Life-Cycle Fertility and Human Capital Accumulation *

George-Levi Gayle Robert A. Miller

May 31 2002

Abstract

In this paper we estimate a choice theoretic model of female labor supply and fertility behavior with panel data, and use our estimates to predict how changes in family policy would affect the behavior of these variables over the life cycle. The motivation for investigating dynamic interactions between fertility and female labor supply comes from broad trends in aggregate behavior and also empirical results from previous empirical work using cross sectional and panel data.

1. INTRODUCTION

In this paper we estimate a choice theoretic model of female labor supply and fertility behavior with panel data, and use our estimates to predict how changes in family policy would affect the behavior of these variables over the life cycle. The motivation for investigating dynamic interactions between fertility and female labor supply comes from broad trends in aggregate behavior and also empirical results from previous empirical work using cross sectional and panel data.

At the aggregate level the trends in birth rates and female labor supply and the wages of females in developed countries are striking. Measures of annual total fertility rates (TFRs) provide a useful way of summarizing trends in fertility at

* *This is an incomplete draft. Please do not quote without permission.*

the aggregate level. There are two basic measures, by period and/or by cohort. A period TFR predicts the total lifetime number of birth if a representative woman realized the age -specific fertility rates that prevail in a given year. A cohort TFR measures the number of children born to a particular birth cohort. Both measures used for the US and most developed countries show there has been a substantial decline in total completed fertility from the beginning to the end of the twentieth century.[1] For example at the beginning of the century a typical woman in the USA who reached the age of 45 bore, on average, four children over her lifetime and by the end of the century, that number has fallen to only 1.9. The period TFR was 4.0 in Sweden in 1905 and has declined to 1.4 by the mid 1980's.[2]. Although the majority of women in developed countries eventually bear a child, an increasing fraction of them bear no children. The incidence of childlessness has fluctuated over the twentieth century but seems to have increased towards the end of the century. For example, Hotz et al (1996) report that over the last 20 years the incidence of childlessness has almost doubled in the US, going from 9% of women who reached age 40-45 in 1978 to 18% for comparably aged women in 1994. [3] The changes in total fertility rates over the last century have been accompanied by changes in the life-cycle timing of childbearing by women who were of fecundity age during the era. The so called baby boom of US was in essence fueled by women shifting their childbearing to earlier ages and the subsequent bust was largely the result of the tendency for childbearing to be delayed.[4]

Parallel to this marked decline in childbearing has been a rise over time in female labor force participation in both developed and developing countries.[5] In the US the participation of all wives increased by 36% over the last 25 years, the rates of mothers with children under the age of three increased by 83% and by 91% for women with children one year old or younger. The rise in female participation in the labor force has been accompanied by a decline in the difference between male and female earnings of full time workers.

These aggregate trends can be rationalized by simple economic models of household choices. The parents' demand for children depend on prices of the inputs used in raising children, and levels household income and/or wealth. The

[1]See Hotz et. al.(1996) and reference there in for a survey on this issue.

[2]See Walker(1995) for details.

[3]Some of this increase reflects the decline in the fraction of women who are married over the same period. But the incidence of childlessness has risen among married women as well. See Hotz et al (1996) and reference therein for more details.

[4]See Hotz et al (1996) for details on this shift.

[5]See Heckman and Killingsworth(1996).

opportunity cost for the mother's care and nurture for her offspring is the female wage rate, and as this increases female labor supply increases and the fertility rate falls. Controlling for the wage, a simple model predicts that families with more wealth have more children. If child would it should not be surprising that the can be prices of other goods and services, including child quality, the mother's market wage rate(s) However the sign and magnitudes of these effects are most often than not, not unambiguously indicated by theory which mean that these issues are empirical questions.

Empirical investigations of micro data bear out these predictions. The mother's wage is negatively related to the demand for children all type of models of fertility (Ward and Butz (1979, 1980), Hotz and Miller (1988)).[6] The relationship between household wealth and fertility controlling for the opportunity cost of the mother's time is harder to document. Schultz (1976) reports many earlier studies that found a positive relationship between family income and/or consumption and parental fertility. Several more recent studies also found a positive relationship between parental fertility and husband income or other household income (Ward and Butz (1980), Wolpin (1984) Hotz and Miller (1988)). On the other hand Willis(1973) finds a U shape relationship between completed fertility and husband income while Heckman and Walker(1990) found that there weak if any relationship between husband's income (and also female wages) and the incidence of childlessness.

A further complication to interpreting the evidence about the effects of higher female wages on fertility was introduced by Becker (1965), who argued that parents not only choose the quantity of offspring but also their quality. Thus highly educated parents might choose to have a lower number of children, but invest more inputs in them. This modification to the basic model provides another reason why women with higher opportunity costs have lower fertility rates. Therefore the almost unanimous finding, that parental demand for children are negatively related to the educational level of the mothers, is hardly surprising. This result holds in static models of completed fertility (Willis (1973))[7], reduced form dynamic models (Walker (1996), Hotz and Miller (1988), Hotz, Heckman and Walker (1990)) and structural dynamic models of fertility and contraceptive practices (Wolpin (1984), Hotz and Miller (1993)). In a more direct test of the role of the mother's inputs on measures of childhood achievement, Michaels (1992) finds that after controlling for labor supply the offspring of more highly educated women perform better at

[6]See Paul Schultz (1976) for a summary of some the earlier studies.

[7]See Paul Schulz (19..) survey for a comprehensive summary of these results.

school.

While the basic models of fertility can be used to explain the relationship between total fertility rates and measures of female labor supply and household wealth, they have much less to say about the timing of births and how this is related to female labor supply over the life cycle. Empirical dynamic models of fertility find that the time costs young children impose on their mothers help to rationalize the spacing of later births (Hotz and Miller (1988). Similarly there is strong evidence from dynamic models of labor supply and human capital accumulation that, in addition to providing wages, work experience is a form of investment in human capital that increases the future wage rate (Eckstein and Wolpin (1989), Miller and Sanders (1997), Altug and Miller (1998)). Thus the costs of staying home to raise children are significantly greater than the current wages forgone.

These empirical results suggest that the patterns of investing for work force through current labor force participation is intertwined with decisions about the timing and the amount of offspring a household chooses. This study then is an attempt to combine both forms of human or family capital within a unified framework. Only by capturing both kinds of choices can one reasonably expect to answer policy questions that bear upon how households will respond their contraceptive and labor force behavior to changes in provisions for maternity leave, child care facilities and the tax treatment of dependents, to name just three examples of topical interest.

The next two sections provide the theoretical underpinnings to our empirical investigations. Section 2 lays out a life cycle model of labor supply and fertility. Then in Section 3 we derive the conditions implied by dynamic optimization that form the basis for identification and estimation. Sections 4 through 7 are the heart of the estimation. First we explain our estimation strategy in Section 4. Then we briefly summaraize the sample of households used in our empirical work, which is drawn from the Panel Study of Income Dynamics (PSID). In Section 6 we report our estimates of the wage equation from wage and labor supply data. The wealth effects of the household are estimated in Section 7 from data on consumption. In Section 8 we estimate from data on labor supply and births the parameters that determine preferences for children, as well as the direct intertemporal effects of labor supply or leisure on household utility. The last two sections of the paper explore the quantitative implications of our model. They use the estimates obtained from the body of the estimation to conduct some policy simulations and summarize our findings.

2. A FRAMEWORK

The model is set in discrete time, and measures the woman's age beyond adolescence with periods denoted by $t \in \{0, 1, \ldots, T\}$. It analyzes the accumulation of two kinds of human capital, offspring, and labor market experience.

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 demographic variables. The latter, denoted by z_{nt}, include such variables as age, formal education, regional location, ethnicity and race. It is assumed that z_{nt} is independently distributed over the population with cumulative probability distribution function $F_0(z_{nt+1} \mid z_{nt})$. Let h_{nt} denote the proportion of time worked in period t as a fraction of the total time available in the period, let d_{nt} denote participation in period t, that is an indicator if $h_{nt} > 0$. We assume that the mapping from experience to the current wage rate is given by:

$$w_{nt} = g(d_{nt-\rho}, ..., d_{nt-1}, h_{nt-\rho}, ..., h_{nt-1}, z_{nt}) \tag{2.1}$$

for some positive integer ρ. Thus Equation 2.1 shows that, in addition to the demographic variables, the current wage depends on past participation and past hours up to ρ periods ago.

The birth of a child at period t is denoted by the indicator variable $b_{nt} \in \{0, 1\}$. It contributes 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 can be summarized by the random variable, U_{0nt}, defined as:

$$U_{0nt} = b_{nt}(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt}) \tag{2.2}$$

where $N_{nt} = \sum_{s=0}^{t} b_{ns}$

Raising children requires market expenditure and parental time. We assume that the discounted cost of expenditures of raising a child is π, a parameter that varies with household demographics, and that a k year old requires nurturing time of ρ_k. Letting c_{nt} denote the fraction of time the n^{th} household spend nurturing children in the household, our assumption about nurturing implies:

$$c_{nt} = \sum_{k=0}^{t} \rho_k b_{n,t-k} \tag{2.3}$$

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:

$$1 = h_{nt} + l_{nt} + c_{nt} \tag{2.4}$$

Apart from having utility for children, household utility also comes from its consumption of market goods, denoted x_{nt}, and leisure, l_{nt}. We assume that preferences are additive in consumption and leisure but not separable with respect to leisure at different dates. To model this dependence, define $z_{nt}^* = (b_{nt-M}, ..., b_{nt-1}, h_{nt-\rho}, ..., h_{nt-1}, z'_{nt})$, where the first ρ elements of z_{nt}^* capture the dependence of the current household state on lagged labor supply and birth choices, and the remaining elements are the set of observed demographics. Letting U_{1nt} represent the fixed utility costs of the n^{th} female from working in period t, we assume:

$$U_{1nt} = u_1(z_{nt}^*, l_{nt}) + u_2(z_{nt}^*, d_{nt})$$

This formulation incorporates both fixed and variable utility costs associated with working. It models the variable costs of working as a mapping of observed household characteristics alone, but allows participation in the work force to be determined by observed factors, entering through $u_2(z_{nt}^*, d_{nt})$. We assume that $u_1(z_{nt}^*, 1 - h_{nt} - c_{nt})$ is a concave increasing function in l_{nt}.

Let first recast the decision process of the individual by defining the following indecator variables.

$$\mathbb{I}_{0nt} = \left\{ \begin{array}{cc} 1 & \text{if } d_{nt} = 0 \text{ and } b_{nt} = 0 \\ 0 & \text{otherwise} \end{array} \right\} \tag{2.5}$$

$$\mathbb{I}_{1nt} = \left\{ \begin{array}{cc} 1 & \text{if } d_{nt} = 1 \text{ and } b_{nt} = 0 \\ 0 & \text{otherwise} \end{array} \right\} \tag{2.6}$$

$$\mathbb{I}_{2nt} = \left\{ \begin{array}{cc} 1 & \text{if } d_{nt} = 0 \text{ and } b_{nt} = 1 \\ 0 & \text{otherwise} \end{array} \right\} \tag{2.7}$$

$$\mathbb{I}_{3nt} = \left\{ \begin{array}{cc} 1 & \text{if } d_{nt} = 1 \text{ and } b_{nt} = 1 \\ 0 & \text{otherwise} \end{array} \right\} \tag{2.8}$$

Let ε_{ntk} $(k = \{0, ..., 3\})$ demographic, psychological factors that determined the precise timing of birth and participation in the labour force that is unobserved by the econometrican and that $(\varepsilon_{nt1}, ..., \varepsilon_{nt1})$ is identically and independently distributed across (n, t) with multivariate probability distribution $F_{14}(\varepsilon_{nt1}, ..., \varepsilon_{nt4})$ ..

The third component in utility is derived from current consumption. We denote by $U_{3nt} \equiv u_3(x_{nt}, z_{nt}, \varepsilon_{5nt})$ the current utility from consumption by household n in period t, and assume $u_2(x_{nt}, z_{nt}, \varepsilon_{5nt})$ is concave increasing in x_{nt} for all values of the observed and unobserved demographic variables $(z_{nt}, \varepsilon_{5nt})$. Analogous to the assumptions made for the other unobserved variables, we assume ε_{5nt} is identically and independently distributed across (n, t) with bivariate probability distribution $F_5(\varepsilon_{5nt})$.

The period t utility for household n utility is found by summing over the three components. Let $\beta \in (0, 1)$ denote the common subjective discount factor, and write $E_t(.)$ as the expectation conditional on information available to household n at period t. The expected lifetime utility of household n is then:

$$E_0 \left\{ \sum_{t=0}^{T} \beta^t \left[\sum_{k=0}^{3} \mathbb{I}_{knt}(U_{0ntk} + U_{1ntk} + U_{3nt} + \varepsilon_{ntk}) \right] \right\} \tag{2.9}$$

Table 1 displays the notation defining the main elements of our model.

3. Optimal Decision Making

In this paper we assume there are no distortions within the labor supply and consumer goods markets. This approach was recently utilized by Altug and Miller (1998) to estimate a life-cycle model of how work experience, affected female wages and labor supply. Indeed, over the last decade an empirical literature has emerged that tests for deviations from Pareto optimal allocations using panel data on households. (See, for example, Altug and Miller (1990), Altonji, Hayashi and Kotilikoff (1995), Cochrane (1991), Mace (1991), Miller and Sieg (1997), and Townsend (1994).) Taken together, this body of work shows that the restrictions imposed by Pareto optimal allocations are quite hard to reject with panel data on households, 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. In addition, the limited empirical work that exists on incomplete markets cannot easily be generalized beyond the highly stylized frameworks that are investigated. So while few economists believe that the real world supports a rich set of Arrow Debreu securities spanning the commodity space, in the absence of clear guidance about precisely how gains from trade are left unfulfilled, the assumption of ignoring such impediments to trade is a useful working hypothesis. In this case the assumption, that observed allocations are Pareto optimal, allows us to derive optimality tractable conditions from a model in which households deal with the complex interactions that arise from spacing births, given the time commitment to their young, while simultaneously determining labor supply in a world where labor force attachment impacts on future wages.

The Pareto optimal allocations are derived as the solution to a social planner's problem for a large population of households n in a cohort defined on the $[0, 1]$ interval, in which the integral of the weighted, expected discounted utilities of each household are maximized subject to an aggregate feasibility or resource constraint. Therefore the objective function for the social planner is formed from the individual utilities defined by Equation 2.9 and the social weights attached to each individual n, which we denote by η_n^{-1}. The planner is constrained by the time available to each household n in the period $t \in \{1, 2, \ldots, T\}$ that cohort is active, which is Equation 2.4, and must respect the time required to nurture children, as indicated by Equation 2.3. At the margin consumption goods are produced the value of marginal product function for labor, which is Equation 2.1, and net transfers between members of the cohort and others (including other income generating family members and public transfers) are exogenously set to W_t in period t. We now state the social planner's constrained optimization problem more formally.

The aggregate feasibility condition equates aggregate consumption at each date t to the sum of output produced by all individuals $n \in [0, 1]$ and the aggregate endowment W_t. Defining $\mathcal{L}$ as the Lebesque measure which integrates over the cohort population, this constraint requires:

$$\int_0^1 (x_{nt} + \pi b_{nt} - w_{nt} h_{nt}) \, d\mathcal{L}(n) \leq W_t \tag{3.1}$$

The Pareto optimal allocations are found by maximizing

$$E_0 \left\{ \int_0^1 \sum_{t=0}^{T} \eta_n^{-1} \beta^t [\sum_{k=0}^{3} \mathbb{I}_{knt}(U_{0ntk} + U_{1ntk} + U_{3nt} + \varepsilon_{ntk})] \right\} d\mathcal{L}(n) \tag{3.2}$$

subject to aggregate budget inequality 3.1 and also the individual household time constraints 2.4 with respect to sequences (of random variables that are successively measurable with respect to the information known at periods $t = 0, 1, \ldots$) for consumption and labor supply $\{x_{nt,} h_{nt}, b_{nt}\}_{t=0}^{T}$ chosen for all the cohort members $n \in [0,1]$. For future reference we denote the optimal choices by $\{x_{nt,}^o h_{nt}^o, b_{nt}^o\}_{t=0}^{T}$ writing $d_{nt}^o = 1$ whenever $h_{nt}^o > 0$.

The necessary conditions characterizing the optimal consumption, labor supply and birth allocations are the basis for the estimation procedure. Turning to the derivation of optimal consumption first, we define $\beta^t \lambda_t$ as the Lagrange multiplier associated with the aggregate feasibility constraint. Differentiating the Lagrangian formed from 3.2 and 3.1, the optimal consumption allocations satisfy the necessary conditions:

$$\frac{\partial u_3(x_{nt,}^o, z_{nt}, \varepsilon_{5nt})}{\partial x_{nt}} = \eta_n \lambda_t, \tag{3.3}$$

for all $n \in [0,1]$ and $t \in \{0, 1, ...\}$. Notice that η_n corresponds to the for each household n. In our empirical work we estimate Equation 3.3 using data on consumption and household demographics to obtain estimates of the social weight by inverting an estimate of the estimated marginal utilities of wealth.

At the heart of the decision maker's trade off between working career and family, is the fact that births and work reduce the amount of time left for leisure time, so they cannot be solved independently. Nevertheless it is logically possible, and pedagogically useful, to analyze the determination of optimal labor supply choices given optimal fertility behavior, and vice versa, rather than tackle the joint problem simultaneously. We first consider life cycle labor supply given the optimal birth choices. Since each person's labor supply contributes only infinitesimally to aggregate output, it is valued at a constant rate each period λ_t by the social planner. Thus we may represent the conditional valuation functions for household n associated with each discrete decision in period t as:

$$V_{jnt} + \varepsilon_{jnt} \equiv \max_{\{h_{nr}(\mathbb{I}_{knr})_{k=0}^{3}\}_{r=t}^{T}} E_t \left[\begin{array}{c} \sum_{r=t}^{T} \beta^{r-t} (\sum_{k=0}^{3} \mathbb{I}_{knt}(U_{0ntk} + U_{1ntk} + \eta_n \lambda_r w_{nr} h_{nr} \\ -\eta_n \lambda_r \pi_k(z_{nr}) b_{nr} + \varepsilon_{ntk})) \mid \mathbb{I}_{jnt} = 1 \end{array} \right] \tag{3.4}$$

$for\ j \in \{0, 1, 2, 3\}$.

Up to the household specific factor of proportionality η_n, the term $V_{knt} + \varepsilon_{knt}$, denotes the social value from n choosing option k at date t, conditional upon all the information available to the social planner (and the household) at the beginning of time t. Whether each individual choice is optimal or not depends on:

$$\mathbb{I}^o_{knt} = \begin{cases} 1, & \text{if } V_{knt} + \varepsilon_{knt} \geqslant V_{jnt} + \varepsilon_{jnt}\ \forall j \neq k \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

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:

$$p_{knt} = E[\mathbb{I}^o_{knt} = 1 \mid z^*_{nt}] \tag{3.6}$$

This definition shows that if a representation for $V_{knt} - V_{jnt}$ can be readily obtained in terms of the variables and parameters that characterize the household's problem, the parameters can be estimated using standard approaches to estimating discrete choice models with labor supply and other demographic data, including data on births. Our estimation approach uses the fact that p_{knt} can be estimated nonparametrically and that V_{knt}'s have the recursive representations:

$$V_{jnt} + \equiv \max_{h_{nt}>0} E_t \left[\begin{array}{c} U_{0ntk} + U_{1ntk} + \eta_n \lambda_t w_{nt} h_{nt} - \eta_n \lambda_t \pi(z_{nt}) b_{nt} \\ +\beta E_t[\sum_{k=0}^{3}\ p_{knt+1}(V_{knt+1} + \varepsilon_{knt+1})] \mid \mathbb{I}_{jnt} = 1 \end{array} \right] \tag{3.7}$$

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 $\partial V_{3nt} \backslash \partial h_{nt} = 0$. Thus if $\mathbb{I}^o_{knt} = 1$ for $k = \{1, 3\}$, then h^o_{nt} solves:

$$\frac{\partial U_{knt}}{\partial h_{nt}} - \eta_n \lambda_t w_{nt} = \beta E_t \left\{ \sum_{k=0}^{3} [\ p_{knt+1} \frac{\partial (V_{knt+1} + \varepsilon_{knt+1})}{\partial h_{nt}} + (V_{knt+1} + \varepsilon_{knt+1})\ \frac{\partial p^h_{n,t+1}}{\partial h_{nt}}] \right\} \tag{3.8}$$

The left side of Equation 3.8 gives the current benefits and costs of a spending a marginal hour working, comprising a utility cost in terms of leisure forgone, 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 ride side of Equation 3.8.

4. An Estimation Strategy

This framework is amenable to a multi-stage estimation strategy. First, there is contemporaneous separability of consumption from labor supply and birth in the utility function. Second, wages are assumed to noisy measures of individual-specific marginal products of labor, which are determined by our two forms of human capital accumulation, namely formal education, past labor market participation and number of hours worked, plus other individual characteristics. Provided the measurement error in wages is uncorrelated with current and past and past labor supply and birth choices (an assumption we can readily set for providing a set of overidnetifying instruments exist), the consumption and wages equations can be estimated separately from the hours, participation and birth equations to provide estimates of the determinants of household consumption and the effects of human capital accumulation on the individual wages.

The representation of individuals' valuation functions defined by (3.4),and (3.8) imply that the fixed cost of participation, benefit of a birth and cost of a birth can be recovered from a model in which the income generated by the decision to work (jointly with the decision to have a birth) is evaluated using the product of shadow value of consumption λ_t and the time-invariant individual-specific effect η_n. However, the existence of fixed costs of participation, birth benefits, birth and the effect of endogenous labor market participation, birth decision, and the optimal choice of hours implies that techniques developed for dynamic discrete choice models must be used to estimate the hours, participation and birth conditions. In principle one could use one of the many maximum likelihood estimation (ML) procedures available(see e.g. Miller(1984), Wolpin(1984), Pakes(1986), Rust(1987), etc.). This however involves the derivation of the valuation function as a mapping of the state and parameter space to calculate the probability of the sample outcome. Our model is very complicated in that it allow nonseparability of both birth and labor supply decision, this would many the computational cost of employing ML in this setting very prohibitive. For this reason, we adapted a conditional choice probability (CCP) estimator, which does not require us to solve the valuation functions.

The CCP estimator forms an alternative representation for the conditional valuation functions that enter individuals' optimizing conditions by multiplying current utilities, evaluated at respective state for a given parameter value and corrected for dynamic selection bias, with the probability that the state in question occurs, and summing over states. The probabilities are estimated nonparametrically and then substituted into a criterion function that is optimized over the structural parameters. Although CCP estimators are far more tractable than ML, the computational burden of estimating conditional choice probabilities at every node in the decision tree is great. We exploit the property of finite state dependence, enjoyed by our model, to use the representation results from Altug and Miller (1998).

4.1. Conditional value functions

The alternative representation for the valuation functions hinges on the observation that the difference between the conditional valuation functions for two discrete outcomes can be expressed as a function of the conditional choice probability of selecting on outcome versus. the other. Define the vector Ψ_{nt} as $\Psi_{nt} \equiv (z_{nt}^{*\prime}, \eta_n \lambda_t w_{nt})'$.By Proposition 1 of Hotz and Miller (1993), there exists a mapping $q : [0,1] \rightarrow R$, which gives the difference between the conditional valuation functions as a function of the conditional choice probabilities, and $\varphi_k : [0,1] \rightarrow R$, which measures the expected value of the unobservables in current utility, conditional on an action $k \in \{0,1,2,3\}$ being taken

$$q_{kj}(p_k(\Psi_{nt})) = V_k(\Psi_{nt}) - V_j(\Psi_{nt}), \tag{4.1}$$

$$\varphi_k(p_k(\Psi_{nt})) = E[\varepsilon_{knt} \mid \Psi_{nt}, I_{knt}^0 = 1], \quad k \in \{0,1,2,3\} \tag{4.2}$$

4.2. Representation for labor supply decision

In our application, the property of finite dependence is exploited to derive an alternative representation for elements of $q_{kj}(\underline{p}(\Psi_{nt}))$. To illustrate this, define the $(\rho + K + M + 1)-$ dimensional vectors $z_{0nt}^{(s)}$ $z_{1nt}^{(s)}$ $z_{2nt}^{(s)}$ and $z_{3nt}^{(s)}$as

$$z_{0nt}^{(s)} \equiv (h_{n,t-\rho+s}, ..., h_{n,t-1}, b_{n,t-M+s}, ..., b_{n,t-1} 0, ..., 0, 0, ..., 0, z_{n,t+s})', \tag{4.3}$$

$$z_{1nt}^{(s)} \equiv (h_{n,t-\rho+s}, ..., h_{n,t-1}, b_{n,t-M+s}, ..., b_{n,t-1}, h_{nt}^{*}, ..., 0, 0, ..., 0, z_{n,t+s})' \tag{4.4}$$

$$z_{2nt}^{(s)} \equiv (h_{n,t-\rho+s}, ..., h_{n,t-1}, b_{n,t-M+s}, ..., b_{n,t-1}, 0, ..., 0, 1, ..., 0, z_{n,t+s})' \tag{4.5}$$

$$z_{3nt}^{(s)} \equiv (h_{n,t-\rho+s}, ..., h_{n,t-1}, b_{n,t-M+s}, ..., b_{n,t-1}, h_{nt}^*, ..., 0, 1, ..., 0, z_{n,t+s})' \qquad (4.6)$$

for $s = 1, ..., \nu$, where h_{nt}^* is the fraction of time a woman chooses to spend at work conditional on participating, $b_{n,t} = (b_{nt}, b_{nt-1}, ..., b_{nt-M})$. Similarly define $\psi_{0nt}^{h(s)} \equiv (z_{0nt}^{(s)}, \mu_n\eta_n)'$ and $\psi_{1nt}^{h(s)} \equiv (z_{1nt}^{(s)}, \mu_n\eta_n)$. The state vector $\psi_{0nt}^{h(s)}$ is the state for a woman at date $t+s$ who has accumulated the work history $(h_{n,t-\nu+s}, ..., h_{n,t-1})'$ up to period t and then chooses not to participate at date t and for $s-1$ periods following period t. The state vector $\psi_{0nt}^{h(\nu)}$ corresponds to the labor market history in which the woman does not participate between t and $t+\nu$. Likewise $\psi_{1nt}^{h(s)}$ is the state vector for a woman at time $t+s$ who accumulates the history $(h_{n,t-\nu+s}, ..., h_{n,t-1})'$ up to time t, chooses to participate at date t, and then chooses not to participate for $s-1$ periods following period t.

Define

$$u_k(\Psi_{nt}) = \begin{cases} u_1(z_{nt}^*, 1) + u_2(z_{nt}^*, 0) & \text{for } k = 0 \\ u_1(z_{nt}^*, 1 - h_{nt}^*) + u_2(z_{nt}^*, 1) + \eta_n\lambda_t w_{nt} h_{nt}^* & \text{for } k = 1 \\ u_1(z_{nt}^*, 1 - c_{nt-1} - \rho_0) + u_2(z_{nt}^*, 0) & \\ +(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt}) - \eta_n\lambda_t\pi(z_{nt}) & \text{for } k = 2 \\ u_1(z_{nt}^*, 1 - c_{nt-1} - \rho_0 - h_{nt}^*) + u_2(z_{nt}^*, 1) - \eta_n\lambda_t\pi(z_{nt}) & \\ +(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt}) + \eta_n\lambda_t w_{nt} h_{nt}^* & \text{for } k = 3 \end{cases}$$

Let$\overline{\rho} \equiv \max(\rho, M)$,Then the condtional valuation functions can be expressed as

$$\begin{aligned} V_k(\Psi_{nt}) &= u_k(\Psi_{nt}) + E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s [u_0(\Psi_{knt}^{(s)}) + \varphi_0(p_0(\Psi_{knt}^{(s)}))]\} \qquad (4.7) \\ & +E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \sum_{j=1}^{3} p_j(\Psi_{knt}^{(s)})[q_{j0}(p_j(\Psi_{knt}^{(s)})) + \varphi_j(p_j(\Psi_{knt}^{(s)})) - \varphi_0(p_0(\Psi_{knt}^{(s)}))]\} \\ & +E_t\{\beta^{\overline{\rho}+1}[V_0(\Psi_{knt}^{(\overline{\rho}+1)}) + \varphi_0(p_0(\Psi_{knt}^{(\overline{\rho}+1)})) + \sum_{j=1}^{3} p_j(\Psi_{knt}^{(\overline{\rho}+1)})[q_{j0}(p_j(\Psi_{knt}^{(\overline{\rho}+1)})) \\ & +\varphi_j(p_j(\Psi_{knt}^{(\overline{\rho}+1)})) - \varphi_0(p_0(\Psi_{knt}^{(\overline{\rho}+1)}))]]\} \end{aligned}$$

This result, which follows directly from Altug and Miller (1998), provides a tractable representation of the difference between the respective valuation functions and Euler equation.

4.3. Hours' Euler Equation

Equation(4.7) also gives a alternative representation for the euler equations for labor supply

$$\begin{aligned} 0 &= -\frac{\partial u_1(z_{nt}, 1-h^*_{nt})}{\partial h_{nt}} - \eta_n \lambda_t w_{nt} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \frac{\partial[u_0(\Psi^{(s)}_{1nt}) + \varphi_0(p_0(\Psi^{(s)}_{1nt}))]}{\partial h_{nt}}\} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \sum_{j=1}^{3} p_j(\Psi^{(s)}_{1nt}) \frac{\partial[q_{j0}(p_j(\Psi^{(s)}_{1nt})) + \varphi_j(p_j(\Psi^{(s)}_{1nt})) - \varphi_0(p_0(\Psi^{(s)}_{1nt}))]}{\partial h_{nt}}\} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \sum_{j=1}^{3} [q_{j0}(p_j(\Psi^{(s)}_{1nt})) + \varphi_j(p_j(\Psi^{(s)}_{1nt})) - \varphi_0(p_0(\Psi^{(s)}_{3nt}))] \frac{\partial p_j(\Psi^{(s)}_{1nt})}{\partial h_{nt}}\} \end{aligned} \tag{4.8}$$

and

$$\begin{aligned} 0 &= -\frac{\partial u_1(z_{nt}, 1-h^*_{nt} - c_{nt-1} - \rho_0)}{\partial h_{nt}} - \eta_n \lambda_t w_{nt} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \frac{\partial[u_0(\Psi^{(s)}_{3nt}) + \varphi_0(p_0(\Psi^{(s)}_{3nt}))]}{\partial h_{nt}}\} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \sum_{j=1}^{3} p_j(\Psi^{(s)}_{3nt}) \frac{\partial[q_{j0}(p_j(\Psi^{(s)}_{3nt})) + \varphi_j(p_j(\Psi^{(s)}_{3nt})) - \varphi_0(p_0(\Psi^{(s)}_{3nt}))]}{\partial h_{nt}}\} \\ &\quad -E_t\{\sum_{s=1}^{\overline{\rho}.} \beta^s \sum_{j=1}^{3} [q_{j0}(p_j(\Psi^{(s)}_{3nt})) + \varphi_j(p_j(\Psi^{(s)}_{3nt})) - \varphi_0(p_0(\Psi^{(s)}_{3nt}))] \frac{\partial p_j(\Psi^{(s)}_{3nt})}{\partial h_{nt}}\} \end{aligned} \tag{4.9}$$

In order to evaluate the terms $\frac{\partial p_j(\Psi^{(s)}_{1nt})}{\partial h_{nt}}$ and $\frac{\partial p_j(\Psi^{(s)}_{3nt})}{\partial h_{nt}}$ which appears in the Euler equation, define

$$f^{(s),j}_{1nt} = f_1(\Psi^{(s)}_{1nt} \mid \mathbb{I}_{jn,t+s} = 1) \tag{4.10}$$

as the probability density function for$\Psi^{(s)}_{1nt}$, conditional on choosing option $j$date $t+s$. Let

$$f^{(s),j}_{nt} = f(\Psi^{(s)}_{1nt}) \tag{4.11}$$

as the related probability density that does not condition on choosing option j in period $t+s$. This both for $s=1,...,\nu$. Denote their derivatives with respect to h_{nt} by $f_{1nt}^{\prime(s)}$ and $f_{nt}^{\prime(s)}$. Then

$$\frac{\partial p_j(\Psi_{1nt}^{(s)})}{\partial h_{nt}} = \left[\frac{f_{1nt}^{\prime(s)}}{f_{1nt}^{(s)}} - \frac{f_{nt}^{\prime(s)}}{f_{nt}^{(s)}}\right] p_j(\Psi_{1nt}^{(s)}) \tag{4.12}$$

A similar structure will be imposed on $\frac{\partial p_j(\Psi_{3nt}^{(s)})}{\partial h_{nt}}$.

5. 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 member; $YKID_{nt}$, the number of children less than six years; $OKID_{nt}$, 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 1.

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 have 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.

6. Wages

We assume the wage rate, or value of marginal product function, Equation(2.1) can be parameterized as:

$$g(d_{nt-\rho}, ..., d_{nt-\rho}, ..., d_{nt-1}, h_{nt-\rho}, ..., h_{nt-1}, z_{nt}) = \omega_t \mu_n \exp[\sum_{s=1}^{\nu}(\delta_{1s} h_{n,t-s} + \delta_{2s} d_{n,t-s}) + z'_{nt} B_3)] \tag{6.1}$$

where μ_n is an unobserved individual-specific effect. We further 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:

$$\widetilde{w}_{nt} = g(\widetilde{A}_{nt}) \exp(\widetilde{\varepsilon}_{nt}) \tag{6.2}$$

where the multiplicative error term in equation (6.2) is conditionally independent over people, the covariates in the wage equation and the labor supply decision. Taking logarithms on both side of Equation (6.2), and then differencing, yields $\triangle \widetilde{\varepsilon}_{nt} = \triangle \ln(\widetilde{w}_{nt}) - \sum_{s=1}^{\nu}(\delta_{1s} \triangle h_{n,t-s} + \delta_{2s} \triangle d_{n,t-s}) - \triangle z'_{nt} B_3$:

$$\begin{aligned} \triangle \widetilde{\varepsilon}_{nt} &= \triangle \ln(\widetilde{w}_{nt}) - \sum_{s=1}^{\nu}(\delta_{1s} \triangle h_{n,t-s} + \delta_{2s} \triangle d_{n,t-s}) - \triangle z'_{nt} B_3 && (6.3) \\ &\equiv \triangle \ln(\widetilde{w}_{nt}) - Z_{nt} \Theta_1 && (6.4) \end{aligned}$$

where $\Theta_1 \equiv (\delta_{11}, ..., \delta_{1\nu}, \delta_{21}, ..., \delta_{2\nu}, B'_3)$ denote the $(2\nu + k)-$ dimensional vector of identifiable parameters, and Z_{nt} is are the covariates.

An instrumental variables estimator was used to estimate Equation (6.3). Defining Y_n and Z_n as:

$$\begin{aligned} Y_n &\equiv (\triangle \ln(\widetilde{w}_{n2}), ..., \triangle \ln(\widetilde{w}_{nT}))' \\ Z_n &\equiv (Z_{n2}, ..., Z_{nT})' \end{aligned}$$

we estimated Θ_1 with:

$$\Theta_1^N = [N^{-1}\sum_{n=1}^{N} Z_n'\widetilde{W}_n^{-1}Z_n]^{-1}[N^{-1}\sum_{n=1}^{N} Z_n'\widetilde{W}_n^{-1}Y_n] \tag{6.5}$$

where $\widetilde{W}_n$ is a consistent estimator of:

$$W_n \equiv E[(Y_n - Z_n'\Theta_1)(Y_n - Z_n'\Theta_1)' \mid Z_n].$$

Assuming the regressors are valid instuments for the wage equation, that is to say $E(\triangle\widetilde{\varepsilon}_{nt}\,|Z_n) = 0$ for each t, then Θ_1^N has the lowest asymptotic covariance within the class of GMM estimators.

Our estimates of the wage equation, displayed in Table 3, are comparable to those reported in Miller and Sanders (1997) for the National Longitudinal Survey for Youth (NLSY) Altug and Miller (1998), also used 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 laged two years is only 0.03. Also the further back the work experience is, the less the timing matters; an extra hour worked two year inthe 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, repespectively, is less that 40%.

Another measure of the effect of past labour 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 $\overline{h}_t$. This 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 for t-ρ 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, who works less that 1000 hours last year does not get this increased in wages, this may be this effect is capturing to the disencouragement effect from the standard job search model, although we do not explicitly model this type of search cost in our model we can pick the lower bound of this effect. This means that not everybody gets the benefit from past job experience, there is a thresshold number of hours of about 1500 for this positive effect to kick in. This will impact fertility behavior even more that if there were all positive benefit from all levels of past hours, since a mother could reduce her hours and still continue to enjoy the benefit high wages. We will come back to this point in the empirical findings section where will have an estimate of the fraction of time a mother spend naturingg her first born.

7. Preferences over Consumption and Wealth Effects

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 This section reports our estimates of the parameters determining the utility from consumption and the marginal utility of wealth parameter, to be used in the labor supply and fertility equations that follow.

7.1. The first order condition

We assume that preferences over consumption take the parametric form:

$$u_3(x_{nt}, z_{nt}, \varepsilon_{5nt}) = \exp(z'_{nt}B_2 + \varepsilon_{5nt})x_{nt}^{\alpha}/\alpha \tag{7.1}$$

where the concavity parameter $\alpha < 1$. Substituting Equation 7.1 into Equation(3.3), taking logarithms and then first differencing yields:

$$(1-\alpha)^{-1} \triangle \varepsilon_{5nt} = \triangle \ln(x_{nt}) - (1-\alpha)^{-1} \triangle z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t) \tag{7.2}$$

Let Θ_2 denote the $(K+T-1)$ dimensional vector of parameters to be estimated, defined:

$$\Theta_2 = \begin{pmatrix} (1-\alpha)^{-1}B_2 \\ (1-\alpha)^{-1}\ln(\lambda_2) \\ \cdots \\ (1-\alpha)^{-1}\ln(\lambda_T) \end{pmatrix}$$

We also define $Y_n = (\triangle \ln(x_{n2}), ..., \triangle \ln(x_{nT}))'$ as a vector of endogenous variables, and Z_n the exogenous variables as

$$Z_n = \begin{bmatrix} \triangle z'_{n2} & D_2 & ... & 0 \\ . & . & . & . \\ \triangle z'_{nT} & 0 & ... & D_T \end{bmatrix}$$

where D_t denotes a time dummy for $t \in \{2, \ldots, T\}$. The assumptions in Section 2 imply the unobserved variable ε_{5nt} is independent of individual specific characteristics. Therefore $E((1-\alpha)^{-1} \triangle \varepsilon_{5nt} \,|z_{nt}) = 0$. Substituting for $(1-\alpha)^{-1} \triangle \varepsilon_{5nt}$ using Equation 7.2 one can obtain a set of orthogonality conditions:

$$E\left[\left(Y_n - Z_n\Theta_2\right) Z_n\right] = 0$$

which can be exploited here to estimate Θ_2 using a similar method to the regression procedures that estimated the wage function.

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 details in the data appendix). The elements of z_{nt} used in the stage of the estimation are defined as FAM_{nt}, $YKID_{nt}$, $OKID_{nt}$, AGE^2_{nt}, NC_{nt} and SO_{nt}. The estimates in Table 4 show that the consumption increases with family size and children consume less than adults, sine 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. Figure IV display the estimated aggregate component of shadow value on consumption, along with its 99% of confidence interval. This shows that these 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.

7.1.1. Individual-specific Effects

Estimation of the labor supply and fertility equations also requires estimates of individual-specific effects $\eta_n\mu_n$ and η_n .There are two approaches that we shall employ for estimating there quantities, the traditional fixed effects estimators and the regression approach of Macurdy (1982), extended by Altug and Miller

(1998) to handle nonlinearities using nonparametric estimation techniques. The traditional estimators are simple to simple to compute. They are however, subject to small sample bias arising from short panel length, although the limited Monte Carlo evidence provided in Hotz and Miller (1988) suggests that small sample bias might not greatly affect the estimates of the other parameters. On the other hand the nonparametric estimator achieves consistency of the cross section of the panel data set but can only deal with any unobserved permanent characteristic that is a mapping of observed random variables.

The traditional fixed effect are estimated as follows. Let T_1 denote the number of time periods for which the wage equation is estimated, and T_2 be the number of time periods for which the marginal utility of consumption equation is estimated.

Let:

$$\begin{aligned}\phi_{1n} \equiv & \sum_{t=1}^{T_1}\left[\ln(w_{nt}) - \sum\nolimits_{s=1}^{\nu}(\delta_{1s}h_{n,t-s} + \delta_{2s}d_{n,t-s}) - z'_{nt}B_3\right]/T_1 \\ & + \sum_{t=1}^{T_1}[\ln(x_{nt}) - (1-\alpha)^{-1}z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t)]/T_2 \end{aligned} \tag{7.3}$$

and

$$\phi_{2n} \equiv \sum_{t\in T_1}[\ln(x_{nt}) - (1-\alpha)^{-1}z'_{nt}B_2 + (1-\alpha)^{-1}\ln(\lambda_t)]/T_2 \tag{7.4}$$

Then estimates of $\eta_n\mu_n$ and η_n are then estimated by simple time averages of the estimated residuals of the consumption and wage equations.

Suppose that $\eta_n\mu_n$ and η_n can be expressed as functions of a Q-dimensional vector of regressors z_n, which is assumed to represent the permanent characteristic of individual n. Let the vector z_n be observed and satisfies the conditions, $E[z_n(\phi_{1n} - \eta_n\mu_n)] = 0$ and $E[z_n(\phi_{2n} - \eta_n)] = 0$. Here z_n could include such observed observable demographic characteristics as religion, marital status, the age distribution of children, home ownership, educational level and geographical location. Let δ_N denote the bandwidth of the proposed kernel estimator and J the normal kernel on R^Q. Then our estimators are

$$\eta_n^N\mu_n^N = \frac{\sum_{m=1,m\neq n}^{N}\phi_{1m}J[\delta_N^{-1}(z_m - z_n)]}{\sum_{m=1,m\neq n}^{N}J[\delta_N^{-1}(z_m - z_n)]} \tag{7.5}$$

and

$$\eta_n^N = \frac{\sum_{m=1,m\neq n}^{N} \phi_{2m} J[\delta_N^{-1}(z_m - z_n)]}{\sum_{m=1,m\neq n}^{N} J[\delta_N^{-1}(z_m - z_n)]} \tag{7.6}$$

The distribution of the estimated fixed-effects estimates for, the wage equation, the consumption equation and the combined individual effects and time effects of the consumption, are displayed in figures I,II,III and IV respectively. The most strictly features of these plots is that the nonparametric estimates are significantly more smooth than the traditional fixed effects. The range of the distribution of the shadow price of consumption all lies above zero, while only a very small portion of the estimated (inverse)social weights lies below zero in the traditional case.

8. Participation, Hours and Birth

8.1. A Parametrization

It is now assumed that $\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt}$ and ε_{3nt} are identically and independently distributed over (n, t) as Type I Extreme Value random variables. This assumption implies that:

$$q_{jk}(\underline{p}) = \ln\left(\frac{p_j(\Psi_{nt})}{p_k(\Psi_{nt})}\right) \tag{8.1}$$

and

$$\varphi_j(p_j(\Psi_{nt})) = \zeta - \ln(p_j(\Psi_{nt})) \tag{8.2}$$

Which implies the

$$q_{jk}(\underline{p}) = -[\varphi_j((p_j(\Psi_{nt})) - \varphi_k((p_k(\Psi_{nt}))] \tag{8.3}$$

The remaining components of the utility function u_1 and u_2 are parametrized as:

$$u_2(z_{nt}^*, d_{nt}) = d_{nt} z_{nt}' B_0$$

and

$$u_1(z_{nt}^*, l_{nt}) = z_{nt}' B_1 l_{nt} + \sum_{s=0}^{\rho} \delta_s l_{nt} l_{nt-s}$$

In these expressions, B_0 are parameters that characterize the fixed-costs of participating in the work force, B_{11} shows the effect of exogenous time-varying characteristics on the marginal disutility of work and B_{12} shows the effect of exogenous time-varying characteristics on the marginal disutility of time spent nuturing children. Preferences are concave decreasing in time spent at work if $z'_{nt}B_{11} + 2\theta_0 l_{nt} + \sum_{s=1}^{\nu} \theta_s l_{nt-s} < 0$ and $\theta_0 < 0$. The parameters θ_s for $s = 1, ..., \nu$ capture intertemporal non-separabilities in preferences with respect to labor supply choices. A value of $\theta_s < 0$ for $s = 1, ..., \nu$ means that hours worked s periods ago increases the marginal disutility of work, and results in less work today. Equivalently, a finding of $\theta_s < 0$ implies that current and past leisure time are substitutes where as $\theta_s > 0$ implies that current and past leisure time are complements. This parametrization also shows that all the terms in equation(**??**) and (4.7) that involve choosing not to participate in future periods are zero.

The distributional assumption for the idiosyncratic stocks and the parameterization of the utility function implies that the Euler equation can be written as

$$0 = \eta_n \lambda_t w_{nt} - z'_{nt}B_1 - 2\delta_0(1 - h_{nt}) - \sum_{s=1}^{\rho} \delta_s l_{nt-s} + E_t[\sum_{s=1}^{\overline{\rho}} \beta^s [p_0(\Psi_{1nt}^{(s)})^{-1} \frac{\partial p_0(\Psi_{1nt}^{(s)})}{\partial h_{nt}}]] \tag{8.4}$$

$$0 = \eta_n \lambda_t w_{nt} - z'_{nt}B_1 - 2\delta_0(1 - h_{nt} - c_{nt-1} - \rho_0) - \sum_{s=1}^{\rho} \delta_s l_{nt-s} + E_t[\sum_{s=1}^{\overline{\rho}} \beta^s [p_0(\Psi_{3nt}^{(s)})^{-1} \frac{\partial p_0(\Psi_{3nt}^{(s)})}{\partial h_{nt}}]] \tag{8.5}$$

The difference in valuation functions associated with $q_{10}(\underline{p})$, $q_{20}(\underline{p})$, $q_{30}(\underline{p})$, $q_{12}(\underline{p})$, $q_{13}(\underline{p})$ and $q_{23}(\underline{p})$ give rise to the following conditions:

$$\begin{aligned} 0 &= \ln(\frac{p_{1nt}}{P_{0nt}}) - z'_{nt}B_0 + z'_{nt}B_1 h_{nt} - \delta_0(h_{nt}^2 - 2h_{nt}) + \sum_{s=1}^{\rho} \delta_s h_{nt}^* l_{nt-s} \\ &\quad - \eta_n \lambda_t w_{nt} h_{nt}^* - E_t[\sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0(\Psi_{0nt}^{(s)})}{p_0(\Psi_{1nt}^{(s)})}\right)] \end{aligned}$$

$$\begin{aligned} 0 &= \ln(\frac{p_{2nt}}{P_{0nt}}) + z'_{nt}B_1(c_{nt-1} + \rho_0) - \delta_0[(c_{nt-1} + \rho_0)^2 - 2(c_{nt-1} + \rho_0)] + \sum_{s=1}^{\rho} \delta_s (c_{nt-1} + \rho_0) l_{nt-s} \\ &\quad -(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N_{nt}^2) + \eta_n \lambda_t \pi(z_{nt}) - E_t[\sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0(\Psi_{0nt}^{(s)})}{p_0(\Psi_{2nt}^{(s)})}\right)] \end{aligned}$$

$$
\begin{aligned}
0 &= \ln(\frac{p_{3nt}}{P_{0nt}}) - z'_{nt}B_0 + z'_{nt}B_1(c_{nt-1}+\rho_0+h^*_{nt}) + \delta_0[(c_{nt-1}+\rho_0+h^*_{nt})^2 - \\
&\quad 2(c_{nt-1}+\rho_0+h^*_{nt})] + \sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0+h^*_{nt})l_{nt-s} - \eta_n\lambda_t w_{nt}h_{nt} + \eta_n\lambda_t\pi(z_{nt}) \\
&\quad -(\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) - E_t[\sum_{s=1}^{\overline{\rho}}\beta^s \ln\left(\frac{p_0(\Psi^{(s)}_{0nt})}{p_0(\Psi^{(s)}_{3nt})}\right)]
\end{aligned}
$$

$$
\begin{aligned}
0 &= \ln(\frac{p_{1nt}}{p_{2nt}}) - z'_{nt}B_0 + z'_{nt}B_1(c_{nt-1}+\rho_0-h^*_{nt}) - \delta_0[(c^2_{nt-1}+\rho_0^2+h^{*2}_{nt} - 2\rho_0 c_{nt-1} \\
&\quad +2c_{nt-1}+2\rho_0-2h^*_{nt})] + \sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0-h^*_{nt})l_{nt-s} - \eta_n\lambda_t w_{nt}h_{nt} \\
&\quad +(\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) - \eta_n\lambda_t\pi(z_{nt}) - E_t[\sum_{s=1}^{\overline{\rho}}\beta^s \ln\left(\frac{p_0(\Psi^{(s)}_{2nt})}{p_0(\Psi^{(s)}_{1nt})}\right)]
\end{aligned}
$$

$$
\begin{aligned}
0 &= \ln(\frac{p_{1nt}}{p_{3nt}}) - z'_{nt}B_1(c_{nt-1}+\rho_0) - \delta_0[(c_{nt-1}+\rho_0.)^2 - 2(1-h_{nt})(c_{nt-1}+\rho_0)] \\
&\quad -\sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0)l_{nt-s} + (\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) \\
&\quad -\eta_n\lambda_t\pi(z_{nt}) - E_t[\sum_{s=1}^{\overline{\rho}}\beta^s \ln\left(\frac{p_0(\Psi^{(s)}_{3nt})}{p_0(\Psi^{(s)}_{1nt})}\right)]
\end{aligned}
$$

$$
\begin{aligned}
0 &= \ln(\frac{p_{2nt}}{p_{3nt}}) + z'_{nt}B_0 - z'_{nt}B_1 h_{nt} - \delta_0[2h_{nt}(1-c_{nt-1}-\rho_0) - h^2_{nt}] \\
&\quad -\sum_{s=1}^{\rho}\delta_s h_{nt}l_{nt-s} + \eta_n\lambda_t w_{nt}h_{nt} - E_t[\sum_{s=1}^{\overline{\rho}}\beta^s \ln\left(\frac{p_0(\Psi^{(s)}_{3nt})}{p_0(\Psi^{(s)}_{2nt})}\right)]
\end{aligned}
$$

8.2. Nonparametric estimation of the Conditional Choice Probabilities

Then the probabilities p_{knt} can be computed as nonlinear regressions of participation index $\mathbb{I}_{knt}$ on the current state Ψ^N_{nt}, where the N superscript denotes an estimated quantity. Define $K[\delta_N^{-1}(\Psi_{nt}^N - \Psi_{mr}^N)]$ as a given kernel , where δ_N is the bandwidth associated with each argument. Then the nonparametric estimate, p_{knt}^N , is then computed using kernel estimator

$$p_{knt}^N \equiv \frac{\sum_{m=1,m\neq n}^N \sum_{r=1,r\neq t}^T \mathbb{I}_{kmr} K[\delta_N^{-1}(\Psi_{nt}^N - \Psi_{mr}^N)]}{\sum_{m=1,m\neq n}^N \sum_{r=1,r\neq t}^T K[\delta_N^{-1}(\Psi_{nt}^N - \Psi_{mr}^N)]} \tag{8.6}$$

The conditional choice probabilities $p_j(\Psi_{knt}^{(s)})$ are also estimated as nonlinear regressions of a choice index on the appropriate state variables. Define the variable

$$\mathbb{I}_{knt}^{(s)} = \mathbb{I}_{kn,t-s} \prod_{\ell=1}^{s-1} (1 - \mathbb{I}_{kn,t-\ell}), \quad k \in \{0,1,2,3\} \tag{8.7}$$

Notice that $\mathbb{I}_{knt}^{(s)} = 1$ if the person choose option k at $t-s$ but then did not option k for$s-1$ periods. Thus , $\mathbb{I}_{knt}^{(s)}$is an index variable that allows us to condition on the behavior of individuals with the labor market and birth histories defined by $z_{knt}^{(s)}$. The conditional choice probabilities $p_j^N(\Psi_{knt}^{(s)})$ are computed

$$p_j^N(\Psi_{knt}^{(s)}) \equiv \frac{\sum_{m=1,m\neq n}^N \sum_{r=1,r\neq t}^T \mathbb{I}_{jmr} \mathbb{I}_{kmr}^{(s)} K[\delta_N^{-1}(\Psi_{knt}^{(s)N} - \Psi_{kmr}^{(s)N})]}{\sum_{m=1,m\neq n}^N \sum_{r=1}^T \mathbb{I}_{kmr}^{(s)} K[\delta_N^{-1}(\Psi_{knt}^{(s)N} - \Psi_{kmr}^{(s)N})]} \tag{8.8}$$

where $\Psi_{knt}^{(s)N} \equiv (z_{knt}^{(s)\prime}, \mu_n^N \eta_n^N)'$ for $k \in \{0,1,2,3\}$ is the state vector for individual n.

8.3. Hours, Participation and Birth Conditions

So combining the estimate we obtain in section 7.1.1 gives

$$\begin{aligned} \xi_{nt}^N &= (1-\alpha)^{-1} \ln(\lambda_t \eta_n)^N \\ &\equiv (1-\alpha)^{-1} \ln(\eta_n)^N + (1-\alpha)^{-1} \triangle \ln(\lambda_t).^N \end{aligned} \tag{8.9}$$

Note that we can obtain an estimate of $\eta_n \lambda_t$ and η_n respectively as:

$$\eta_n^N \lambda_t^N \equiv \exp((1-\alpha)\xi_{nt}^N) \tag{8.10}$$

So combining the estimate we obtain in section 7.1.1 gives

$$
\begin{aligned}
\xi_{nt}^{N} &= (1-\alpha)^{-1}\ln(\lambda_t\eta_n)^N \\
&\equiv (1-\alpha)^{-1}\ln(\eta_n)^N + (1-\alpha)^{-1}\triangle\ln(\lambda_t).^N
\end{aligned} \tag{8.11}
$$

Note that we can obtain an estimate of $\eta_n\lambda_t$ and η_n respectively as:

$$
\eta_n^N\lambda_t^N \equiv \exp((1-\alpha)\xi_{nt}^N) \tag{8.12}
$$

So we have

$$
\begin{aligned}
D_{1nt}^N &= \exp((1-\alpha)\xi_{nt}^N)w_{nt} + \sum_{s=1}^{\overline{\rho}}\beta^s p_0^N(\Psi_{1nt}^{(s)})^{-1}\left[\frac{f_{1nt}^{\prime 1(s),N}}{f_{1nt}^{1(s),N}} - \frac{f_{nt}^{\prime 1(s)N}}{f_{nt}^{1(s),N}}\right]p_0^N(\Psi_{1nt}^{(s)}) \\
&\quad -z_{nt}^{\prime}B_{11} - 2\delta_0(1-h_{nt}) - \sum_{s=1}^{\rho}\delta_s l_{nt-s}
\end{aligned} \tag{8.13}
$$

$$
\begin{aligned}
D_{2nt}^N &= \exp((1-\alpha)\xi_{nt}^N)w_{nt} + \sum_{s=1}^{\overline{\rho}}\beta^s p_0^N(\Psi_{3nt}^{(s)})^{-1}\left[\frac{f_{1nt}^{\prime 3(s),N}}{f_{1nt}^{3(s),N}} - \frac{f_{nt}^{\prime 3(s)N}}{f_{nt}^{3(s),N}}\right]p_0^N(\Psi_{3nt}^{(s)}) \\
&\quad -z_{nt}^{\prime}B_{11} - 2\delta_0(1-h_{nt}-c_{nt-1}-\rho_0) - \sum_{s=1}^{\rho}\delta_s l_{nt-s}
\end{aligned} \tag{8.14}
$$

$$
\begin{aligned}
D_{3nt}^N &= \ln(\frac{p_{1nt}^N}{p_{0nt}^N}) - \exp((1-\alpha)\xi_{1nt}^N)w_{nt}h_{nt}^* - \sum_{s=1}^{\overline{\rho}}\beta^s\ln\left(\frac{p_0^N(\Psi_{0nt}^{(s)})}{p_0^N(\Psi_{1nt}^{(s)})}\right) \\
&\quad -z_{nt}^{\prime}B_0 + z_{nt}^{\prime}B_1 h_{nt} - \delta_0(h_{nt}^2 - 2h_{nt}) + \sum_{s=1}^{\rho}\delta_s h_{nt}^* l_{nt-s}
\end{aligned} \tag{8.15}
$$

$$
\begin{aligned}
D_{4nt}^N &= \ln(\frac{p_{2nt}^N}{p_{0nt}^N}) + \exp((1-\alpha)\xi_{nt}^N)\pi(z_{nt}) - \sum_{s=1}^{\overline{\rho}}\beta^s\ln\left(\frac{p_0^N(\Psi_{0nt}^{(s)})}{p_0^N(\Psi_{2nt}^{(s)})}\right) \\
&\quad +z_{nt}^{\prime}B_1(c_{nt-1}+\rho_0) - \delta_0[(c_{nt-1}+\rho_0)^2 - 2(c_{nt-1}+\rho_0)] \\
&\quad +\sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0)l_{nt-s} - (\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N_{nt}^2)
\end{aligned} \tag{8.16}
$$

$$D^N_{5nt} = \ln(\frac{p^N_{3nt}}{p^N_{0nt}}) - \exp((1-\alpha)\xi^N_{1nt}) w_{nt} h^*_{nt} + \exp((1-\alpha)\xi^N_{nt})\pi(z_{nt}) - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p^N_0(\Psi^{(s)}_{0nt})}{p^N_0(\Psi^{(s)}_{3nt})}\right)$$
$$-z'_{nt}B_0 + z'_{nt}B_1(c_{nt-1}+\rho_0+h^*_{nt}) + \delta_0[(c_{nt-1}+\rho_0+h^*_{nt})^2 - 2(c_{nt-1}+\rho_0+h^*_{nt})]$$
$$+\sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0+h^*_{nt})l_{nt-s} - (\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) \qquad (8.17)$$

$$D^N_{6nt} = \ln(\frac{p^N_{1nt}}{p^N_{2nt}}) - \exp((1-\alpha)\xi^N_{1nt}) w_{nt} h^*_{nt} - \exp((1-\alpha)\xi^N_{nt})\pi(z_{nt}) - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p^N_0(\Psi^{(s)}_{2nt})}{p^N_0(\Psi^{(s)}_{1nt})}\right)$$
$$+z'_{nt}B_1(c_{nt-1}+\rho_0-h^*_{nt}) - \delta_0[(c^2_{nt-1}+\rho^2_0+h^{*2}_{nt} - 2\rho_0 c_{nt-1} + 2c_{nt-1} + 2\rho_0 - 2h^*_{nt})]$$
$$-z'_{nt}B_0 + \sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0-h^*_{nt})l_{nt-s} + (\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) \quad (8.18)$$

$$D^N_{7nt} = \ln(\frac{p^N_{1nt}}{p^N_{3nt}}) - \exp((1-\alpha)\xi^N_{nt})\pi(z_{nt}) - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0(\Psi^{(s)}_{3nt})}{p_0(\Psi^{(s)}_{1nt})}\right)$$
$$-z'_{nt}B_1(c_{nt-1}+\rho_0) - \delta_0[(c_{nt-1}+\rho_0.)^2 - 2(1-h_{nt})(c_{nt-1}+\rho_0)]$$
$$-\sum_{s=1}^{\rho}\delta_s(c_{nt-1}+\rho_0)l_{nt-s} + (\gamma_0 + \sum_{k=1}^{M}\gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) \quad (8.19)$$

$$D^N_{8nt} = \ln(\frac{p^N_{2nt}}{p^N_{3nt}}) + \exp((1-\alpha)\xi^N_{1nt}) w_{nt} h^*_{nt} - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p^N_0(\Psi^{(s)}_{3nt})}{p^N_0(\Psi^{(s)}_{2nt})}\right)$$
$$+z'_{nt}B_0 - z'_{nt}B_1 h_{nt} - \delta_0[2h_{nt}(1-c_{nt-1}-\rho_0) - h^2_{nt}] - \sum_{s=1}^{\rho}\delta_s h_{nt} l_{nt-s} \quad (8.20)$$

Since $(1-\alpha)$ and the ρ_k's enters nonlinearly into our moment functions with all other parameters that are to be estimated entering linearly, we will use a 2-step estimation procedure which will be computationally more convenient. Let us define as follows

$$\Theta_3 \equiv (B'_0, B'_{11}, \delta_0, ..., \delta_\rho . \gamma_0, ..., \gamma_M, \phi_1, , \phi_2, \pi).$$

and

$$\Gamma \equiv (\alpha, \rho_0, ..., \rho_M)$$

Then define the $(2K+\rho+M+3)-$dimensional vectors X^N_{knt} $k \in \{1,...,8\}$ conditional on Γ by

$$\begin{aligned}
X^N_{1nt} &\equiv (0, z'_{nt},\ 2(1-h_{nt}), l_{nt-1}, ..., l_{nt-\rho}, 0,0,0,0) \\
X^N_{2nt} &\equiv (0, z'_{nt},\ 2(1-h_{nt}-c_{nt-1}-\rho_0), l_{nt-1}, ..., l_{nt-\rho}, 0,0,0,0) \\
X^N_{3nt} &\equiv (z'_{nt},\ \text{-}z'_{nt}h_{nt}, (h^2_{nt}-2h_{nt}),\ \text{-}h_{nt}l_{nt-1}, ..., \text{-}h_{nt}l_{nt-\rho},\ 0,\ 0,0,0) \\
X^N_{4nt} &\equiv (0, -z'_{nt}(c_{nt-1}+\rho_0), (c_{nt-1}+\rho_0)^2 - 2(c_{nt-1}+\rho_0), -(c_{nt-1}+\rho_0)l_{nt-1} \\
&\quad , ..., \text{-}(c_{nt-1}+\rho_0)l_{nt-\rho}, 1, b_{nt-1}, ..., b_{nt-M}, N_{nt}, -\exp((1-\alpha)\xi^N_{nt})\ ) \\
X^N_{5nt} &\equiv (z'_{nt}, -z'_{nt}(c_{nt-1}+\rho_0+h^*_{nt}), -(c_{nt-1}+\rho_0+h^*_{nt})^2 + 2(c_{nt-1}+\rho_0+h^*_{nt}), \\
&\quad \text{-}(c_{nt-1}+\rho_0+h^*_{nt})l_{nt-1}, ..., \text{-}(c_{nt-1}+\rho_0+h^*_{nt})l_{nt-\rho},\ 1, b_{nt-1}, ..., \\
&\quad b_{nt-M}, N_{nt}, -\exp((1-\alpha)\xi^N_{nt})\ )) \\
X^N_{6nt} &\equiv (z'_{nt},\ \text{-}z'_{nt}(c_{nt-1}+\rho_0-h^*_{nt}), (c^2_{nt-1}+\rho^2_0+h^{*2}_{nt} - 2\rho_0 c_{nt-1} + 2c_{nt-1} + 2\rho_0 - 2h^*_{nt}), \\
&\quad \text{-}(c_{nt-1}+\rho_0-h^*_{nt})l_{nt-1}, ..., \text{-}(c_{nt-1}+\rho_0-h^*_{nt})l_{nt-\rho},\ \text{-}1, -b_{nt-1}, ..., -b_{nt-M}, -N_{nt}, \\
&\quad \exp((1-\alpha)\xi^N_{nt})\ )) \\
X^N_{7nt} &\equiv (0, z'_{nt}(c_{nt-1}+\rho_0), (c_{nt-1}+\rho_0.)^2 - 2(1-h_{nt})(c_{nt-1}+\rho_0),\ (c_{nt-1}+\rho_0)l_{nt-1}, ..., \\
&\quad (c_{nt-1}+\rho_0)l_{nt-\rho},\ \text{-}1, -b_{nt-1}, ..., -b_{nt-M}, -N_{nt}, \exp((1-\alpha)\xi^N_{nt})\ )) \\
X^N_{8nt} &\equiv (-z'_{nt}, z'_{nt}h_{nt}, [2h_{nt}(1-c_{nt-1}-\rho_0) - h^2_{nt}], h^*_{nt}l_{nt-1}, ..., h^*_{nt}l_{nt-\rho}, 0,0,0,0)
\end{aligned}$$

Also define 8×1 vector $y^N_{nt} \equiv (y^N_{1nt}, ..., y^N_{8nt})'$ where

$$y^N_{1nt} \equiv \exp((1-\alpha)\xi^N_{nt})w_{nt} + \sum_{s=1}^{\overline{\rho}} \beta^s p^N_0(\Psi^{(s)}_{1nt})^{-1} \left[\frac{f'^{1(s),N}_{1nt}}{f^{1(s),N}_{1nt}} - \frac{f'^{1(s)N}_{nt}}{f^{1(s),N}_{nt}}\right] p^N_0(\Psi^{(s)}_{1nt})$$

$$y^N_{2nt} \equiv \exp((1-\alpha)\xi^N_{nt})w_{nt} + \sum_{s=1}^{\overline{\rho}} \beta^s p^N_0(\Psi^{(s)}_{3nt})^{-1} \left[\frac{f'^{3(s),N}_{1nt}}{f^{3(s),N}_{1nt}} - \frac{f'^{3(s)N}_{nt}}{f^{3(s),N}_{nt}}\right] p^N_0(\Psi^{(s)}_{3nt})$$

$$y^N_{3nt} \equiv \ln(\frac{p^N_{1nt}}{p^N_{0nt}}) - \exp((1-\alpha)\xi^N_{1nt})w_{nt}h^*_{nt} - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p^N_0(\Psi^{(s)}_{0nt})}{p^N_0(\Psi^{(s)}_{1nt})}\right)$$

$$y^N_{4nt} \equiv \ln(\frac{p^N_{2nt}}{p^N_{0nt}}) - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p^N_0(\Psi^{(s)}_{0nt})}{p^N_0(\Psi^{(s)}_{2nt})}\right)$$

$$y_{5nt}^N \equiv \ln(\frac{p_{3nt}^N}{p_{0nt}^N}) - \exp((1-\alpha)\xi_{1nt}^N)w_{nt}h_{nt}^* - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0^N(\Psi_{0nt}^{(s)})}{p_0^N(\Psi_{3nt}^{(s)})}\right)$$

$$y_{6nt}^N \equiv \ln(\frac{p_{1nt}^N}{p_{2nt}^N}) - \exp((1-\alpha)\xi_{1nt}^N)w_{nt}h_{nt}^* - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0^N(\Psi_{2nt}^{(s)})}{p_0^N(\Psi_{1nt}^{(s)})}\right)$$

$$y_{7nt}^N \equiv \ln(\frac{p_{1nt}^N}{p_{3nt}^N}) - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0(\Psi_{3nt}^{(s)})}{p_0(\Psi_{1nt}^{(s)})}\right)$$

$$y_{8nt}^N \equiv \ln(\frac{p_{2nt}^N}{p_{3nt}^N}) + \exp((1-\alpha)\xi_{1nt}^N)w_{nt}h_{nt}^* - \sum_{s=1}^{\overline{\rho}} \beta^s \ln\left(\frac{p_0^N(\Psi_{3nt}^{(s)})}{p_0^N(\Psi_{2nt}^{(s)})}\right)$$

Using this notation, the idiosyncratic errors associated with Euler,choice equations can be written as

$$D_{int}^N = y_{int}^N - X_{int}^{N\prime}\Theta_3 \qquad i = 1, \rightarrow 8. \tag{8.21}$$

Let $D_{nt}^N \equiv (D_{1nt}^N, ..., D_{8nt}^N)'$ and let T_3 denote the set of periods for which the hours, participation and births conditions are valid. Define the vector $D_n^{N\prime} \equiv (D_{n1}^N, ..., D_{nT_3}^N)$. Similarly define $X_{nt}^N \equiv (X_{1nt}^{N\prime}, ..., X_{8nt}^{N\prime})'$, $X_n^N \equiv (X_{n1}^{\alpha\prime}, ..., X_{nT_3}^{\alpha\prime})'$ and $\Phi_n^N \equiv E_t[D_n^N D_n^{N\prime}]$. Then the optimal instrumental variables estimator for Θ_3 conditional on Γsatisfies

$$\Theta_3^N \equiv [\frac{1}{N}\sum_{n=1}^{N} X_n^{N\prime}(\Phi_n^N)^{-1}X_n^N]^{-1}[\frac{1}{N}\sum_{n=1}^{N} X_n^{N\prime}(\Phi_n^N)^{-1}y_n^N] \tag{8.22}$$

Step-2: We next define a MD estimator for Γ as

$$\Gamma^N = \arg\min_{\Gamma} \left[\frac{1}{N}\sum_{n=1}^{N}\widetilde{D_n^N}\right] A_N \left[\frac{1}{N}\sum_{n=1}^{N}\widetilde{D_n^N}\right]' \tag{8.23}$$

where

$$\widetilde{D_{int}^N} = y_{int}^{N\alpha} - \Psi_{int}^{\alpha\prime}\Theta_3^N \qquad i = 1, \rightarrow 8., t = 1, \rightarrow T_3 \tag{8.24}$$

Then $\widetilde{D_n^N}$ is formed similar to $D_n^{N\prime}$above and A_N is a weight matrix appropriately chosen.

8.4. Empirical Findings

Tables V, VI and VII contains estimates of alternative estimators of our participation cost, utiliy of leisure and birth equation. Column (1) reports estimates of the birth preference and cost parameters that are based on nonparametric estimates of individual effects η_n and $\mu_n \eta_n$, while estimates in column (2) are based on the standard effects estimators of η_n and $\mu_n \eta_n$. The significant feature of our results are the strictly similarities between both sets of estimates when they are statistically significant. These results are very seminal in that there is not a lot of literature to compare our results to. Most of results on labour markets particiapation are similar to other studies , however our are more precisey estimated and and significant sign for for some effects that other studies could not estimate. However the major contribution of this work in the birthor/ fertility component.

First, we estimate a sequence of parameters that characterizes the optimal timing of births and a parameter ρ_0, of how this relates to female labor supply over the life-cycle. We concurred with the classical literature that children are good and not bad, since we find a positive utility up to the 6th birth. The parameter on the timing of birth for example, would imply that the optimal space of a two children family would be 3 to 4 years apart. So, having kids too close or too far apart are less desirable. Turning to the cost of a child, we find that both sets of estimates give similar results. There is a positive cost discounted life-time cost to having a child, we find that having at least a high school education significantly increases that cost. After controlling for education, we find that Hispanic have a significantly higher cost than white for both estimates. Black also seems to have a higher cost than white, but this was only significant when we use the non parametric estimators. 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. The human component , give estimates of ρ_0 of between .15 and 0.08 , coupled with the nonlinearity effect found in the wage equation on past work means there is a significant wage cost which is imposed in order to have a child.

9. Policy Simulation

In this section we used the estimates from our model to solve the individual's optimization problem with different parameter values corresponding to different policy proposals. We first outline the model at the individual level which will be used to do this exercise, next we outline the solution method used. Then we outline the numerical specifics that we used to compute our solution and present the different policy questions and answers.

9.1. Optimization

Therefore can rewrite the perperiod utility to depend on $k \in \{0,1,2,3\}$. Let

$$U_{nt0}(z^*_{nt}, x_{nt}) = z'_{nt}B_1 + \sum_{s=0}^{\rho} \delta_s l_{nt-s} + \varepsilon_{0nt} \tag{9.1}$$

$$U_{nt1}(z^*_{nt}, x_{nt}) = z'_{nt}B_0 + z'_{nt}B_{11}(1-h_{nt}) + \sum_{s=0}^{\nu} \delta_s(1-h_{nt})l_{nt-s} + \eta_n\lambda_t w_{nt}h^*_{nt} \ + \varepsilon_{1nt} \tag{9.2}$$

$$\begin{aligned} U_{nt2}(z_{nt}, x_{nt}, h_{nt}) &= z'_{nt}B_{11}(1-c_{nt_{-1}}-\rho_0) + \sum_{s=0}^{\nu} \delta_s(1-c_{nt_{-1}}-\rho_0)l_{nt-s} && (9.3) \\ &+(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) - \eta_n\lambda_t\pi(z_{nt}) + \varepsilon_{2nt} && (9.4) \end{aligned}$$

$$\begin{aligned} U_{nt3}(z_{nt}, x_{nt}, h_{nt}) &= z'_{nt}B_0 + z'_{nt}B_{11}(1-c_{nt_{-1}}-\rho_0-h_{nt}) + \sum_{s=0}^{\nu} {}_s\delta_s(1-c_{nt_{-1}}-\rho_0-h_{nt})h_{nt-s} \\ & \eta_n\lambda_t w_{nt}h^*_{nt} + +(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k} + \phi_1 N_{nt} + \phi_2 N^2_{nt}) - \eta_n\lambda_t\pi(z_{nt}) + \varepsilon_{3nt} && (9.5) \end{aligned}$$

Then let

$$Z^* \equiv [z'_{nt}, h_{nt}, ..., h_{nt+1-\rho}, b_{nt}, ..., b_{nt+1-M}, \mu_n\eta_n] \tag{9.6}$$

and

$$\Pr[Z^*_{nt+1} = Z^* \mid Z^*_{nt}, x_{nt}, h_{nt}, \mathbb{I}_{ntk} = 1] \equiv F_{ntk}(Z^{**} \mid Z^*_{nt}, x_{nt}, h_{nt}). \tag{9.7}$$

Let x^o_{nt}, h^o_{nt} and $\mathbb{I}^0_{ntk}$ be the optimal action conditional on the current state Z^*_{nt}. Then we can recast the problem recursively. Th that end, the value function,

$V_{nt}(Z^*)$ is defined for each $(t, Z^*) \in \{0, ..., T\} \times \mathbb{Z}^*$ by substituting the optimal decision rule back into the expected lifetime utility function:

$$V_{nt}(Z^*) \equiv E_t[\sum_{s=t}^{T} \sum_{k=0}^{3} \mathbb{I}^o_{nsk} \beta^{t-s} U_{nsk}(z_{ns}, x^o_{ns}, h^o_{ns})] \tag{9.8}$$

For notational convinence define the reduced form utilities

$$U_{ntk}(Z^*) \equiv U_{ntk}(x^o_{knt}, h^o_{knt}, Z^*) \tag{9.9}$$

and the transition probabilities

$$F_{ntk}(Z^{**} \mid Z^*_{nt}) \equiv F_{ntk}(Z^{**} \mid Z^*_{nt}, x^o_{knt}, hk^o_{nt}). \tag{9.10}$$

The optimal discrete choice $\mathbb{I}^0_{ntk}$ maximizes:

$$U_{ntk}(Z^*) + \int V_{nt+1}(Z^{**}) dF_{ntk}(Z^{**} \mid Z^*_{nt}) \tag{9.11}$$

over $\{0, 1, 2, 3\}$, patently a finite discrete choice problem. Since

$$Z^{**} \equiv [z'_{nt+1}, h_{nt}, ..., h_{nt+1-\rho}, b_{nt}, ..., b_{nt+1-M}, \mu_n \eta_n] \tag{9.12}$$

and since $\Pr[Z^*_{nt+1} = Z^* \mid Z^*_{nt}, x_{nt}, h_{nt}, \mathbb{I}_{ntk} = 1] \equiv F_{ntk}(Z^{**} \mid Z^*_{nt}, x_{nt}, h_{nt})$ then the transition density is deterministic in our model. So let Z^{**}_k the future state variable then the optimal choice $\mathbb{I}^0_{ntk}$ maximizes:

$$U_{ntk}(Z^*) + V_{nt+1}(Z^{**}_k) \tag{9.13}$$

9.2. Solving the value Function

To solve the for the valuation function we shall used a hibrid method which is a combination of the the finite horizon and infinite horizon contraction .

9.3. Finite Horizon Problem

The standard solution method is the Bellman's(1957) perspective of the backward induction. Suppose the problem has a finite horizon T, and consider the choices facing an individual entering the last period with state variables Z^*_{nT} her valueation function is simple $V_T(Z^*_{nT}) = \max\{U_{nTk}(Z^*_{nT})\}$. Taking expectation of $V_T(Z^*_{nT})$ on period before when her state variables on Z^* yields

$$\begin{aligned} g_{T-1}(Z^*) &= E[V_T(Z^*_{nT}) \mid Z^*] \\ &= E[\max\{U_{nTk}(Z^*_{nT})\} \mid Z^*] \end{aligned}$$

Lets for notational convenient denote

$$U_{ntk}(Z^*_{nt}) \equiv \overline{U_{ntk}(Z^*_{nt})} + \varepsilon_{knt}$$

Then

$$g_{T-1}(Z^*) = \sum_{k=0}^{3} p_k(Z^*_{nt})[\overline{U_{ntk}(Z^*_{nt})} + E(\varepsilon_{knt} \mid Z^*_{nt}, \mathbb{I}_{knt} = 1)]$$

By the extreme value assumption

$$\begin{aligned} p_k(Z^*_{nt}) &= \Pr[\overline{U_{ntk}(Z^*_{nt})} + \varepsilon_{knt} > \overline{U_{ntj}(Z^*_{nt})} + \varepsilon_{jnt}, \forall k \neq j] \\ &= \frac{\exp(\overline{U_{ntk}(Z^*_{nt})})}{\sum_{s=0}^{3} \exp(\overline{U_{nts}(Z^*_{nt})})} \end{aligned} \tag{9.14}$$

and

$$E(\varepsilon_{knt} \mid Z^*_{nt}, \mathbb{I}_{knt} = 1) = \zeta - \ln(p_k(Z^*_{nt})) \tag{9.15}$$

These assumption inply that

$$g_{T-1}(Z^*) = \zeta + \frac{\sum_{k=0}^{3} \exp(\overline{U_{ntk}(Z^*)}) \ln[\sum_{j=0}^{3} \exp(\overline{U_{ntj}(Z^*)})]}{\sum_{s=0}^{3} \exp(\overline{U_{nts}(Z^*)})} \tag{9.16}$$

$$= \zeta + \ln[\sum_{j=0}^{3} \exp(\overline{U_{ntj}(Z^*)})] \tag{9.17}$$

Having calculated $g_{T-1}(Z^*)$, at the beginning of the period T-1 the person chooses the maximum over the different options and obtain a valuation function of:

$$V_{T-1}(Z^*) = \max\{\overline{U_{ntk}(Z^*_{nt})} + \varepsilon_{knt} + \beta g_{T-1}(Z^{**})\} \tag{9.18}$$

In a iterative fashion lets

$$g_{T-2}(Z^*_{nT-2}) = E[V_{T-1}(Z^*_{nT-1})]. \tag{9.19}$$

By successively soluing for the functions $g_{T-1}(Z^*)$ through $g_1(Z^*)$, the value function $V_t(Z^*)$ is derived numerically for all $t \in \{0, ..., T\}$.

9.4. Infinite Horizon

Here we will used the contraction mapping theory extends the idea od iteration on the value function to the infinite horizon case. Let $H_0(Z^{**})$ be any real value bounded continuous function defined on the coordinates Z^* and define the real valued mapping as:

$$C[H(Z^{**})] = E[\max_k\{\overline{U_k(Z^*)} + \varepsilon_k + \beta H(Z^{**})\}] \tag{9.20}$$

where the itegration is over ε_k , an extreme value Type I random variable. The mapping $C[.]$ is a contraction mapping, and satisfies the fixed point property that a uniques $H(Z^{**})$ solves:

$$H(Z^{**}) = C[H(Z^{**})] \tag{9.21}$$

In this case $H(Z^{**})$ is interpreted as the expected value function for the infinite horizon problem:

$$H(Z^{**}) = E[V(Z^*)].$$

By a corrollary of the contraction mapping fixed point theorem we obtain an upper bounde on the sequence of iteration approximating function. In particular for an initial $H_0(Z^{**})$, then

$$\|C^s[H_0(Z^{**})] - H(Z^{**})\| \leq (1-\beta)^{-1} \left\|C^s[H_0(Z^{**})] - C^{s-1}[H_0(Z^{**})]\right\| \tag{9.22}$$

In our specific case with the extreme value Type I assumption on the errors, we have:

$$C[H(Z^{**})] = \zeta + \frac{\sum_{k=0}^{3} \exp(\overline{U_{ntk}(Z^*)} + \beta H(Z^{**})) \ln[\sum_{j=0}^{3} \exp(\overline{U_{ntj}(Z^*)} + \beta H(Z^{**}))]}{\sum_{s=0}^{3} \exp(\overline{U_{nts}(Z^*)} + \beta H(Z^{**}))} \tag{9.23}$$

$$\zeta + \ln[\sum_{j=0}^{3} \exp(\overline{U_{ntj}(Z^*)} + \beta H(Z^{**}))] \tag{9.24}$$

We can then use a newton fixed method to numerically solue for the fixed point. The newton iteration is of the form:

$$H^{s+1} = H^s - \frac{C[H^s]}{C'[H^s]}, \qquad s > 0$$

this gives the following

$$H^{s+1} = H^s - \beta^{-1}\ln[\sum_{j=0}^{3}\exp(\overline{U_j} + \beta H^s)], \qquad s > 0 \tag{9.25}$$

Although convergence is global, an itelligent choice for the intila function $H_0()$ reduces the number of iterations required to achieve convergence. One such choice , is to conbine the finite horizon problem with the infinite horizon problem and set

$$H_0() = (1-\beta)g_{T-1}(Z) \tag{9.26}$$

which is the discounted life time utility for a household one period before the terminal horizon. This choice works very well in our application and achieve convergence in a 4 to 7 iterations.

9.5. Policy Experiments(Incomplete)

To be presented at conference.

10. 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 intertemporal nature of utility from leisure). In addition there are a small number of studies of fertility behavior that suggested 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 empirical results reaffirm findings from previous work, and provide a set of parameters that capture the costs and benefits of having children within a dynamic structural framework. More specifically, our estimates reaffirm the importance of nonseparabilites in labor supply choices. Wages are increased by experience up to four years in the past, recent experience counting the most. In addition we reject the null hypoethesis that leisure is intertemporally separable, our estimates suggesting that there is also learning by doing in home making activities.

With regards fertility previous work estimating linear index functions and fertility hazards had found that the timing of later births depended on the ages of older siblings. Although there is not much in the literature with which we can directly compare our findings, our estimated costs and benefits of children are plausible. They imply that households view children as a good, not a bad, thus suggesting that households limit their size because of the time and money costs assocated with raising offspring, not because adults do not like having offspring. Our estimates show that there is an optimal gestion period with respect to births that is partly determined by the same economic factors that have played a role in reducing TFR over the past two generation or two, a finding which is consistent with previous work on the spacing of births over the life cycle.

Our study was motivated by the fact that the decline in childbearing coincides with higher wages for females, who are raising fewer children, participating in the labor force in greater numbers and are working longer hours. Our model is uniquely suited to analyzing whether shifts in public policy towards child support could affect these trends, and in what ways. We are currently in the midst of running policy simulations to explore such questions, and the results of these simulations will be made available before the June conference.

11. Data Appendix

In 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 contain 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 in 1985 to 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 are in the survey for four years are less, this selection criteria eliminated further 1,946 individuals from our sample. We then drop all individuals who were older than 45 in 1967, the eliminated an additional 1,531 individuals. We then drop all individuals that were less than 14 years old in 1991, this eliminated and additional 385 individuals.

The corresponding number of observations for the interviewing year 1968 through 1992 interviewing year 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 study or they were mover-out non-respondents in year 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 occur for the different measure 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 meal 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 dollars while the relevant value for this variable is the subsequent years and for the value of food consumed at home and eating out is 9,999 dollars. 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 this 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 value 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 such 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 the 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 at the value of the 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 other 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 measures 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. An 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.

References

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

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

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

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

[5] 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.

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

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

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

[9] Eckstein anf K. Wolpin (1989), *Dynamic labour Force Participation of married womem and Endogenous Work Experirence*, Review of Economic Studies, Vol. 56, 1989, 375-590.

[10] 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.

[11] 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.

[12] 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 .

[13] 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.

[14] 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.

[15] 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.

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

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

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

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

[20] 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 Polpulation Economics, Vol. 11 , 29-51.

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

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

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

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

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

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

[27] 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.

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

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

Table I: Notations

w_{nt}	individual marginal product of labor
x_{nt}	consumption of market goods
z_{nt}	demographic variables
h_{nt}	proportion of time worked in period t
	as a fraction of the total time available in the period
l_{nt}	leisure in period t :balance of time not spent at work or nurturing children
d_{nt}	labor force participation dummy
b_{nt}	indicator of the birth of a child at period t
γ_0	additional lifetime expected utility a household receives for its first child
$\gamma_0 + \gamma_k$	utility from having a second child when the first born is k years old
$\gamma_0 + \gamma_k + \gamma_j$	utility from having a third child when the first two are aged k and j years old
U_{0nt}	benefits from offspring to the n^{th} household in period t
U_{1nt}	utility costs of the n^{th} female from working in period t
U_{3nt}	current utility from consumption by household n in period t
π	discounted cost of expenditures of raising a child
ρ_k	nurturing time required by a k year old child
c_{nt}	fraction of time the n^{th} household spend nurturing children in the household
η_n^{-1}	social weight attached to each individual n
W	aggregate endowment or output from the exogenous production process
d_{nt}^o	optimal labor forceparticipation decision at date t
h_{nt}^*	optimal labor supply
h_{nt}^o	optimal labor supply conditional on participation
b_{nt}^0	optimal birth decision
λ_t	shadow value of consumption
μ_n	time-invariant individual-specific effect of marginal product

TABLE II

Characteristic of main sample

Years	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
No.of obs.	5,370	5,546	5,731	5,906	6,133	6,326	6,490	6,674	6,855	7,071	7,210	7,282
Demographics												
Age	14.69	15.51	16.36	17.22	18.10	19.00	19.92	20.86	21.81	22.79	23.78	24.78
Family size	9.13	8.70	8.11	7.66	7.17	6.61	6.17	5.76	5.39	5.07	4.80	4.51
No.of child. <6	2.19	2.05	1.92	1.80	1.68	1.58	1.49	1.40	1.31	1.22	1.15	1.09
No.of child. 6 to 14	.498	.494	.488	.485	.466	.451	.429	.411	.400	.387	.381	.374
Prop.married	.340	.350	.361	.368	.379	.392	.396	.395	.398	.398	.399	.412
Years of compl. edu.	11.26	11.26	11.23	11.21	11.19	11.17	11.15	11.12	11.11	11.11	11.11	11.13
Prop. black	.438	.437	.434	.434	.430	.429	.426	.425	.423	.420	.419	.418
Prop. hisp.	.045	.045	.045	.045	.045	.046	.046	.046	.046	.046	.046	.046
Hours,Earnings, Participation and Consumption												
Household income	28,585	30,930	31,967	31,831	32,597	33,942	35,213	34,143	32,772	34,157	34,881	35,608
Total labor income	777	869	1,022	1,125	1,224	1,327	1,516	1,702	1,908	2,112	2,351	2,718
Annual hrs worked	410	426	456	461	470	475	505	506	519	535	558	598
Hourly earnings	1.91	1.95	2.20	2.30	2.55	2.72	2.94	3.26	8.07	3.80	4.17	4.32
Prop. who worked	.543	.555	.562	.559	.556	.543	.550	.549	.546	.538	.547	.577
Food Consumption	7,758	7,485	7,372	7,390	7,187	7,618	7,777	7,244	7,125	6,612	6,409	6,469

TABLE II

Characteristic of main sample (to be continued)

Years	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
No.of obs.	7,353	7,408	7,491	7,562	7,600	7,662	7,612	7,566	7,511	7,459	7,341	7,200	7,063
Demographics													
Age	25.78	26.78	27.78	28.78	29.78	30.78	31.78	32.78	33.78	34.78	35.78	35.78	35.78
Family size	4.33	4.18	4.03	3.91	3.79	3.65	3.53	3.42	3.31	3.22	3.15	3.09	3.05
No.of child. <6	1.03	.972	.913	.850	.794	.734	.673	.619	.566	.508	.455	.405	.351
No.of child. 6 to 14	.374	.376	.379	.387	.390	.395	.413	.430	.442	.458	.477	.487	.504
Prop.married	.421	.425	.430	.440	.448	.457	.462	.467	.474	.484	.492	.497	.504
Years of compl. edu.	11.15	11.17	11.19	11.22	11.24	11.26	11.28	11.31	11.34	11.38	11.38	11.38	11.40
Prop. black	.418	.416	.415	.414	.412	.411	.409	.405	.402	.399	.397	.395	.392
Prop. hisp.	.046	.046	.046	.045	.046	.046	.046	.046	.046	.046	.046	.047	.047
Hours,Earnings, Participation and Consumption													
Household income	35,276	34,092	32,482	31,393	32,447	33,776	34,012	34,667	35,102	36,426	36,175	35,481	34,898
Total labor income	3,109	3,446	3,843	4,174	4,631	5,196	5,775	6,323	6,960	7,752	8,298	8,945	9,550
Annual hrs worked	615	637	648	659	680	750	773	801	836	874	899	931	947
Hourly earnings	4.82	5.18	5.64	5.97	6.71	6.64	7.11	7.35	7.86	8.25	8.81	9.20	9.72
Prop. who worked	.594	.607	.607	.611	.625	.656	.665	.687	.711	.728	.755	.772	.773
Food Consumption	6,268	6,099	5,840	5,722	5,866	5,946	5,821	5,756	5,828	5,702	5,616	5,552	5,632

Table III. Wage Equation

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

Variable	Parameter	Estimate
Lags of hours worked		
$\triangle h_{n,t-1}$	δ_{11}	9.1680 (0.1324)
$\triangle h_{n,t-2}$	δ_{12}	6.6241 (0.1383)
$\triangle h_{n,t-3}$	δ_{13}	5.2116 (0.1364)
$\triangle h_{n,t-4}$	δ_{14}	3.1320 (0.1266)
Lags of participation		
$\triangle d_{n,t-1}$	δ_{21}	-2.1693 $(2.28e-02)$
$\triangle d_{n,t-2}$	δ_{22}	-1.6858 $(2.48e-02)$
$\triangle d_{n,t-3}$	δ_{23}	-0.9500 $(2.47e-02)$
$\triangle d_{n,t-4}$	δ_{24}	-0.6819 $(2.22e-02)$
Socioeconomic Variables		
$\triangle AGE^2_{nt}$	B_{31}	-0.0058 $(2.0e-04)$
$\triangle(AGE_{nt} \times EDU_{nt})$	B_{32}	$5.11e-02$ $(8.0e-04)$

(Standard Errors in parenthesis)

Table IV: Consumption Equation

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

Variable	Parameter	Estimate
Socieconomic variables		
$\triangle FAM_{nt}$	$(1-\alpha)^{-1}B_{21}$	$3.30e-02$ $(3.0e-04)$
$\triangle YKID_{nt}$	$(1-\alpha)^{-1}B_{22}$	$-3.33e-02$ $(1.5e-03)$
$\triangle OKID_{nt}$	$(1-\alpha)^{-1}B_{23}$	$-1.18e-02$ $(1.1e-03)$
$\triangle AGE^2_{nt}$	$(1-\alpha)^{-1}B_{24}$	$-1.0e-04$ $(1.0e-07)$
Region Dummies		
$\triangle NC_{nt}$	$(1-\alpha)^{-1}B_{25}$	$-6.7e-03$ $(3.0e-03)$
$\triangle SO_{nt}$	$(1-\alpha)^{-1}B_{26}$	$-1.39e-02$ $(2.7e-03)$

(Standard Errors in parenthesis)

Table V: Fixed Utility Cost of Labour Force Participation

$u_{10}(z_{nt}) = B_0 z'_{nt}$

Variable	*Parameter*	*Nonparametric Estimates*	*Traditional Estimates*
$CONSTANT$	B_{00}	1.0369 (4.7575)	3.8003 (4.2458)
AGE_{nt}	B_{01}	-17.2338 (1.2669)	$-1.1189e-02$ (0.58034)
AGE^2_{nt}	B_{02}	$8.9032e-02$ $(2.7328e-02)$	$2.334e-02$ $(1.4484e-02)$
$AGE_{nt} \times EDUC_{nt}$	B_{03}	1.0255 (0.14464)	0.10462 $(7.3445e-02)$
$MART.STATUS_{nt}$	B_{04}	-71.0032 (7.3629)	-28.8882 (6.4046)
$BLACK_{nt}$	B_{05}	$1.3449e+02$ (2.1974)	7.5066 (4.6588)
$HISPANIC_{nt}$	B_{06}	10.6969 (12.8036)	-13.1433 (12.5981)
$YKID_{nt}$	B_{07}	-30.8942 (7.551)	-7.0725 (3.3293)
$OKID_{nt}$	B_{08}	66.2822 (5.5852)	5.2438 (3.7016)
$RISK\ AV.$	α	0.125 $2.5e-03$	0.091 $3.1e-03$

(Standard Errors in parenthesis)

Table VI: Utility Cost of Leisure

$$U_{11}\left(z_{nt}^{*}, l_{nt}\right) = z_{nt}' B_{11}(1 - h_{nt} - c_{nt}) + \sum_{s=0}^{\rho} \theta_0 h_{n,t-s} h_{nt}$$

Variable	*Parameter*	*Nonparametric Estimates*	*Traditional Estimates*
h_{nt}	B_{110}	$1.5024e+03$ $(2.3174e+02)$	84.6155 $(1.1945e+02)$
$AGE_{nt} \times h_{nt}$	B_{111}	$-1.3770e+02$ (6.6123)	-0.23512 (5.7547)
$AGE_{nt}^2 \times h_{nt}$	B_{112}	0.65743 (0.28421)	-0.16932 (0.15327)
$AGE_{nt} \times EDUC_{nt} \times h_{nt}$	B_{113}	8.2049 (1.4732)	0.7542 (0.77288)
$MART.STATUS_{nt} \times h_{nt}$	B_{114}	$7.0259e+02$ (81.0569)	$2.2536e+02$ (75.6154)
$BLACK_{nt} \times h_{nt}$	B_{115}	$1.1524e+03$ (33.205)	29.9664 (61.4587)
$HISPANIC_{nt} \times h_{nt}$	B_{116}	-56.5003 $(1.6632e+02)$	$-2.3438e+02$ $(1.6396e+02)$
$YKID_{nt} \times h_{nt}$	B_{117}	$5.7925e+02$ (87.9956)	19.3398 (40.0993)
$OKID_{nt} \times h_{nt}$	B_{118}	$-3.0994e+02$ (81.9573)	-61.5767 (51.2033)
$NURT.\ TIME$	ρ_0	0.15066 $1.0814e-03$	$4.9731e-03$ $1.9338e-03$
h_{nt}^2	θ_0	$-7.6067e+03$ $(4.5587e+02)$	$1.3496e+02$ $(3.5763e+02)$
$h_{nt}h_{nt-1}$	θ_1	$4.2192e+04$ $(1.9252e+02)$	$-4.2933e+02$ $(1.3727e+03)$
$h_{nt}h_{nt-2}$	θ_2	$-2.6934e+04$ $(9.7361e+02)$	$-3.3903e+03$ $(8.5779e+02)$
$h_{nt}h_{nt-3}$	θ_3	$8.2431e+03$ $(1.0929e+03)$	$5.1877e+03$ $(6.7510e+02)$
$h_{nt}h_{nt-4}$	θ_4	$-8.1166e+03$ $(3.6815e+02)$	$-8.6618e+02$ $(2.8484e+02)$

(Standard Errors in parenthesis)

Table VII: Birth Effects

$$U_{0nt} = b_{nt}\left(\gamma_0 + \sum_{k=1}^{M} \gamma_k b_{nt-k}\right) + b_{nt}\varepsilon_{1nt} + (1 - b_{nt})\,\varepsilon_{2nt}$$
$$\pi_{nt}\,(z_{nt}) = \pi_0 + z'_{nt}\pi_1$$

Variable	*Parameter*	*Nonparametric Estimates*	*Traditional Estimates*
b_{nt}	γ_0	7.3008 $(3.1843e-02)$	6.4047 $(5.3733e-02)$
$b_{nt}b_{nt-1}$	γ_1	-0.41214 $(2.0614e-02)$	-0.41375 $(2.3618e-02)$
$b_{nt}b_{nt-2}$	γ_2	2.2951 $(2.1092e-02)$	$-1.6705e-02$ $(2.7973e-02)$
$b_{nt}b_{nt-3}$	γ_3	2.3311 $(2.136e-02)$	0.26757 $(2.9038e-02)$
$b_{nt}b_{nt-4}$	γ_4	1.1563 $(2.2099e-02)$	0.49679 $(2.7186e-02)$
$b_{nt}b_{nt-5}$	γ_5	-0.6471 $(2.1729e-02)$	-0.86401 $(2.4547e-02)$
$CONSTANT$	π_0	$5.8631e-03$ $(5.2509e-04)$	$8.8262e-03$ $(2.5147e-03)$
$NO\ HIGH\ SCH_{nt}$	π_1	$7.8272e-02$ $(5.7429e-04)$	$1.3019e-02$ $(2.6559e-03)$
$BLACK_{nt}$	π_2	$3.1473e-02$ $(7.567e-04)$	$3.7921e-04$ $(3.098e-03)$
$HISPANIC_{nt}$	π_3	0.26004 $(2.0344e-03)$	$2.9552e-02$ $(9.4566e-03)$

(Standard Errors in parenthesis)

Figure I-a: Traditional Estimates of Fixed Effects of Marginal Products

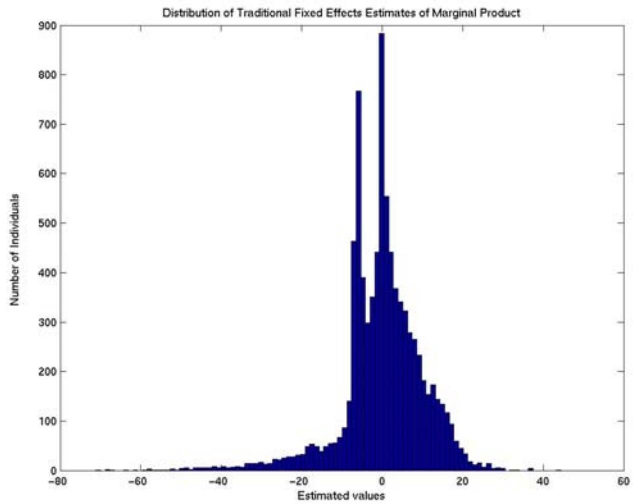

Figure I-b: Nonparametric Estimates of Fixed Effects of Marginal Products

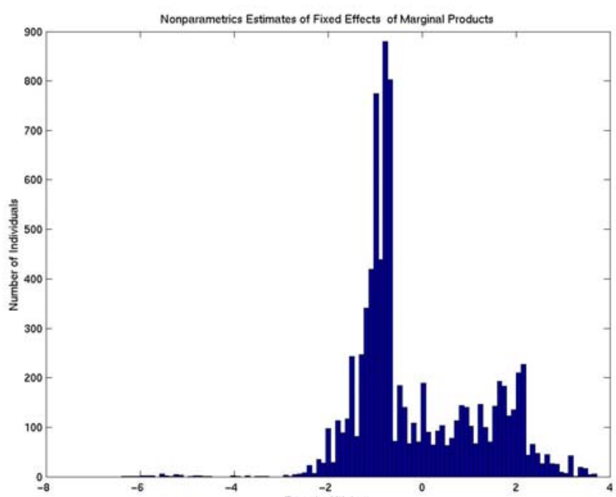

Figure II: Aggregate Prices

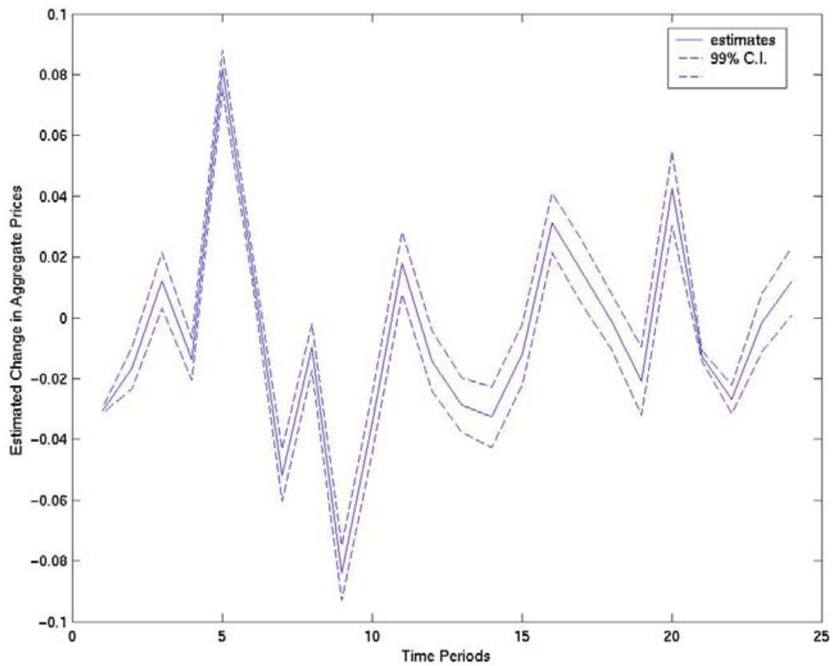

Figure III -a: Traditional Fixed Effects Estimates of Social Weights

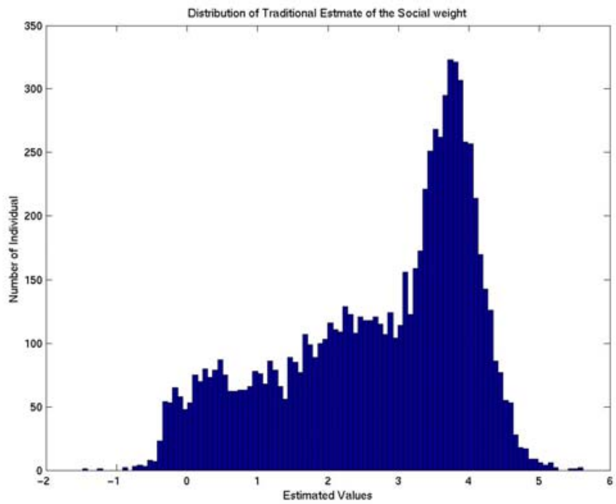

Figure III-b: Nonparametric Fixed Effects Estimates of Social Weights

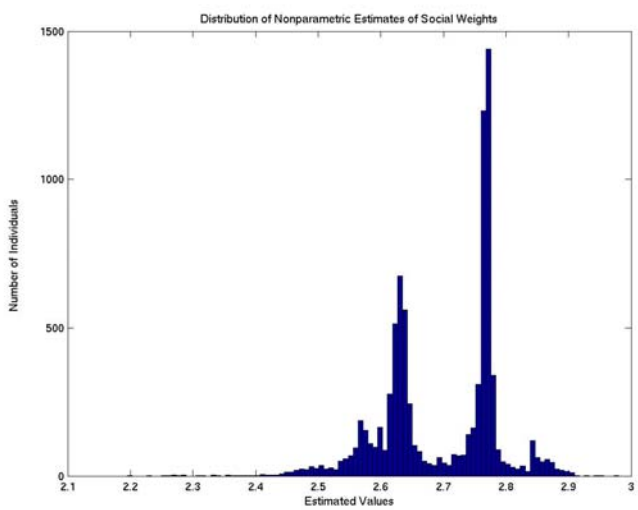

Figure IV-a: Traditional Shadow Prices of Consumption

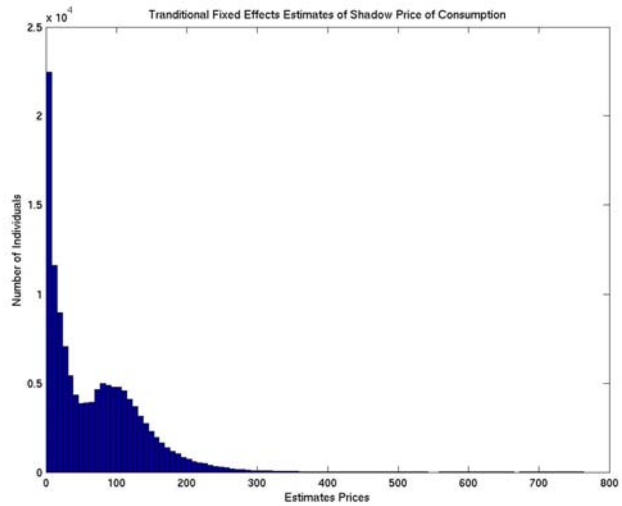

Figure IV-b: Nonparametric Shadow Prices of Consumption