>	0.897	0.912	0.913	0.916	0.915	0.921
Hispanic							
M	<	0.612	0.634	0.632	0.625	0.623	0.618
	HS	0.722	0.745	0.739	0.738	0.737	0.735
	>	0.823	0.856	0.842	0.845	0.835	0.812
U	<	0.732	0.742	0.692	0.693	0.695	0.683
	HS	0.752	0.765	0.745	0.746	0.748	0.746
	>	0.824	0.878	0.867	0.857	0.856	0.872
White							
M	<	0.678	0.693	0.687	0.598	0.662	0.597
	HS	0.897	0.876	0.874	0.873	0.878	0.871
	>	0.912	0.927	0.921	0.928	0.926	0.923
U	<	0.753	0.734	0.727	0.714	0.701	0.692
	HS	0.857	0.876	0.767	0.798	0.845	0.855
	>	0.866	0.857	0.867	0.849	0.867	0.856

TABLE IX(Percentage Change)

Predicted Annual Labor Force Participation Rates under Different Policies

	Marital	Education	Actual	Estimation	Percentage Change of Estmation Expenses	Daycare	Wages	Retraining
Black								
	M	$<$	0.57	0.452	-3.54	-6.86	5.31	-6.42
		HS	0.673	0.772	-6.48	-5.18	-6.22	-6.35
		$>$	0.781	0.729	2.19	1.78	0.41	0.41
	U	$<$	0.678	0.616	-1.62	1.79	-2.44	-4.06
		HS	0.723	0.763	-1.57	-1.83	-0.26	0.00
		$>$	0.897	0.912	0.11	0.44	0.33	0.99
Hispanic								
	M	$<$	0.612	0.634	-0.32	-1.42	-1.74	-2.52
		HS	0.722	0.745	-0.81	-0.94	-1.07	-1.34
		$>$	0.823	0.856	-1.64	-1.29	-2.45	-5.14
	U	$<$	0.732	0.742	-6.74	-6.60	-6.33	-7.95
		HS	0.752	0.765	-2.61	-2.48	-2.22	-2.48
		$>$	0.824	0.878	-1.25	-2.39	-2.51	-0.68
White								
	M	$<$	0.678	0.693	-0.87	-13.71	-4.47	-13.85
		HS	0.897	0.876	-0.23	-0.34	0.23	-0.57
		$>$	0.912	0.927	-0.65	0.11	-0.11	-0.43
	U	$<$	0.753	0.734	-0.95	-2.72	-4.50	-5.72
		HS	0.857	0.876	-12.44	-8.90	-3.54	-2.40
		$>$	0.866	0.857	1.17	-0.93	1.17	-0.12