Review of Economic Studies (1998) **65**, 45–85 0034-6527/98/00030045$02.00

The Effect of Work Experience on Female Wages and Labour Supply

SUMRU ALTUĞ
Koç University

and

ROBERT A. MILLER
Carnegie-Mellon University

First version received February 1995; *final version accepted August* 1997 (*Eds.*)

This paper develops and implements a semiparametric estimator for investigating, with panel data, the importance of human capital and time nonseparable preferences to females when aggregate shocks are present. It provides a set of conditions for making statistical inferences about agents' expectations of their correlated future choices, from a short panel. Under the assumption that observed allocations are Pareto optimal, a dynamic model of female labour supply and participation is estimated, in which experience on the job raises future wages, and time spent off the job in the past directly affects current utility (or, indirectly through productivity in the non-market sector).

1. INTRODUCTION

A growing body of empirical research finds current wages and marginal utilities of leisure are affected by past labour supply choices. These dynamic interactions have been found by Miller (1984), Hotz, Kydland and Sedlacek (1988), Eichenbaum, Hansen and Singleton (1988), Shaw (1989), Eckstein and Wolpin (1989a) and Miller and Sanders (1997). The result is robust to differences about how investment in human capital is modelled (whether through labour market participation, hours worked, or job match), assumptions about financial markets, and the data used (aggregate time series vs. panels; data on males vs. females). There is, however, little agreement beyond the stylized fact that some form of human capital accumulation (in the market and/or home sector) is empirically significant. For example, Hotz, Kydland and Sedlacek (1988) conclude that current leisure and a stock of accumulated leisure are substitutes in current utility, whereas Eichenbaum, Hansen and Singleton (1988) and Miller and Sanders (1997) find evidence indicating complementarity. Yet, as Kydland and Prescott (1982) and Altuğ (1989) demonstrate, such features can play an important role in propagating aggregate shocks.

This paper develops and implements a semiparametric estimator for investigating, with panel data, the importance of human capital and time separable preferences in the presence of aggregate shocks. It analyses how past decisions about labour market participation and hours spent at work by women affect their current wages and employment, and estimates the depreciation that occurs when women temporarily or permanently withdraw from the labour force. While our framework does not statistically nest the models that have been estimated previously, three of its distinguishing features have potentially important empirical consequences.

First, in contrast to existing studies, we estimate parameters that characterize both nonseparabilities in preferences with respect to leisure and learning-by-doing on the job. The panel data studies of Miller (1984), Shaw (1989) and Eckstein and Wolpin (1989a) impose the form of preferences directly. Miller (1984) and Shaw (1989) assume leisure is additively separable over time while Eckstein and Wolpin (1989a) assume that only the number of years a person has worked affects the leisure stock. Eichenbaum, Hansen and Singleton (1988) ignore marketable human capital. Hotz, Kydland and Sedlacek (1988) avoid taking a stand on the role of human capital in determining wages because they do not estimate the Euler equation for the optimal (interior) choice for hours. Their results are based on the Euler equation for consumption, and account for intertemporal nonseparabilities in leisure only if leisure enters preferences nonseparably with consumption.

Second, the panel data studies cited above identify the structural parameters from variation in individual behaviour and either ignore the existence of aggregate shocks (as in Shaw, 1989 or Eckstein and Wolpin, 1989a), or account for them in an inconsistent manner (as in Hotz, Kydland and Sedlacek, 1988). Eichenbaum, Hansen and Singleton (1988) exploit variation in aggregate time series data and impose strong functional form restrictions to achieve aggregation (ruling out, for example, the boundary choice of not participating in the labour market). These differences might not be cause for concern if they were not reflected in such different findings mentioned above. Our study exploits a panel data set to capture heterogeneity throughout the population while simultaneously accounting for the effects of aggregate shocks on current wages and prices and on the future value of market vs. nonmarket work. We employ two approaches for dealing with population heterogeneity. The first follows MaCurdy (1981), who assumes that unobserved individual characteristics can be described as a parametric function of observed variables. We extend MaCurdy's approach by modelling the unobserved fixed effect as a nonparametric function of observables. Our second approach to dealing with unobserved population heterogeneity follows Heckman and MaCurdy (1980), who use a standard fixed-effects estimator which estimates time-invariant individual characteristics using the time dimension of a panel data set.

A third distinguishing feature of our analysis is that we deal with restrictions that arise from the decision to participate or not, and conditional on participating, the number of hours worked. This is in contrast to the other studies mentioned above, which deal with the participation choice or the hours choice but not both. Our approach allows us to estimate such features as the fixed costs of participating in the labour market using information on both participation behaviour and hours worked, for example. Unlike most earlier studies, we do not employ dynamic programming methods in order to evaluate the valuation functions associated with the decision to participate or not. Instead, we extend the approach in Hotz and Miller (1993) to derive alternative representations of the conditional valuation functions in terms of known functions of conditional choice probabilities, which show the probability of a woman participating conditional on some given history. These conditional choice probabilities are nonparametrically estimated, and treated as incidental parameters in the orthogonality conditions used to estimate the remaining parameters of interest.

One of the implicit assumptions in some of the studies that we cited above, including Kydland and Prescott (1982) and Eichenbaum, Hansen and Singleton (1988), is that observed consumption and leisure allocations are Pareto optimal, and are generated as the solution to a social planner's problem that maximizes a weighted sum of individuals' utilities subject to an economy-wide resource constraint. This assumption is used here to allow for the effects of aggregate shocks, and to account for the nonconvexity induced by the fixed costs of labour market participation. In our framework, aggregate shocks are transmitted to individual behaviour through prices or shadow values alone. However, their effects cannot be captured by simply inserting time dummies in estimated equations. In our application, the value of accumulating an additional unit of human capital depends on the future (shadow) value of consumption and future aggregate wages. Thus, the decision to acquire additional human capital today requires an estimate of the future benefits of this choice. One way to form an estimate of the future (random) benefits of a current action is to derive the probability distribution that characterizes the state variables of the model. However, characterizing the endogenous distribution of the state variables in models of dynamic discrete choice, is, in general, difficult.

Our solution to this problem is based on the intuition under Pareto optimality, the logarithm of an individual's shadow value of consumption decomposes additively into two terms, a time-invariant individual-specific effect and an aggregate (or common) or shadow value of consumption. If the Pareto optimal allocations can be supported as a competitive equilibrium, the individual-specific effects correspond to individuals' marginal utilities of wealth and the shadow values of consumption are prices to contingent claims for consumption (Altuğ and Miller (1990)). Thus, it is possible to predict the behaviour of a wealthy agent living in economic slumps by observing that of a poorer person living in a prosperous world. Hence, the probability distribution describing the behaviour of an individual in some future state can be inferred by nonparametrically estimating the current behaviour of individuals she may later mimic, weighted by the probability of this state occurring. Rather than estimate the whole probability distribution, however, we adapt the techniques of simulation estimation to this nonparametric context.

The next section describes our framework. Section 3 describes the data, while Sections 4 through 6 are devoted to the estimation of each component of the model. These include the wage and consumption equations and time-invariant individual-specific effects, the nonparametrically estimated conditional choice probabilities, and the participation and hours conditions; a detailed overview of our empirical analysis is provided in Section 2.3.

2. A FRAMEWORK

2.1. *Preferences and technology*

We begin by describing individual preferences and the marginal productivity to working. We assume that there exists a continuum of agents on the unit interval [0, 1]. The consumption allocation to individual n at date $t \in \{0, 1, \dots\}$ is denoted c_{nt}. The other choice variable is l_{nt}, the time spent at work by agent n in period t. There is a fixed amount of time in each period that is available for working, which implies that the amount of time worked in each period can be normalized as $0 \leqq l_{nt} \leqq 1$. If $l_{nt}=0$, the agent does not work at time t. Otherwise, the agent works the fraction of time $l_{nt}>0$. For convenience, a participation indicator d_{nt} is also defined, where $d_{nt}=1$ if and only if $l_{nt}>0$, and 0 otherwise.

Preferences are additive in consumption and leisure but not separable with respect to leisure at different dates. To capture this dependence, define $(l_{nt-\rho}, \dots, l_{nt-1})'$ as the ρ-dimensional vector of past labour supply outcomes that affect current utility. It is assumed that there are both observed and unobserved exogenous, time-varying characteristics that determine the utility associated with alternative consumption and leisure allocations. Denote the former by the $K \times 1$ vector x_{nt} and the latter by the 2×1 vector $(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})'$.

It is assumed that x_{nt} is independently distributed over the population with a known distribution function $F_0(x_{nt+1} \mid x_{nt})$; the vector $(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})'$ is independent across (n, t) and drawn from a population with a distribution function $F_1(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt})$. Define $z_{nt} \equiv (l_{nt-\rho}, \ldots, l_{nt-1}, x'_{nt})'$. The first ρ elements of z_{nt} capture the dependence of the individual's state on lagged labour supply choices.

The current-period utility function at date t for individual n is defined as

$$U_{nt} \equiv d_{nt} U_0(x_{nt}) + d_{nt}\varepsilon_{1nt} + (1 - d_{nt})\varepsilon_{0nt} + U_1(z_{nt}, l_{nt}) + U_2(c_{nt}, x_{nt}, \varepsilon_{2nt}), \tag{2.1}$$

where $U_1(z_{nt}, l_{nt})$ is assumed to be concave decreasing in l_{nt} for any z_{nt}, and $U_2(c_{nt}, x_{nt}, \varepsilon_{2nt})$ is assumed to be concave increasing in c_{nt} for any z_{nt} and ε_{2nt}. Intuitively, U_0 represents the fixed utility costs from working (which depends only on the observed individual-specific characteristics x_{nt}), U_1 is the utility cost of working a greater fraction of time (which varies with current time spent at work as well as observed individual-specific characteristics and time spent at work in the past), and U_2 is the current utility from consumption (which depends on consumption and the observed characteristics of the individual and the idiosyncratic shocks). Let $\beta \in (0, 1)$ denote the common subjective discount factor, and $E_0(\cdot)$ expectation conditional on information available at date 0. The preferences of individual n are defined as

$$E_0\{\textstyle\sum_{t=0}^{\infty} \beta^t U_{nt}\}. \tag{2.2}$$

There are two ways of obtaining the output of the consumption good in this economy. The first is a linear production technology that is available to each individual; if an individual works, she produces the quantity of output $w_{nt} l_{nt}$ in each period, where w_{nt} is the person-specific marginal product of labour at date t. We assume that w_{nt} can be written in terms of the exogenously determined productivity of a standard efficiency unit of labour in period t denoted ω_t, an individual-specific, time-invariant productivity effect v_n, and a function $\gamma(z_{nt})$, which reflects the socio-economic characteristics and the stock of human capital accumulated by the individual. Symbolically

$$w_{nt} = \omega_t v_n \gamma(z_{nt}). \tag{2.3}$$

Thus, $v_n \gamma(z_{nt})$ is the efficiency units a person produces per unit of time or a measure of their individual-specific human capital, while ω_t depends on aggregate effects alone. A second source of consumption is through an exogenous production process that yields e_t units of output each period.

This specification of preferences and choices over leisure differs from those found in the existing literature in several important aspects. In contrast to previous static models of female labour supply (discussed in Killingsworth (1983) and Smith (1980) for example), and their dynamic counterparts (estimated by Heckman and MaCurdy (1980) and MaCurdy (1981)), investment in human capital and intertemporally nonseparable preferences play an important role here. Another distinctive feature of our formulation is that conditional on working, labour supply plans depend on observables alone.[1] The limited role unobservables play in self-selection, coupled with the prominence attached to dynamic considerations, is foreshadowed in Mroz (1987), who reports that "among potential specifications found unimportant are . . . controls for self-selection when experience is treated as endogenous" (p. 795) whereas assuming the wife's wage is exogenous induces "upward bias in the estimated wage effect" (p. 795).

1. This assumption is made for tractability in estimation; if an unobserved component entered the marginal utility of leisure, this would complicate greatly estimation of the Euler equation.

Our model also assumes that preferences are additively separable over current leisure and consumption as well as additively separable with respect to the leisure time of husbands and wives. The evidence on these types of nonseparabilities is mixed. Although most empirical work on female labour supply routinely ignores consumption data and vice versa, the small number of studies that have investigated both, notably Altuğ and Miller (1990) and Browning and Meghir (1991), find evidence favouring preferences that are contemporaneously nonseparable. Altuğ and Miller (1990) also find some evidence against preferences that are additively separable with respect to husbands' and wives' leisure time. Neither study, however, asks how their findings would change if preferences were not time additive so in this respect, our work is complementary to these earlier studies.

While the work cited above is related, the structural framework analysed here most closely resembles Eckstein and Wolpin (1989a) and Miller and Sanders (1997). Eckstein and Wolpin estimate their dynamic model of female labour force participation and endogenous work experience by solving a dynamic programme and using maximum likelihood methods. Miller and Sanders investigate female labour force and welfare participation using the conditional choice probability estimator of Hotz, Miller, Sanders and Smith (1994), which is similar to the one developed here. There are several features differentiating the underlying economic frameworks. First, our framework endogenizes labour market participation and the choice of hours worked whereas their studies only investigate discrete choices. Second, the other two studies assume there is neither saving nor borrowing, and that all disturbances are identically and independently distributed over individuals and time. Our study assumes all resouces are allocated optimally, and that fluctuations in aggregate variables lead to correlated activity at the individual level. A third difference stems from precisely how the effects of past participation and labour supply are modelled. Despite the fact that none of the under-lying economic models is nested within another, the estimation techniques differ, and the empirical analyses are based on three different data sets, the papers are nonetheless quite comparable. For this reason, we compare our empirical findings with those of the other two papers.

2.2. *Optimal consumption and labour supply allocations*

The approach of analysing the empirical implications of economies in which resource allocation is assumed to satisfy the Pareto optimality criterion was recently utilized by Altuğ and Miller (1990) to estimate a life-cycle model of consumption and male labour supply with aggregate shocks. Other recent papers that discuss Pareto optimal allocations and estimate frameworks based on this assumption include Card (1990), Mace (1991), Townsend (1994), Miller and Sieg (1997), Altonji, Hayashi and Kotlikoff (1996), among others. While the assumption of Pareto optimality is controversial, it is empirically tractable and serves as a useful benchmark. In this paper, we extend our previous analysis by considering a model in which there are aggregate shocks, fixed costs of participation, and intertemporal nonseparabilities in preferences and wages. The assumption that observed allocations are Pareto optimal allows us to deal with the effects of aggregate shocks on individual allocations, and to derive conditions for the determination of the optimal participation and hours decisions that are tractable in estimation.

The Pareto optimal allocations are derived as the solution to a social planner's problem, in which the weighted sum of the expected discounted utilities of each individual n are maximized subject to an aggregate feasibility or resource constraint. Let η_n^{-1} denote the social weight attached to each individual n. The primitives of the social planner's problem are the individual utilities defined by equation (2.1), the social weights n_n^{-1} for

$n \in [0, 1]$, the marginal product function (2.3), and the technology parameters described in Section 2.1, namely, the individual-specific, time-invariant productivity parameter v_n, the marginal product of a unit of standardized human capital in period t, ω_t, and the output from the exogenous production process, e_t.

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 e_t

$$\int_0^1 [c_{nt} - w_{nt} l_{nt}] d\mathscr{L}(n) \leqq e_t, \qquad t \in \{0, 1, \ldots\}. \tag{2.4}$$

In this expression, $\mathscr{L}$ is the Lebesgue measure which integrates over the population. The Pareto optimal allocations are found by maximizing

$$E_0 \left\{ \int_0^1 \sum_{t=0}^{\infty} \beta^t \eta_n^{-1} U_{nt} d\mathscr{L}(n) \right\}, \tag{2.5}$$

subject to (2.4) with respect to sequences for consumption and labour supply $\{c_{nt}, l_{nt}\}_{t=0}^{\infty}$ for all individuals $n \in [0, 1]$.

The necessary conditions characterizing the optimal consumption and labour supply allocations form the basis for estimation. Define $\beta^t \lambda_t$ as the Lagrange multiplier associated with the aggregate feasibility constraint in equation (2.4). The optimal consumption allocations satisfy the necessary conditions

$$\frac{\partial U_2(c_{nt}, x_{nt}, \varepsilon_{2nt})}{\partial c_{nt}} = \eta_n \lambda_t, \tag{2.6}$$

for all $n \in [0, 1]$ and $t \in \{0, 1, \ldots\}$. As in Altuğ and Miller (1990), the assumption of Pareto optimality for observed allocations yields a simple representation for individuals' shadow value of consumption. With the contemporaneous separability of consumption from labour supply choices, the condition in (2.6) can be used to solve for individuals' Frisch demand functions, which determine consumption in terms of the time-varying characteristics x_{nt}, the idiosyncratic shock to preferences ε_{2nt}, and the shadow value of consumption $\eta_n \lambda_t$. More precisely, the optimal consumption allocation of individual n at each date t can be expressed as $c_{nt}^o = c^o(x_{nt}, \varepsilon_{2nt}, \eta_n \lambda_t)$. Such Frisch demand functions are used in Section 5.1 to estimate the parameters of the function U_2, and to derive observable measures of the individual-specific weights η_n.

Characterizing the optimal labour market participation and hours of work decisions is more involved. The conditional valuation functions associated with the work decision of individual n at time t are building blocks of this analysis. Noting that each person's labour supply contributes an infinitesimal addition to aggregate output, define the conditional valuation functions associated with the decision to work or not as

$$V_{1nt} + \varepsilon_{1nt} \equiv \max_{\{l_{nr}\}_{r=t}^{\infty}} E_t \{ \textstyle\sum_{r=t}^{\infty} \beta^{r-t} [d_{nr} U_0(x_{nr}) + d_{nr} \varepsilon_{1nr} + (1 - d_{nr}) \varepsilon_{0nr}$$
$$+ U_1(z_{nr}, l_{nr}) + \eta_n \lambda_r w_{nr} l_{nr}] \mid l_{nt} > 0 \}, \tag{2.7}$$

$$V_{0nt} + \varepsilon_{0nt} \equiv \max_{\{l_{nr}\}_{r=t}^{\infty}} E_t \{ \textstyle\sum_{r=t}^{\infty} \beta^{r-t} [d_{nr} U_0(x_{nr}) + d_{nr} \varepsilon_{1nr} + (1 - d_{nr}) \varepsilon_{0nr}$$
$$+ U_1(z_{nr}, l_{nr}) + \eta_n \lambda_r w_{nr} l_{nr}] \mid l_{nt} = 0 \}. \tag{2.8}$$

Up to a factor of proportionality, the term $V_{1nt} + \varepsilon_{1nt}$, denotes the social value from n working at date t while the term V_{0nt} denotes the social value from this individual not participating. Let l^o_{nt} denote optimal labour supply. If $l^o_{nt} \in (0, 1)$, then define l^*_{nt} as the optimal interior choice. Define d^o_{nt} as the optimal participation decision at date t. Thus l^*_{nt} and d^o_{nt} satisfy the necessary conditions

$$\frac{\partial V_{1nt}}{\partial l_{nt}} = 0, \tag{2.9}$$

$$d^o_{nt} = \begin{cases} 1, & \text{if } V_{1nt} + \varepsilon_{1nt} \geqq V_{0nt} + \varepsilon_{0nt}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

The optimal labour supply decision in period t denoted l^o_{nt} affects the conditional valuation functions both through the future labour market histories $(l_{n,r-\rho}, l_{n,r-\rho+1}, \ldots, l_{n,r-1})'$ and also through the future marginal products w_{nr} for $r = t+1, \ldots,$. The above discussion implies that $l^o_{nt} = d^o_{nt} l^*_{nt}$ or unity. To complete the characterization of l^*_{nt}, the conditional valuation function V_{1nt} is expressed recursively. For this purpose, define p_{nt} as the conditional participation rate in period t

$$p_{nt} \equiv \int_{-\infty}^{V_{1nt} - V_{0nt}} (\varepsilon_{0nt} - \varepsilon_{1nt}) dF(\varepsilon_{0nt}, \varepsilon_{1nt}, \varepsilon_{2nt}). \tag{2.11}$$

Then V_{1nt} can be expressed as

$$V_{1nt} = \max_{l_{nt}} \{ U_0(x_{nt}) + U_1(z_{nt}, l_{nt}) + \eta_n \lambda_t w_{nt} l_{nt} + \beta E_t [p_{n,t+1} V_{1n,t+1} + (1 - p_{n,t+1}) V_{0n,t+1}] \mid l_{nt} > 0 \}. \tag{2.12}$$

Thus, the value of participating in the labour market at date t depends on the current net benefits of participating, defined as $U_0(x_{nt}) + U_1(z_{nt}, l_{nt}) + \eta_n \lambda_t w_{nt} l_{nt}$, plus the future benefits from current participation, which are captured by the terms on the last line of (2.12). From (2.12), notice that if an interior solution exists

$$l^*_{nt} = \underset{l_{nt} \in \{(0,1)}{\operatorname{argmax}} \{ U_0(x_{nt}) + U_1(z_{nt}, l_{nt}) + \eta_n \lambda_t w_{nt} l_{nt} + \beta E_t [p_{n,t+1} V_{1n,t+1} + (1 - p_{n,t+1}) V_{0n,t+1}] \}. \tag{2.13}$$

In this case, the condition $\partial V_{1nt} / \partial l_{nt} = 0$ evaluated at l^*_{nt} is

$$\frac{\partial U_1(z_{nt}, l^*_{nt})}{\partial l_{nt}} + \eta_n \lambda_t w_{nt} = -\beta E_t \left\{ \frac{\partial [V_{0n,t+1} + p_{n,t+1}(V_{1n,t+1} - V_{0n,t+1})]}{\partial l_{nt}} + (V_{1n,t+1} - V_{0n,t+1}) \frac{\partial p_{n,t+1}}{\partial l_{nt}} \bigg| l_{nt} = l^*_{nt} \right\}. \tag{2.14}$$

The existence of an interior optimum for the fraction of time spent at work requires that the objective function in (2.8) is locally concave at the stationary point which satisfies (2.14). Since V_{1nt} includes the future wage benefits from working today, an interior solution will not exist if, for all $l_{nt} < 1$, the marginal benefit of working exceed the current and future utility costs. In this case, the individual specializes in working, meaning $l^o_{nt} = 1$, provided the fixed costs of working are sufficiently small.

Reviewing (2.14), note that with endogenous labour market participation, the Euler equation must be modified to take into account the value an extra hour's work creates through its effect on the probability of future participation. If the participation decision were exogenous, then $\partial p_{n,t+1}/\partial l_{nt}=0$, implying that the second set of terms on the right-side of (2.14) is zero at the optimal value l^*_{nt}. In this case, the Euler equation reduces to the type of intertemporal Euler equations considered by Eichenbaum, Hansen and Singleton (1988) or Hotz, Kydland and Sedlacek (1988).[2] Without such an exogeneity assumption, the optimal choice of hours varies with the effect of additional work at date t on the probability of future labour market participation.

Another feature of (2.14) worth noting is that in a panel data context, the forecast error obtained by replacing future values of variables in (2.14) with their realized values is correlated across individuals. (See Chamberlain (1984). Our earlier work showed how the assumption of complete markets could be used to provide a consistent way of estimating the parameters of time-separable models using cross-sectional averages of individuals. In the current framework, we follow a similar approach. However, the existence of intertemporal nonseparabilities in preferences and wages implies that future values of endogenously determined variables affect current choices. Moreover, with aggregate uncertainty in the economy due to aggregate variation in technology or endowments, the future values of such endogenously determined variables will vary with the aggregate state of the economy. To summarize the effects of the unknown future aggregates prices λ_{t+s} and ω_{t+s} on current decisions using a finite set of variables that are known at time t, we make the following assumption:

Assumption 1. Define $\pi_t\equiv(\lambda_t\omega_t)/(\lambda_{t-1}\omega_{t-1})$. The sequence of random variables $\{\pi_t\}_{t=0}^{\infty}$ is an independently and identically distributed process with distribution function G.

Thus, the shadow price of a standardized unit of human capital denoted $\lambda_t\omega_t$ is assumed to follow a first-order Markov process. Consequently, current realizations of $\lambda_t\omega_t$ are sufficient to predict its future values.

This assumption allows us to characterize the Frisch demands for the optimal labour market participation decision and the fraction of time spent at work. For notational convenience, define the vector ψ_{nt} as $\psi_{nt}\equiv(z'_{nt}, v_n\eta_n\lambda_t\omega_t)'$, where $z_{nt}=(l_{n,t-\rho},\ldots,l_{n,t-1},x'_{nt})'$. It follows from (2.8) and (2.9) that l^*_{nt} and d^o_{nt} can be written as

$$l^*_{nt}=l^*(\psi_{nt}), \tag{2.15}$$

$$d^o_{nt}=d^o(\psi_{nt},\varepsilon_{1nt}). \tag{2.16}$$

Likewise, the conditional valuation functions can be written as a function of the vector ψ_{nt} as $V_{0nt}\equiv V_0(\psi_{nt})$ and $V_{1nt}\equiv V_1(\psi_{nt})$.

To see the effects of aggregate shocks on individual choices, consider two individuals n and m who face the aggregate prices $\lambda_t\omega_t$ and $\lambda_s\omega_s$ at dates t and s, respectively, and who have the same labour supply histories and time-varying characteristics $z_{nt}=z_{ms}$, where $z_{nt}=l_{n,t-\rho},\ldots,l_{n,t-1},x'_{nt})'$ and z_{ms} is defined analogously. It follows from (2.15) that the fraction of time that individuals n and m optimally choose to spend at work at dates t and s, respectively, are the same if and only if the product of their time-invariant individual-specific characteristics and the aggregate prices are also equal

$$l^*_{nt}=l^*_{ms} \qquad \text{if and only if } (z'_{nt}, v_n\eta_n\lambda_t\omega_t)'=(z'_{ms}, v_m\eta_m\lambda_s\omega_s)'. \tag{2.17}$$

2. Note that the effects of intertemporal nonseparabilities in preferences and wages are captured by the term $\partial[V_{0n,t+1}+p_{n,t+1}(V_{1n,t+1}-V_{0n,t+1})]/\partial l_{nt}$. There are no separate terms showing the effect of hours worked today on future wages in (2.14) because the conditional valuation functions in (2.8) and (2.9) are defined to include the future wage benefits from current labour supply choices.

If, in addition, individuals are hit with the same value of the idiosyncratic disturbance ε_{1nt} that affects the fixed costs of participation, then their participation decision would also match. A similar argument can be made for the optimal consumption choice.

These results provide a convenient way to account for the effects of aggregate shocks. In particular, they say that the behaviour of an individual m in some future period s who responds to an aggregate shock $\lambda_s \omega_s$ can be inferred by forming an appropriate comparison group of choices actually observed in the earlier period t.[3] Rather than estimate the entire probability distribution, however, we use simulation techniques to simulate a hypothetical sequence of aggregate shocks for each observation, computing nonparametric estimates of their behavioural responses to the simulated shock sequence, and substituting it for their actual responses into sample moments derived from the Euler and participation conditions.

The only remaining issue is whether the first-order Markov process assumed for $\lambda_t \omega_t$ is consistent with the solution to the social planner's problem described by equation (2.5). Appendix A establishes this by construction. In particular, Appendix A.1 exhibits an exogenous dividend process for a physical asset such that Assumption 1 is satisfied for the solution to the social planner's for the economy with the primitives defined by equations (2.1) through (2.3).

2.3. *Overview of the empirical analysis*

This framework is amenable to a multi-stage estimation strategy. First, there is contemporaneous separability of consumption from labour supply in the utility function. Second, wages are assumed to be noisy measures of the individual-specific marginal products of labour, which are determined as a function of variables that are known at time t such as past labour market participation and the number of hours worked plus individual characteristics. Provided the measurement error in wages is uncorrelated with current and past labour supply choices, the consumption and wages equations can be estimated separately from the hours and participation equations to provide estimates of the determinants of household consumption and the effects of past labour market experience on individual wages.

The representation for individuals' valuation functions defined by (2.8) and (2.9) imply that the fixed costs of participation can be recovered from a model in which the income generated by the decision to work is evaluated using the product of the common shadow value of consumption λ_t and the time-invariant individual-specific effect η_n. However, the existence of fixed costs of participation and the effect of endogenous labour market participation on the optimal choice of hours implies that the techniques developed for dynamic discrete choice models must be used to estimate the participation and hours conditions. In principle, one could follow the approach in Miller (1984), Wolpin (1984), Pakes (1986), Rust (1987), and others and use maximum likelihood estimation (ML). This involves deriving the valuation function as a mapping of the state and parameter space to calculate the probability of the sample outcome.[4] Yet our problem allows hours worked in the past to enter both preferences and wages, accounts for the effects of aggregate shocks, and includes idiosyncratic shocks in preferences that affect the utility associated with different consumption and labour supply choices. In contrast to many of the applications described in the literature, the computational costs of employing ML in this setting

3. Our assumptions imply that the product of the fixed effects $v_n \eta_n$ is dense in $[0, \infty)$.

4. See the expository surveys of Eckstein and Wolpin (1989b), Rust (1993) and Miller (1996).

would be prohibitive. For this reason, a conditional choice probability (CCP) estimator is used, extending its applicability to a class of models beyond those considered by Hotz and Miller (1993).

The remaining sections are organized as follows. The data are described in Section 3. Sections 4.1 and 4.2 provide estimates of the consumption and wage equations, and Section 4.3 discusses the estimates of the aggregate prices. Section 4.4 describes the estimation of the time-invariant individual-specific effects in preferences and wages. Section 5.1 derives alternative representations for individuals' valuation functions and the Euler equation that form the basis for the conditional choice probability estimator. Section 5.2 describes the derivation of the nonparametric estimates of the conditional choice probabilities. This section also presents some diagnostics about how well our model fits the data by examining the behaviour of the nonparametric estimates. Section 6 is devoted to the estimation of the participation and hours conditions. Section 6.1 parameterizes the remaining components of the utility function, and imposes a distributional assumption on the idiosyncratic shocks to the fixed costs of participation. Section 6.2 uses the alternative representations of the Euler and participation conditions together with the parametric assumptions in Section 6.1 to formulate orthogonality condition estimators with and without aggregate shocks, while Section 6.3 reports our empirical findings.

3. DATA

The data consist of observations on wives and female household heads taken from the Michigan Panel Study of Income Dynamics (PSID) Family-Individual File, Waves I through XIX, for 1967 to 1985. The main advantage of working with the Family-Individual File is that it contains a separate record for each member of all households in the survey in any given year. Consequently, one can track more easily the behaviour of married women who may not have been the household head as well as the behaviour of married or unmarried women who were heads of their own households.

There are 20,437 individuals included in the nineteen-year Family-Individual Respondents File of the PSID. Appendix B describes how we obtained an initial sample of female heads or wives who did not have missing data on any of the variables listed there. The sample on which our estimation results is obtained by imposing the following additional selection criteria: (i) individuals be members of households that had not been surveyed as part of the nonrandom U.S. census sample in 1968, (ii) be a member of a PSID household for at least six consecutive years during the period 1968 to 1985, (iii) participate at least two (not necessarily consecutive) years in the labour market between 1972 and 1985, and (iv) the age of the individual be greater or equal to the value of an experience variable that is defined as the number of years of schooling plus ten.

The requirement that the individual participate at least twice in the labour market allows us to estimate the wage equation in differences. The requirement that the individual has been a member of a PSID household for at least six (consecutive) years allows us to use information on the individual's past labour supply and participation behaviour. Our main sample contains observations on 2169 women for the sample period 1973 to 1985. Its characteristics are displayed in Table I, where the number of observations by year refer to observations on individuals who were in the sample in that year.[5]

5. In this table, the sample mean of the ages of women included in the sample in any given year does not exactly increase by one every year. The reason is that not all the women in our sample are observed every year.

TABLE I

Characteristics of the main sample

	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
No. of observations	1118	1202	1284	1375	1456	1542	1617	1701	1773	1822	1902	1980	2049
Demographic Data													
Age	42·0	42·0	42·0	42·0	42·0	42·0	42·0	42·0	42·0	43·0	43·0	43·0	43·0
Number of children	1·4	1·4	1·3	1·3	1·3	1·2	1·2	1·2	1·1	1·1	1·1	1·1	1·1
Number in family	3·5	3·4	3·4	3·3	3·3	3·2	3·2	3·2	3·1	3·1	3·1	3·1	3·0
Prop. married	0·80	0·80	0·78	0·78	0·77	0·77	0·78	0·77	0·76	0·75	0·75	0·75	0·75
Prop. white	0·89	0·89	0·89	0·90	0·90	0·90	0·90	0·90	0·90	0·90	0·90	0·90	0·90
Educational level	4·4	4·5	4·5	4·5	4·6	4·6	4·7	4·7	4·7	4·7	4·8	4·9	4·9
Hours, Earnings, and Participation													
Prop. who worked	0·72	0·70	0·73	0·73	0·70	0·72	0·73	0·74	0·73	0·72	0·72	0·75	0·74
Annual hours	1,392	1,376	1,327	1,333	1,373	1,374	1,382	1,396	1,416	1,431	1,465	1,513	1,548
Hourly earnings	7·5	7·4	7·0	7·2	7·6	7·4	7·6	7·2	7·2	7·0	7·6	7·2	7·5
Food Consumption, Income, and Housing													
Food consumption	—	5,201	4,807	4,575	4,590	4,841	4,760	4,254	4,328	4,124	4,107	4,226	4,136
Income	32,164	31,336	30,310	31,311	31,226	31,767	31,952	31,717	30,886	30,325	31,393	32,076	32,415
House value	56,774	56,440	57,987	61,319	65,642	70,955	73,053	70,717	67,607	65,681	65,911	67,385	69,529
Rent value	3,002	2,919	3,073	3,085	2,938	3,111	3,022	3,000	3,031	2,909	3,048	3,107	3,202
Prop. homeowners	0·78	0·77	0·76	0·76	0·77	0·76	0·77	0·77	0·76	0·76	0·75	0·75	0·75
Prop. renters	0·20	0·20	0·21	0·21	0·21	0·21	0·20	0·20	0·21	0·21	0·22	0·22	0·22

The variables used in the empirical study are l_{nt}, the annual hours of work by individual n at date t; $\tilde{w}_{nt}$, her reported real average hourly earnings at t; c_{nt}, real household food consumption expenditures; FAM_{nt}, the number of household members; 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 education of the individual at time t; NE_{nt}, NC_{nt}, SO_{nt}, which are region dummies for northeast, north-central, and south, respectively; MAR_{nt} denoting whether a woman is married or not; and other variables denoting the woman's race, whether her household was a homeowner or renter, and house value.

4. CONSUMPTION AND WAGES

4.1. *Consumption*

The estimation of the marginal utility of consumption equation requires a parameterization for the component of preferences describing the effect of current consumption on utility. We assume that the idiosyncratic shock ε_{2nt} is uncorrelated with individual-specific characteristics, that is, $E(\varepsilon_{2nt} \mid x_{nt}) = 0$, and let

$$U_2(c_{nt}, x_{nt}, \varepsilon_{2nt}) = \exp(x'_{nt} B_2 + \varepsilon_{2nt}) c^{\alpha}_{nt}/\alpha, \qquad \alpha \leqq 1. \tag{4.1}$$

This specification implies that food consumption is separable from other types of consumption (such as consumption from the services of durable goods) as well as leisure in preferences. Thus, food consumption and other types of consumption are related only through the constraints imposed by the economy-wide feasibility condition in (2.4). More specifically, the assumption of Pareto optimality implies that all the interactions occur through the social planner's weight for the individual household.

Using this parameterization, the condition that characterizes the optimal consumption allocation is

$$\exp(x'_{nt} B_2 + \varepsilon_{2nt}) c^{\alpha-1}_{nt} = \eta_n \lambda_t. \tag{4.2}$$

Let Δ denote the first-difference operator. Taking logarithms of both sides of this expression, differencing and rearranging implies

$$(1-\alpha)^{-1} \Delta \varepsilon_{2nt} = \Delta \ln(c_{nt}) - (1-\alpha)^{-1} \Delta x'_{nt} B_2 + (1-\alpha)^{-1} \Delta \ln(\lambda_t). \tag{4.3}$$

Define time dummies, D_t, for $t = 2, \ldots, T$, with the $(T-1)$ dimensional vector of coefficients, Γ_1, which represent the variation in the shadow value of consumption, namely, $(1-\alpha)^{-1} \Delta \ln(\lambda_t)$. The $(K+T-1)$ dimensional vector of identified parameters in (4.3) is given by $\Theta_1 = ((1-\alpha)^{-1} B'_2, \Gamma'_1)'$. Define the $T-1$ dimensional vector Y_n, the $(T-1) \times (K+T-1)$ dimensional matrix X_n, and the square $T-1$ matrix W_n as

$$Y_n = (\Delta \ln(c_{n2}), \ldots, \Delta \ln(c_{nT}))',$$

$$X_n = \begin{bmatrix} \Delta x_{n2} & D_2 & \cdots & 0 \\ \vdots & & & \vdots \\ \Delta x_{nT} & 0 & \cdots & D_T \end{bmatrix},$$

$$\varepsilon_n \equiv ((-\alpha)^{-1} \Delta \varepsilon_{2n2}, \ldots, (1-\alpha)^{-1} \Delta \varepsilon_{2nT})',$$

$$W_n \equiv E[(Y_n - X'_n \Theta_1)(Y_n - X'_n \Theta_1)' \mid X_n].$$

It is easy to see that $E(\varepsilon_n | X_n)=0$, and that the GMM estimator with lowest asymptotic covariance within this class is

$$\Theta_1^N=[N^{-1}\sum_{n=1}^N X_n'\tilde{W}_n^{-1}X_n]^{-1}[N^{-1}\sum_{n=1}^N X_n'\tilde{W}_n^{-1}Y_n], \tag{4.4}$$

where $\tilde{W}_n$ is any consistent estimate of W_n. (See Robinson (1987) for example.) The asymptotic covariance of Θ_1^N is $E[X_n'W_n^{-1}X_n]^{-1}$. Since W_n is not restricted to be diagonal, there is arbitrary correlation across the elements of $\tilde{\varepsilon}_n$. However, $\tilde{\varepsilon}_n$ is assumed to be homoscedastic, which implies that $E[\tilde{\varepsilon}_n\tilde{\varepsilon}_n']$, is independent of n. Hence, a consistent estimator of W_n can be obtained by forming cross-sectional averages of the outer product of the residuals defined as $Y_n-X_n'\Theta_1^N$.

The estimates of the consumption equation are based on the main sample of women for the years 1975 to 1985. Since the PSID data tapes do not contain information on the different components of consumption expenditures for the calendar year 1972, the consumption measure used in this study is not defined for the years 1972 and 1973. The reason is that consumption expenditures in a given year are measured by taking 0·25 of the value of this variable for year $t-1$ and 0·75 of its value for year t. The requirement that the consumption equation be estimated in differences implies that the sample period is further reduced by an additional year. The number of observations used in estimation is 18,501. The elements of x_{nt} are defined as FAM_{nt}, YKID_{nt}, OKID_{nt}, AGE_{nt}^2, NC_{nt} and SO_{nt}.

The estimates in Table II show that consumption increases with family size and children consume less than adults, since the coefficients on children between the ages of zero and six, and children between six and fourteen are negative and smaller in absolute value than the coefficient on total household size. Furthermore, the behaviour of consumption over the life-cycle is concave since the coefficient on age squared is negative. All the coefficients characterizing these effects are highly significant. On the other hand, the regional dummies are imprecisely estimated. Estimates of the ratios of the shadow value of consumption, namely $(1-\alpha)^{-1}\Delta \ln(\lambda_t)$, show that there is significant variation in these time effects; the test statistic for the null hypothesis that $(1-\alpha)^{-1}\Delta\ln(\lambda_t)=(1-\alpha)^{-1}\Delta\ln(\lambda_{t-1})$ for $t=1976,\ldots, 1985$ is 316. Under the null, it would be distributed as χ^2 with 10 degrees of freedom, implying rejection of the null at conventional significance levels.

4.2. *Wages*

The wage rate for a woman is given by the individual-specific marginal product of labour defined as $w_{nt}=\omega_t v_n\gamma(z_{nt})$, where $\gamma(z_{nt})$ is the efficiency units of labour for individual n at time t. Assume that $\gamma(z_{nt})$ has the representation

$$\gamma(z_{nt})\equiv v_n \exp[\sum_{s=1}^{\rho}(\gamma_{1s}l_{n,t-s}+\gamma_{2s}d_{n,t-s})+x_{nt}'B_3], \tag{4.5}$$

where v_n is an unobserved individual-specific effect.

Observed wages are assumed to be noisy measures of women's marginal product in the market sector as

$$\tilde{w}_{nt}=\omega_t v_n\gamma(z_{nt})\exp(\tilde{\varepsilon}_{nt}). \tag{4.6}$$

The multiplicative error term in (4.6) is conditionally independent over people, the covariates in the wage equation and the labour supply decisions. Since the error term enters wages log-linearly and is independent of any variables known at time t, including current

TABLE II

The consumption equation†

$\ln(c_{nt}) = 1/(1-\alpha)[x'_{nt}B_2 - \ln(\eta_n\lambda_t) + \varepsilon_{2nt}]$

Variable	Parameter	Estimate	Variable	Estimate
Socioeconomic variables			Aggregate prices	
ΔFAM_{nt}	$(1-\alpha)^{-1}B_{21}$	0·14 (0·005)	$(1-\alpha)^{-1}\Delta\ln(\lambda_9)$	0·02 (0·01)
ΔYKID_{nt}	$(1-\alpha)^{-1}B_{22}$	−0·05 (0·004)	$(1-\alpha)^{-1}\Delta\ln(\lambda_{10})$	−0·005 (0·001)
ΔOKID_{nt}	$(1-\alpha)^{-1}B_{23}$	−0·03 (0·006)	$(1-\alpha)^{-1}\Delta\ln(\lambda_{11})$	−0·06 (0·01)
ΔAGE^2_{nt}	$(1-\alpha)^{-1}B_{24}$	−0·0006 (0·0001)	$(1-\alpha)^{-1}\Delta\ln(\lambda_{12})$	−0·12 (0·01)
			$(1-\alpha)^{-1}\Delta\ln(\lambda_{13})$	−0·05 (0·01)
Region dummies				
ΔNC_{nt}	$(1-\alpha)^{-1}B_{26}$	−0·012 (0·014)	$(1-\alpha)^{-1}\Delta\ln(\lambda_{14})$	−0·003 (0·01)
ΔSO_{nt}	$(1-\alpha)^{-1}B_{27}$	−0·01 (0·014)	$(1-\alpha)^{-1}\Delta\ln(\lambda_{15})$	−0·02 (0·01)
			$(1-\alpha)^{-1}\Delta\ln(\lambda_{16})$	0·003 (0·01)
			$(1-\alpha)^{-1}\Delta\ln(\lambda_{17})$	−0·05 (0·01)
			$(1-\alpha)^{-1}\Delta\ln(\lambda_{18})$	−0·08 (0·01)
			$(1-\alpha)^{-1}\Delta\ln(\lambda_{19})$	−0·04 (0·01)

† Standard errors in parentheses.

labour supply l_{nt}, this representation for wages allows us to circumvent the selection problem referred to in Section 2.1.

The wage equation used in estimation is obtained by substituting for $\gamma(z_{nt})$ in (4.6), taking logarithms of both sides of this equation, and differencing as

$$\Delta\tilde{\varepsilon}_{nt} = \Delta\ln(\tilde{w}_{nt}) - \Delta\ln(\omega_t) - \textstyle\sum_{s=1}^{\rho}(\gamma_{1s}\Delta l_{n,t-s} + \gamma_{2s}\Delta d_{n,t-s}) - \Delta x'_{nt}B_3. \tag{4.7}$$

This equation is estimated using generalized method moments (GMM) estimation with the optimal instruments, using the conditional independence of the disturbances with the covariates. A time dummy for each year denoted D_t for $t = 2, \ldots, T$ was included to capture the change in wages due to aggregate or common effects, $\Delta\ln(\omega_t)$. Let Γ_2 denote the $(T-1)$-dimensional vector of coefficients corresponding to these time dummies. Recall that the dimension of B_3 is K, and there are ρ lagged values of labour supplies and participation. Then the parameter vector to be estimated is the $(2\rho+K+T-1)$-dimensional vector, $\Theta_2 \equiv (\gamma_{11}, \ldots, \gamma_{1\rho}, \gamma_{21}, \ldots, \gamma_{2\rho}, B'_3, \Gamma'_2)'$. Define the $T-1$ dimensional vector Y_n, the $(T-1)\times(2\rho+K+T-1)$ matrix X_n, and the square $T-1$ matrix W_n as follows

$$Y_n \equiv (\Delta\ln(\tilde{w}_{n2}), \ldots, \Delta\ln(\tilde{w}_{nT}))',$$

$$\tilde{\varepsilon}_n \equiv (\Delta\tilde{\varepsilon}_{n2}, \ldots, \Delta\tilde{\varepsilon}_{nT})'$$

$$X_n \equiv \begin{bmatrix} \Delta l_{n1} & \ldots & \Delta l_{n,2-\rho} & \Delta d_{n1} & \ldots & \Delta d_{n,2-\rho} & \Delta x_{n2} & D_2 & \ldots & 0 \\ \vdots & & & & & & & & & \vdots \\ \Delta l_{n,T-1} & \ldots & \Delta l_{n,T-\rho} & \Delta d_{n,T-1} & \ldots & \Delta d_{n,T-\rho} & \Delta x_{nT} & 0 & \ldots & D_T \end{bmatrix}.$$

Thus, the wage equation can be estimated in the same way as the consumption equation.

TABLE III

The wage equation†

$\ln(w_{nt}) = \ln(\omega_t) + \ln(v_n) + \sum_{s=1}^{\rho} (\gamma_{1s} l_{n,t-s} + \gamma_{2s} d_{n,t-s}) + x'_{nt} B_3$

Variable	Parameter	(i) Estimate	(ii) Estimate	Variable	(i) Estimate	(ii) Estimate
Lags of hours worked				Aggregate wages		
$\Delta l_{n,t-1}$	γ_{11}	0·0002	0·0002	$\Delta \ln(\omega_7)$	—	0·045
		(0·00001)	(0·5E−5)			(0·016)
$\Delta l_{n,t-2}$	γ_{12}	0·0001	0·0001	$\Delta \ln(\omega_8)$	—	−0·003
		(0·00001)	(0·000006)			(0·021)
$\Delta l_{n,t-3}$	γ_{13}	0·00008	0·00006	$\Delta \ln(\omega_9)$	−0·013	−0·038
		(0·00001)	(0·000006)		(0·022)	(0·011)
$\Delta l_{n,t-4}$	γ_{14}	0·00006	0·00005	$\Delta \ln(\omega_{10})$	−0·04	0·02
		(0·00001)	(0·000005)		(0·02)	(0·01)
$\Delta l_{n,t-5}$	γ_{15}	0·00002	—	$\Delta \ln(\omega_{11})$	0·11	0·074
		(0·00001)			(0·02)	(0·01)
$\Delta l_{n,t-6}$	γ_{16}	0·00002	—	$\Delta \ln(\omega_{12})$	0·05	0·03
		(0·00001)			(0·02)	(0·02)
Lags of participation				$\Delta \ln(\omega_{13})$	0·05	0·02
					(0·02)	(0·01)
$\Delta d_{n,t-1}$	γ_{21}	−0·11	−0·09	$\Delta \ln(\omega_{14})$	−0·01	−0·03
		(0·021)	(0·01)		(0·02)	(0·01)
$\Delta d_{n,t-2}$	γ_{22}	−0·10	−0·07	$\Delta \ln(\omega_{15})$	0·05	0·02
		(0·021)	(0·01)		(0·02)	(0·01)
$\Delta d_{n,t-3}$	γ_{23}	−0·10	−0·09	$\Delta \ln(\omega_{16})$	−0·02	−0·01
		(0·02)	(0·01)		(0·02)	(0·01)
$\Delta l_{n,t-4}$	γ_{24}	−0·10	−0·07	$\Delta \ln(\omega_{17})$	0·07	0·04
		(0·02)	(0·01)		(0·02)	(0·01)
$\Delta d_{n,t-5}$	γ_{25}	−0·05	—	$\Delta \ln(\omega_{18})$	−0·01	−0·02
		(0·02)			(0·02)	(0·01)
$\Delta d_{n,t-6}$	γ_{26}	−0·03	—	$\Delta \ln(\omega_{19})$	0·082	0·05
		(0·02)			(0·02)	(0·02)
Socioeconomic variables						
ΔAGE_{nt}^2	B_{31}	−0·0005	−0·0002			
		(0·0001)	(0·00005)			
$\Delta(\text{AGE}_{nt} \times \text{EDU}_{nt})$	B_{32}	0·0001	0·0002			
		(0·0004)	(0·0001)			

† Standard errors in parentheses.

The empirical findings are reported in Table III, which reports estimates obtained using the main sample of women described in Section 3 as well as estimates obtained using a smaller sample of women who had been a member of a PSID household for eight consecutive years during the period 1968 to 1985. The latter estimates are reported in column (i) of Table III while the former are reported in column (ii). The elements of x_{nt} are defined as AGE_{nt}^2 and $\text{AGE}_{nt} \times \text{EDU}_{nt}$. Table III shows that the estimates obtained from the larger sample lie within two standard deviations of those from the smaller sample, and as we illustrate below, the predictions implied by the two sets of results are quantitatively similar. This suggests that the more demanding sample selection criteria of the smaller sample, which requires respondents to remain in the panel two extra periods, combined with estimating the extra two lags this smaller sample permits, do not affect the empirical results much.

For both samples, all the parameter estimates for the personal coefficients are significant. The coefficient on age squared is negative, implying that wages are concave in age.

The coefficient on age times education is positive, implying that the effects of education are magnified with age, and estimated significantly for the larger sample. The null hypothesis of no aggregate effects, with $\Delta \ln(\omega_t) = \Delta \ln(\omega_{t-1})$ for $t = 1974, \ldots, 1985$, is rejected. The test statistic for this hypothesis is 148 which under the null, is distributed as χ^2 with 12 degrees of freedom.

The estimated 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 0·2, but the wage elasticity of hours lagged two years is 0·05. Also the further back the working experience is, the less the timing matters: an extra hour worked one year in the past has twice the effect on current wages as an extra hour worked two years in the past, but the difference between the wage effects of an extra hour worked three and four years in the past, respectively, is less than 50%. We note that this finding is confirmed, qualitatively, by Miller and Sanders (1997), who also estimate a wage equation in which current wages depend on hours worked in the past.

As 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 $\bar{l}_t$. With $\rho = 4$, this is given by $\sum_{s=1}^{4} [\gamma_{1s}\bar{l}_{t-s} + \gamma_{2s}] = 0{\cdot}25$. 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 $t-\rho$ to $t-2$ but works the sample average at $t-1$ is $\gamma_{11}\bar{l}_{t-1} + \gamma_{21} = 0{\cdot}20$. Conversely, an individual who had not been working in $t-2$, $t-3, \ldots$, would have had to work less than 500 hours in period $t-1$ for the negative effects of participation to reduce wages in period t relative to period $t-1$.

4.3. *Aggregate effects*

In later parts of this paper, the estimated aggregate effects from the wage and consumption equations are used as part of the estimation of the hours and participation equations with aggregate shocks. As a preliminary way of examining the implications of the model, we compare their behaviour with more traditional measures derived from time series data. One advantage of using the estimated dummies over more traditional measures as a guide for examining past economic performance is that, unlike the aggregate series, these statistics are purged of demographically induced effects (by conditioning upon them in estimation).

Figures 1 and 2 track the estimated aggregate shock series over the 1975–1985 decade, along with four other series. Figure 1 relates two series of consumption growth, that is, *per capita* nondurable goods and services, and *per capita* food consumption in the PSID sample, to our estimates of the growth rate of the shadow value of food consumption, λ_t/λ_{t-1}. The coherence between the three series is striking, especially between the PSID food consumption data and estimated prices. In years when growth in *per capita* food consumption and more generally nondurable consumption was high, the state-contingent price of consumption declined and vice versa. Thus the recession of 1980 through 1982 is marked by low consumption and high values of λ_t/λ_{t-1}. In Figure 1 there is little support for the hypothesis that conditioning on demographics has a noticeable effect, although composition effects probably explain why the broadly defined nondurables and services series exhibits much less variation than the other two.

A different story emerges from Figure 2, which plots the differences in aggregate wages, sample wage rates, and the estimates of wages to standardized labour units, ω_t/ω_{t-1}. While the pattern of very high coherence is repeated, the aggregate wage series has

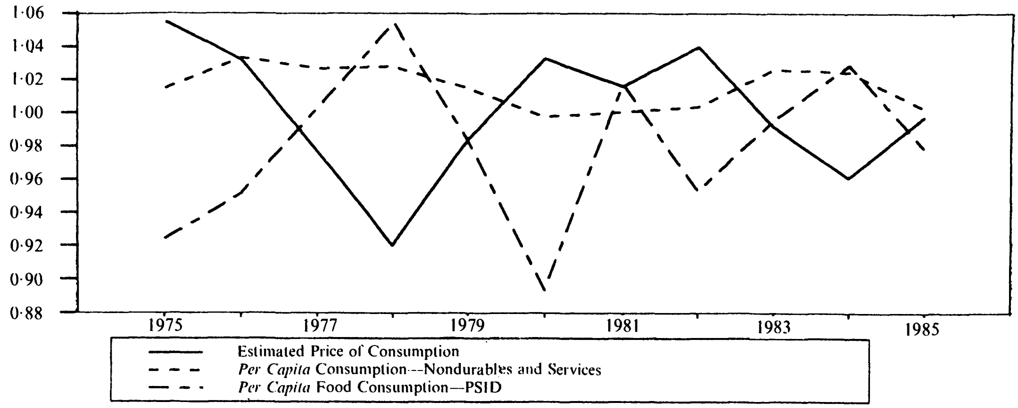

FIGURE 1

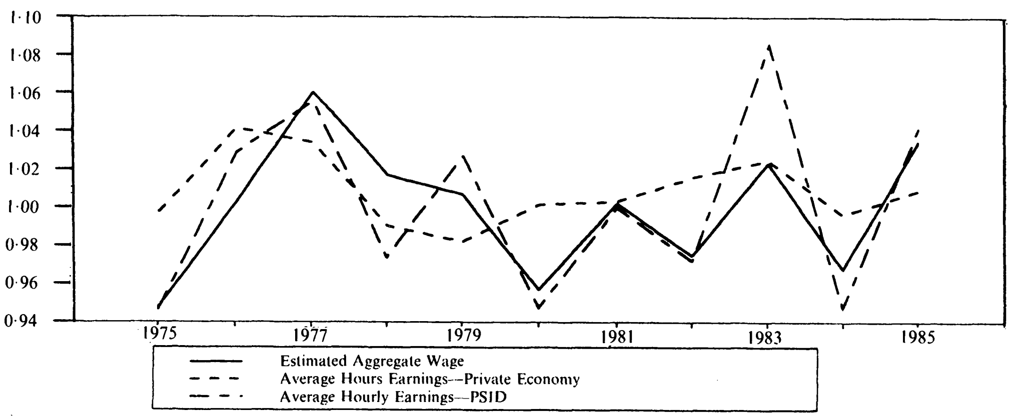

FIGURE 2

the lowest variance, and the ω_t/ω_{t-1} sequence has the highest. This suggests that ignoring heterogeneity across skill levels, experience on the job, and the endogeneity of workforce participation, may disguise cyclical movements in wages. Comparing the diagrams, movements in real wages and contingent prices are positively correlated over this decade (wage changes tending to lead price adjustments), which would imply wages are countercyclical.

4.4. *Individual-specific effects*

Estimation of the participation and hours conditions also requires estimates of the individual-specific effects $\eta_n v_n$ and η_n. Two approaches are employed for estimating these quantities.

The first method is based on a traditional fixed effects estimator, which estimates $v_n \eta_n$ and η_n using the estimated residuals from the log-linear versions of the consumption and the wage equations defined by (4.3) and (4.7), respectively. This estimator allows for unobserved heterogeneity arising from permanent unobserved individual-specific characteristics. Define the quantities ϕ_{1n} and ϕ_{2n} as

$$\phi_{1n}\equiv\sum_{t\in T_1}[\ln(w_{nt})-\ln(\omega_t)-\sum_{s=1}^{\rho}(\gamma_{1s}l_{n,t-s}+\gamma_{2s}d_{n,t-s})-x'_{nt}B_3]/T_1$$
$$+\sum_{t\in T_2}[\ln(c_{nt})-(1-\alpha)^{-1}x'_{nt}B_2+(1-\alpha)^{-1}\ln(\lambda_t)]/T_2, \quad (4.8)$$

$$\phi_{2n}\equiv\sum_{t\in T_2}[\ln(c_{nt})-(1-\alpha)^{-1}x'_{nt}B_2+(1-\alpha)^{-1}\ln(\lambda_t)]/T_2, \quad (4.9)$$

where T_1 and T_2 denotes the set of time periods for which the wage equation and the marginal utility of consumption equation is estimated. The fixed effects estimators of $v_n\eta_n$ and η_n are obtained as simple time averages of the estimated residuals of the consumption and wage equations, which correspond to sample counterparts of ϕ_{1n} and ϕ_{2n} defined in (4.8) and (4.9). One problem with using this estimator is that it is subject to small sample bias arising from the short panel length. On this point, Hotz and Miller (1988) provide Monte Carlo evidence based on a model with an index function for the participation decision as well as a behavioural equation determining the optimal interior choice of hours. They show that the small sample bias caused by using such a fixed-effects estimator on the remaining parameter estimates is quite small for moderate to large sample sizes.[6]

The second method that is used to estimate the individual-specific effects achieves consistency of the cross section of the panel data set but it can only deal with any unobserved permanent characteristic that can be written as a nonparametrically estimated function of observables. This estimator may be viewed as a nonparametric extension of the approach in MaCurdy (1981), pp. 1066–1069. To implement this estimator, suppose that $v_n\eta_n$ and η_n can be expressed as functions of a Q-dimensional vector of regressors x_n, which is assumed to represent the permanent characteristics of individual n.

Assumption 2. The vector x_n is observed and satisfies the conditions $E[x_n(\phi_{1n}-v_n\eta_n)]=0$ and $E[x_n(\phi_{2n}-\eta_n)]=0$.

This assumption allows us to estimate $v_n\eta_n$ and η_n in terms of nonlinear regressions of the variables ϕ_{1n} and ϕ_{2n} on the components of the vector x_n. Let δ_N denote the bandwidth of the proposed kernel estimator and J the normal kernel defined on $\Re^Q$. Our estimator for $v_n\eta_n$ is defined as

$$v_n^N\eta_n^N=\frac{\sum_{m=1,m\neq n}^{N}\phi_{1m}J[\delta_N^{-1}(x_m-x_n)]}{\sum_{m=1,m\neq n}^{N}J[\delta_N^{-1}(x_m-x_n)]}. \quad (4.10)$$

The nonparametric estimate for η_n is similarly defined. Actually, the application does not rely on the existence of the variables ϕ_{1n} and ϕ_{2n}. Instead it exploits the first-order condition for consumption and the wage equation to generate a $N^{1/2}$-uniformly consistent estimator of ϕ_{in} denoted ϕ_{in}^N for $i=1, 2$. However, this approximation has no asymptotic consequences for the estimator of $v_n\eta_n$ or η_n which converges more slowly than $N^{1/2}$, in any case. The components of the regressor x_n consist of race plus a number of characteristics associated with individuals' achievements or situation by different ages. These include such observable demographic characteristics as marital status, the age distribution of children,

6. This is in contrast to Heckman (1981), whose Monte Carlo study suggests substantial bias when modelling the behaviour of an indicator variable which depends on its own lagged values as well as an unobserved fixed effect. Honore (1992) has developed an estimator to avoid the sample small bias induced by using a traditional fixed effects estimator in a discrete choice setting. However, his method cannot be applied here because it requires preferences to be both contemporaneously separable with respect to consumption and leisure as well as additively separable over time.

homeownership, educational level, and geographical location. The bandwidths associated with each characteristic are computed as the sample variance of each characteristic.

5. CONDITIONAL CHOICE PROBABILITIES

The CCP estimator forms an alternative representation for the conditional valuation functions that enter individuals' optimizing conditions by multiplying current utilities, evaluated at the 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 non-parametrically 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 show that a property enjoyed by our model, finite time dependence, allows us to extend the domain of problems to which the CCP may be applied.[7] As developed by Hotz and Miller (1993), the CCP estimator assumes agents live in a stationary environment. The CCP estimator is initially developed for the stationary case. In Section 6, it is shown how this assumption can be relaxed to allow for aggregate shocks.

5.1. *Conditional valuation 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 one outcome vs. the other. By Proposition 1 of Hotz and Miller (1993), there exists a mapping q: $[0,1]\to\mathfrak{R}$, which gives the difference between the conditional valuation functions as a function of the conditional choice probabilities, and $\varphi_k:[0,1]\to\mathfrak{R}$, which measures the expected value of the unobservables in current utility, conditional on an action $k\in\{0,1\}$ being taken

$$q(p(\psi_{nt}))=V_1(\psi_{nt})-V_0(\psi_{nt}), \tag{5.1}$$

$$\varphi_k(p(\psi_{nt}))=E[\varepsilon_{knt}\mid\psi_{nt},d^o_{nt}=k],\qquad k\in\{0,1\}. \tag{5.2}$$

In our application, the property of finite dependence is exploited to derive an alternative representation for the elements of $q(p(\psi_{nt}))$. To illustrate this, define the $(\rho+K)$-dimensional vectors $z^{(s)}_{0nt}$ and $z^{(s)}_{1nt}$ as

$$z^{(s)}_{0nt}\equiv(l_{n,t-\rho+s},\ldots,l_{n,t-1},0,\ldots,0,x'_{n,t+s})', \tag{5.3}$$

$$z^{(s)}_{1nt}\equiv(l_{n,t-\rho+s},\ldots,l_{n,t-1},l^*_{nt},\ldots,0,x'_{n,t+s})', \tag{5.4}$$

for $s=1,\ldots,\rho$, where l^*_{nt} is the fraction of time a woman chooses to spend at work, conditional on participating. Similarly define $\psi^{(s)}_{0nt}\equiv(z^{(s)'}_{0nt},v_n\eta_n)'$ and $\psi^{(s)}_{1nt}\equiv(z^{(s)'}_{1nt},v_n\eta_n)'$. The state vector $\psi^{(s)}_{0nt}$ is the state for a woman at date $t+s$ who has accumulated the history $(l_{n,t+s-\rho},\ldots,l_{n,t-1})'$ up to period t and then chooses not to participate at date t and for $s-1$ periods following period t. The state vector $\psi^{(\rho)}_{0nt}$ corresponds to the labour market history in which the woman does not participate between t and $t+\rho$. Likewise,

7. Other assumptions that ease the computational burden of the CCP estimator include the terminal state property (Hotz and Miller (1993)), stationary Markov models (Aguirregabiria (1994)), and simulation techniques (Hotz, Miller, Sanders and Smith (1994)).

$\psi_{1nt}^{(s)}$ is the state vector for a woman at time $t+s$ who accumulates the history $(l_{n,t-\rho}, \ldots, l_{n,t-1})'$ up to time t, chooses to participate at date t, and then chooses not to participate for the $s-1$ periods following period t.

The results in equations (5.1) and (5.2) are used to establish an alternative representation for the conditional valuation function V_{knt} for the finite dependence case.

Proposition 1. *Define*:

$$u_k(\psi_{nt}) \equiv \begin{cases} U_1(z_{nt}, 0), & \text{for } k=0; \\ U_0(z_{nt}) + U_1(z_{nt}, l_{nt}^*) + \eta_n \lambda_t w_{nt} l_{nt}^* & \text{for } k=1. \end{cases}$$

Let $p_{knt}^{(s)} \equiv p(\psi_{knt}^{(s)})$ *for* $k \in \{0, 1\}$ *and* $s=1, \ldots, \rho$. *For* $k \in \{0, 1\}$, *the conditional valuation functions can be expressed as*

$$\begin{aligned} V_k(\psi_{nt}) &= u_k(\psi_{nt}) \\ &\quad + E_t\{\textstyle\sum_{s=1}^{\rho} \beta^s [u_0(\psi_{knt}^{(s)}) + \varphi_0(p_{knt}^{(s)}) + p_{knt}^{(s)}(q(p_{knt}^{(s)}) + \varphi_1(p_{knt}^{(s)}) - \varphi_0(p_{knt}^{(s)}))] \\ &\quad + \beta^{\rho+1}[V_0(\psi_{knt}^{(\rho+1)}) + \varphi_0(p_{knt}^{(\rho+1)}) + p_{knt}^{(\rho+1)}(q(p_{knt}^{(\rho+1)}) \\ &\quad + \varphi_1(p_{knt}^{(\rho+1)}) - \varphi_0(p_{knt}^{(\rho+1)}))]\}. \end{aligned} \tag{5.5}$$

In equation (5.5), $u_k(\psi_{nt})$ is the current utility from undertaking the k-th action at date t, conditional on the state ψ_{nt}. The remaining terms in the first line of (5.5) show the utility that the individual would obtain, conditional on having the history described by $\psi_{knt}^{(s)}$. The term $u_0(\psi_{knt}^{(s)}) + \varphi_0(p_{knt}^{(s)})$ shows the utility to the individual from not participating at date $t+s$ while the next set of terms account for the fact that it may not be optimal to not participate at $t+s$. If the individual does participate with probability $p_{knt}^{(s)}$, she obtains the utility $q(p_{knt}^{(s)}) + \varphi_1(p_{knt}^{(s)}) - \varphi_0(p_{knt}^{(s)})$, which is just the difference between the value of participating vs. not participating at date $t+s$, conditional on the history $\psi_{knt}^{(s)}$. The terms in the third line of (5.5) are inconsequential from the point of view of the optimal choices because the histories $\psi_{knt}^{(\rho+1)}$ for $k \in \{0, 1\}$ are both equal to the $(\rho+K)$-dimensional vector $(0, \ldots, 0, x_{n,t+\rho}')'$ using the definitions in (5.3) and (5.4). Thus, the proposition shows that the optimality conditions for the labour supply decisions depend on evaluating the probabilities and payoffs for 2ρ dates into the future.

Proposition 1 can also be used to derive an alternative representation for the Euler equation by substituting (5.5) into (2.9) and noting that the derivative of the terms involving the histories $\psi_{knt}^{(\rho+1)}$ are zero.

$$\begin{aligned} \frac{\partial U_1(z_{nt}, l_{nt}^*)}{\partial l_{nt}} + \eta_n \lambda_t w_{nt} &= -\frac{E_t\{\sum_{s=1}^{\rho} \beta^s \partial[U_1(z_{1nt}^{(s)}, 0) + \varphi_0(p_{1nt}^{(s)})]\}}{\partial l_{nt}} \\ &\quad - E_t\left\{\sum_{s=1}^{\rho} \beta^s p_{1nt}^{(s)} \frac{\partial[q(p_{1nt}^{(s)}) + \varphi_(p_{1nt}^{(s)}) - \varphi_0(p_{1nt}^{(s)})]}{\partial l_{nt}}\right. \\ &\quad \left. + \sum_{s=1}^{\rho} \beta^s [q(p_{1nt}^{(s)}) + \varphi_1(p_{1nt}^{(s)}) - \varphi_0(p_{1nt}^{(s)})] \frac{\partial p_{1nt}^{(s)}}{\partial l_{nt}}\right\}. \end{aligned} \tag{5.6}$$

This representation illustrates the endogeneity of the labour force participation decision. The two terms on the left-side of (5.6), $\partial U_1(z_{nt}, l_{nt}^*)/\partial l_{nt} + \eta_n \lambda_t w_{nt}$, are the current marginal benefit from n working l_{nt}^* in period t (denominated in the numeraire). The first group of

ρ terms on the right-side of (5.6), each taking the form

$$\beta^s E_t[\partial(U_1(z^{(s)}_{1nt}, 0) + \varphi_0(p^{(s)}_{1nt}))/\partial l_{nt}],$$

show how marginal utility would be affected if the agent did not work in period $t+s$ (through the effect of additional hours worked at t on the utility from additional leisure at $t+s$). The other terms reflect the fact that not working throughout these periods is not (necessarily) an optimal strategy. Using the chain rule of differentiation, the effects of these terms can be distinguished. The first term, $E_t\{p^{(s)}_{1nt}\partial[q(p^{(s)}_{1nt}) + \varphi_1(p^{(s)}_{1nt}) - \varphi_0(p^{(s)}_{1nt})]/\partial l_{nt}\}$ measures the marginal utility differential (with respect to not participating at time t) multiplied by the conditional choice probability and then summed over the set of choices. The second term, $E_t\{[q(p^{(s)}_{1nt}) + \varphi_1(p^{(s)}_{knt}) - \varphi_0(p^{(s)}_{1nt})]\ \partial p^{(s)}_{1nt}/\partial l_{nt}\}$, accounts for variation in the participation probability induced in period $t+s$ by marginally adjusing labour supply at date t.

The proposition also provides a tractable representation of the participation condition. Using the fact that labour market histories matter only for ρ periods (so that the future utility conditional on the states $\psi^{(\rho+1)}_{0nt}$ and $\psi^{(\rho+1)}_{1nt}$ are equal) and taking the difference of (5.5) for $k=1$ and $k=0$, equation (2.10) can be rewitten as

$$\begin{aligned} q(p_{nt}) = {} & U_0(z_{nt}) + \eta_n \lambda_t w_{nt} l^*_{nt} + U_1(z_{nt}, l^*_{nt}) - U_1(z_{nt}, 0) \\ & + E_t\{\textstyle\sum_{s=1}^{\rho} \beta^s[U_1(z^{(s)}_{1nt}, 0) + \varphi_0(p^{(s)}_{1nt}) - U_1(z^{(s)}_{0nt}, 0) - \varphi_0(p^{(s)}_{0nt})] \\ & + \textstyle\sum_{s=1}^{\rho} \beta^s[p^{(s)}_{1nt}(q(p^{(s)}_{1nt}) + \varphi_1(p^{(s)}_{1nt}) - \varphi_0(p^{(s)}_{1nt})) \\ & - p^{(s)}_{0nt}(q(p^{(s)}_{0nt}) + \varphi_1(p^{(s)}_{0nt}) - \varphi_0(p^{(s)}_{0nt}))]\}. \end{aligned} \tag{5.7}$$

From (5.1), the left-side of (5.7) is the difference in conditional valuation functions for working vs. staying at home. The right-side partitions this difference into three expressions. The effects on current utility are captured by the right-side terms on the first line of equation (5.7). The terms in the middle line represent the difference in utility received at date $t+s$ from not having worked an extra period while the remaining terms are the respective selection correction factors. Specialized in this way, the participation equation illustrates how the conditional choice probabilities capture the dynamic features of this problem. If the subjective discount factor is set to zero and future consequences are ignored, the difference in the conditional valuation functions $q(p_{nt})$ simply equals the current utility differential from participating vs. staying at home, in other words, the familiar static model of random utility.[8] But supposing β is positive and labour force participation in the current period raises the probability of future participation, the final summation in equation (5.7) is positive, and the threshold current utility required to induce her to work is lower because of long term effects. Furthermore, for those who do participate, the marginal disutility of work is pushed beyond the point of offsetting the benefits of current wages, through greater time spent at work.

5.2. *Nonparametric estimation of the conditional choice probabilities*

The alternative representations for the Euler and participation conditions require estimates of the conditional choice probabilities. The probabilities p_{nt} are computed as nonlinear regressions of the participation index d_{nt} on the current state $\psi^N_{nt} \equiv (z'_{nt}, v^N_n \eta^N_n)'$, where the N superscript denotes an estimated quantity. Define $J[\delta^{-1}_N\ (\psi^N_{nt} - \psi^N_{mr})]$ as the normal

8. See Maddala (1983), for example.

kernel, where δ_N is the bandwidth associated with each argument. The nonparametric estimate of p_{nt} denoted p_{nt}^N is computed using the kernel estimator

$$p_{nt}^N = \frac{\sum_{m=1,m\neq n}^N \sum_{r=1}^\tau d_{mr} J[\delta_N^{-1}(\psi_{nt}^N - \psi_{mr}^N)]}{\sum_{m=1,m\neq n}^N \sum_{r=1}^\tau J[\delta_N^{-1}(\psi_{nt}^N - \psi_{mr}^N)]}. \tag{5.8}$$

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

$$d_{knt}^{(s)} = [1 - k(1 - d_{n,t-s})] \prod_{l=1}^{s-1} (1 - d_{n,t-l}), \qquad k \in \{0, 1\}.$$

Notice that $d_{1nt}^{(s)} = 1$ if the person participated at $t-s$ but then did not participate for $s-1$ periods. Likewise, $d_{0nt}^{(s)} = 1$ if the person did not participate between $t-s$ and $t-1$. Thus, $d_{knt}^{(s)}$ is an index variable that allows us to condition on the behaviour of individuals with the labour market histories defined by $z_{knt}^{(s)}$. The conditional choice probabilities $p_{knt}^{(s,N)}$ are computed as

$$p_{knt}^{(s,N)} = \frac{\sum_{m=1,m\neq n}^N \sum_{r=1}^\tau d_{mr} d_{kmr}^{(s)} J[\delta_N^{-1}(\psi_{knt}^{(s,N)} - \psi_{mr}^N)]}{\sum_{m=1,m\neq n}^N \sum_{r=1}^\tau d_{kmr}^{(s)} J[\delta_N^{-1}(\psi_{knt}^{(s,N)} - \psi_{mr}^N)]}, \tag{5.9}$$

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

To evaluate the term $\partial p_{1nt}^{(s)} / \partial l_{nt}$ which appears in the Euler equation, define

$$f_{1nt}^{(s)} \equiv f_1(\psi_{1nt}^{(s)} \mid d_{n,t+s} = 1),$$

as the probability density function for $\psi_{1nt}^{(s)}$, conditional on participating at date $t+s$. Similarly, let $f_{nt}^{(s)} \equiv f(\psi_{1nt}^{s})$ be the related probability density that does not condition on participating in period $t+s$ for $s = 1, \ldots, \rho$. Denote their derivatives with respect to l_{nt}^* by $f_{1nt}^{\prime(s)}$ and $f_{nt}^{\prime(s)}$. It is straightforward to show that

$$\frac{\partial p_{1nt}^{(s)}}{\partial l_{nt}} = \left[\frac{f_{1nt}^{\prime(s)}}{f_{1nt}^{(s)}} - \frac{f_{nt}^{\prime(s)}}{f_{nt}^{(s)}}\right] p_{1nt}^{(s)}, \qquad s = 1, \ldots, \rho.^9 \tag{5.10}$$

The nonparametric estimates of $f_{1nt}^{(s)}$ and $f_{nt}^{(s)}$ are defined, respectively, as the numerators and denominators of $p_{1nt}^{(s,N)}$ in equation (5.9). The estimates of $f_{1nt}^{\prime(s)}$ and $f_{nt}^{\prime(s)}$ are obtained from the derivatives of the estimates $f_{1nt}^{(s,N)}$ and $f_{nt}^{(s,N)}$ with respect to l_{nt}. (See Prakasa Rao (1983).

The nonparametric estimates of the conditional probabilities are now used together with the estimates of the individual-specific effects to provide some diagnostics about our model. The sample counterparts of $p_{knt}^{(s)}$ are calculated by conditioning on the behaviour of individuals who have the corresponding labour market histories. In these calculations, $\rho = 3$ and the elements of x_{nt} consist of age AGE_{nt}, age squared AGE_{nt}^2, age times education $\text{AGE}_{nt} \times \text{EDU}_{nt}$, marital status MAR_{nt}, a dummy for white vs. non whites, two region dummies NE_{nt} and SO_{nt}, and variables showing the number of children younger than six years old and the number of children between six and fourteen years, YKID_{nt} and OKID_{nt}, respectively.

Table IV shows the sample averages of the conditional choice probabilities, their derivatives with respect to hours worked at date t, and the number of observations used to calculate the probability assocated with each possible history $z_{knt}^{(s)}$ for $k = 0, 1$ and

9. Notice that $p_{1nt}^{(s)}$ can be written as $p_{1nt}^{(s)} = \Pr(d_{n,t+s} = 1 \mid \psi_{1nt}^{(s)}) = \Pr(d_{n,t+s} = 1) f_{1nt}^{(s)} / f_{nt}^{(s)}$. Differentiating this expression with respect to l_{nt} yields the expression in the text.

TABLE IV

Sample averages of nonparametric estimates

Variable	Comparison group	Sample mean†	Sample std. dev.†	Sample mean‡	Sample std. dev.‡
p_{nt}	15129	0·74	0·20	0·74	0·20
$p_{nt}^{(1,1)}$	1405	0·92	0·06	0·92	0·06
$p_{nt}^{(0,1)}$	4287	0·29	0·08	0·29	0·09
$p_{nt}^{(1,2)}$	958	0·38	0·11	0·38	0·12
$p_{nt}^{(0,2)}$	3329	0·22	0·07	0·23	0·08
$p_{nt}^{(1,3)}$	733	0·26	0·10	0·26	0·12
$p_{nt}^{(0,3)}$	2596	0·21	0·07	0·21	0·07
$\partial p_{nt}^{(1,1)}/(\partial l_{n,t})$	—	0·00002	0·00001	0·00002	0·00001
$\partial p_{nt}^{(1,2)}/(\partial l_{n,t})$	—	0·000007	0·00002	0·000007	0·00002
$\partial p_{nt}^{(1,3)}/(\partial l_{n,t})$	—	−0·000008	0·00002	−0·000009	0·00002

† Nonparametrically estimated individual effects.
‡ Time-averaged individual effects.

$s=1,\ldots,\rho$. The sample averages of the choice probabilities decline monotonically with s for each $k=0$, 1 and are fairly precisely estimated. For example, conditional on having worked at t, the probability of working at $t+1$ as captured by $p_{1nt}^{(1)}$ is 0·92 whereas if the person had not worked at date t, the probability of working at $t+1$ falls to $p_{0nt}^{(1)}=0{\cdot}29$. This table also shows the change in the probability of working at date $t+s$ due to additional hours worked at date t, conditional on the different labour market histories. The probability of working at $t+1$ increases with time spent at work at t. Likewise, if the woman did not participate at $t+1$, the probability of participation at $t+2$ increases with extra work at t. Thus, the behaviour of the estimated conditional choice probabilities are consistent with a model of human capital accumulation in which past labour market participation increases the probability of future participation.

Table V provides a second set of diagnostics in terms of the response of an individual with the average set of characteristics to changes in her wealth, as captured by changes in her inverse social weight, and skill levels. This table shows the effects of changing η_n and ν_n by a standard deviation in either direction on such variables as the probability of participation, hours worked, wages, and income. The columns in each of the five panels

TABLE V

Behavioural responses†

	$\eta_n=\mu_\eta+\sigma_\eta;\ \nu_n=\mu_\nu$						$\nu_n=\mu_\nu+\sigma_\nu;\ \eta_n=\mu_\eta$					
Variable	$z_1^{(1)}$	$z_1^{(2)}$	$z_1^{(3)}$	$z_0^{(1)}$	$z_0^{(2)}$	$z_0^{(3)}$	$z_1^{(1)}$	$z_1^{(2)}$	$z_1^{(3)}$	$z_0^{(1)}$	$z_0^{(2)}$	$z_0^{(3)}$
p	0·96	0·43	0·24	0·31	0·24	0·23	0·96	0·41	0·23	0·31	0·24	0·24
l	1513	1460	1394	1490	1394	1320	1504	1448	1382	1447	1382	1306
w	7·18	5·71	5·48	5·94	5·38	5·46	8·80	7·00	6·72	7·28	6·60	6·70
λwl	10877	7803	6311	8666	7028	5950	13250	9493	7670	10543	8542	7224
	$\eta_n=\mu_\eta+\sigma_\eta;\ \nu_n=\mu_\nu$						$\nu_n=\mu_\nu+\sigma_\nu;\ \eta_n=\mu_\eta$					
Variable	$z_1^{(1)}$	$z_1^{(2)}$	$z_1^{(3)}$	$z_0^{(1)}$	$z_0^{(2)}$	$z_0^{(3)}$	$z_1^{(1)}$	$z_1^{(2)}$	$z_1^{(3)}$	$z_0^{(1)}$	$z_0^{(2)}$	$z_0^{(3)}$
p	0·96	0·39	0·22	0·29	0·23	0·23	0·96	0·39	0·23	0·30	0·23	0·23
l	1484	1425	1357	1425	1357	1282	1486	1427	1360	1427	1360	1285
w	7·18	5·71	5·48	5·94	5·38	5·46	5·56	4·42	4·24	4·60	4·17	4·23
λwl	10668	7622	6145	8465	6843	5782	8370	5910	4766	6564	5307	4484

† Nonparametrically estimated individual effects.

refer to alternative participation histories; recall $z_k^{(s)}$ means that the person enters period $t+s$ having set d_t equal to k and then stayed out of the labour force for $s-k$ periods. Apart from the participation histories and the values of the estimated incidental parameters, all other individual-specific characteristics are set equal to their mean values in the sample.[10]

The top rows of each of the panels shows that the patterns observed in the estimated conditional choice probabilities that are reported in Table IV across the whole sample are also present for this representative individual. As before, a longer absence from the labour market reduces the probability of future participation; if the woman worked in the previous year, the probability of working in the current one is over 0·95, but this probability tails off to 0·23 after three years of nonparticipation.[11] If the representative individual chooses to work, the time spent at work also increases with past participation; as the second row of each panel shows, hours worked increases by 200 hours per year with three years experience. Similarly the wages of an average individual increase up to 31% with three years experience. Turning to the effects of changing individual's wealth and skill levels, Table V shows that the direction of these effects is consistent with the predictions of economic theory: an increase in η_n, which corresponds to an increase in the marginal utility of wealth, leads to greater participation and labour supply. The substitution effect of a higher wage rate leads to the same qualitative response. However, the magnitudes of these effects are small: changing η_n and ν_n by a standard deviation either way affects the probability of participation by less than 0·03 and annual hours of participants by less than 25.

We can also gauge the importance of aggregate shocks from the responses reported in Table V. The wage function described by equation (4.6) shows that a woman in our model would respond the same way to a permanent increase in her person-specific wage component ν_n as to the same sustained rise in the value of a standard efficiency unit ω_t. This is because ν_n enters into w_{nt} with ω_t multiplicatively, and future changes in ω_t do not depend on its current level. Similarly, unanticipated changes in η_n have the same effect as changes in λ_t, the shadow price of consumption. Consequently, investigating the effects of aggregate wage and price changes reduces to comparing people with different permanent labour market skills and different wealth endowments. In our sample, the sequence of estimated time dummies $\lambda_t\omega_t/(\lambda_{t-1}\omega_{t-1})$ has a standard deviation that is roughly one tenth of the estimated standard deviations for η_n and ν_n respectively. Thus the effects of aggregate shocks are of a lower order of magnitude than the changes exhibited in the table, and indeed do not register at the level of accuracy we are reporting.

6. PARTICIPATION AND HOURS

The above discussion shows that hours worked in previous periods raises the current wage rate, which leads to higher participation rates and labour supply by the work force participant. We now ask whether preferences over leisure are additively separable over time, and if not, whether they mitigate or exacerbate the persistence in hours and participation decisions arising from the effects of past labour market participation. To answer this

10. In what follows, we report results for the case in which the incidental parameters are estimated using the nonparametric estimator described in Section 4.4 but the results are not affected if a traditional fixed-effect estimator is used instead.

11. There are some slight differences in the values of the estimated conditional choice probabilities between Tables IV and V. Table IV shows the *average* response of all individuals in the sample whereas Table V shows the response of a representative individual who has the *average* characteristics of these individuals.

question, we parameterize the distribution of the unobservable preference shocks and the remaining components of the utility function.

6.1. *A parameterization*

It is now assumed that ε_{0nt} and ε_{1nt} are identically and independently distributed over (n, t) as Type 1 extreme value random variables. This distributional assumption produces a very simple form for the differences in the conditional valuation functions, and the expected values of the unobservables when their choice is picked. Specifically

$$q(p) = \ln [p/(1-p)], \tag{6.1}$$

$$\varphi_0(p) = \zeta - \ln [(1-p)], \tag{6.2}$$

$$\varphi_1(p) = \zeta - \ln [p], \tag{6.3}$$

where ζ is Euler's constant, approximately equal to 0·576. In this case, $q(p) = -[\varphi_1(p) - \varphi_0(p)]$, which implies that the conditional valuation functions do not depend on terms that show which choices will be optimal in the future. As a result, the dynamic selection terms (5.7) drop out under this parameterization of the unobservables.

The remaining components of the utility function U_0 and U_1 are parameterized as

$$U_0(z_{nt}) = x'_{nt}B_0, \tag{6.4}$$

$$U_1(z_{nt}, l_{nt}) = x'_{nt}B_1 l_{nt} + \textstyle\sum_{s=0}^{\rho} \delta_s l_{nt} l_{n,t-s}. \tag{6.5}$$

In these expressions, B_0 are parameters that characterize the fixed costs of participation, and B_1 shows the effect of exogenous time-varying characteristics on the marginal disutility of work. Preferences are concave decreasing in time spent at work if $x'_{nt}B_1 + 2\delta_0 l_{nt} + \sum_{s=1}^{\rho} \delta_s l_{n,t-s} < 0$ and $\delta_0 < 0$. The parameters δ_s for $s = 1, \ldots, \rho$ capture intertemporal non-separabilities in preferences with respect to labour supply choices. A value of $\delta_s < 0$ for $s = 1, \ldots, \rho$ means that hours worked s periods ago increases the marginal disutility of work, and results in less work today. Equivalently, a finding of $\delta_s < 0$ implies that current and past leisure time are substitutes whereas $\delta_s > 0$ implies that current and past leisure time are complements.

Recall that the number of hours a woman works is determined by comparing the marginal disutility of work, which depends on the current and expected future utility costs of working, to the marginal benefit from work, captured by current and expected future wage payments. Since the aggregate price process $\{\lambda_t \omega_t\}$ can take on arbitrarily large positive values, the quadratic specification of preferences implies that there may exist some states of the economy in which a nonnegligible fraction of the population will specialize in work (with $l_{nt} = 1$). While this is not an empirically relevant issue from the point of the current analysis, it suggests a possible shortcoming of the quadratic specification in describing the number of hours an individual chooses to spend at work. This parameterization also shows that all the terms in equations (5.6) and (5.7) that involve choosing not to participate in future periods are zero

$$\partial U_1(z_{1nt}^{(s)}, 0)/\partial l_{nt} = 0,$$

$$U_1(z_{1nt}^{(s)}, 0) - U_1(z_{0nt}^{(s)}, 0) = 0.$$

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

$$\eta_n \lambda_t w_{nt} + E_t \left\{ \sum_{s=1}^{\rho} \beta^s (1 - p_{1nt}^{(s)})^{-1} \left(\frac{f_{1nt}^{\prime(s)}}{f_{1nt}^{(s)}} - \frac{f_{nt}^{\prime(s)}}{f_{nt}^{(s)}} \right) p_{1nt}^{(s)} \right\} + x_{nt}' B_1 + 2\delta_0 l_{nt} + \sum_{s=1}^{\rho} \delta_s l_{n,t-s} = 0. \quad (6.6)$$

Likewise, the participation condition can be written as

$$\ln\left(\frac{p_{nt}}{1-p_{nt}}\right) - \eta_n \lambda_t w_{nt} l_{nt} - E_t \left\{ \sum_{s=1}^{\rho} \beta^s \ln\left(\frac{1-p_{1nt}^{(s)}}{1-p_{0nt}^{(s)}}\right) \right\}$$
$$- x_{nt}' B_0 - x_{nt}' B_1 l_{nt} - \sum_{s=0}^{\rho} \delta_s l_{nt} l_{n,t-s} = 0. \quad (6.7)$$

Our estimation procedure uses sample counterparts of the Euler and participation equations in (6.6) and (6.7) to derive estimates of the remaining unknown parameters of the model, namely, B_0, B_1, and δ_s for $s = 0, \ldots, \rho$.

6.2. *Estimation*

The above arguments show that there are two sources of error in evaluating the sample counterparts of (6.6) and (6.7). The first is the forecast error that arises from replacing expectations of future variables with their realizations. The second is the approximation error that arises due to replacing the true values of the conditional choice probabilities and the time-invariant individual-specific effects with their estimates. Let h_{1nt}^N and h_{2nt}^N denote the disturbances in (6.6) and (6.7) that correspond to such sources of error as

$$h_{1nt}^N \equiv \eta_n^N \lambda_t^N w_{nt} + \sum_{s=1}^{\rho} \beta^s (1 - p_{1nt}^{(s,N)})^{-1} \left(\frac{f_{1nt}^{\prime(s,N)}}{f_{1nt}^{(s,N)}} - \frac{f_{nt}^{\prime(s,N)}}{f_{nt}^{(s,N)}} \right) p_{1nt}^{(s,N)}$$
$$+ x_{nt}' B_1 + 2\delta_0 l_{nt} + \sum_{s=1}^{\rho} \delta_s l_{n,t-s}, \quad (6.8)$$

$$h_{2nt}^N \equiv \ln\left(\frac{p_{nt}^N}{1-p_{nt}^N}\right) - \eta_n^N \lambda_t^N w_{nt} l_{nt} - \sum_{s=1}^{\rho} \beta^s \ln\left(\frac{1-p_{1nt}^{(s,N)}}{1-p_{0nt}^{(s,N)}}\right)$$
$$- x_{nt}' B_0 - x_{nt}' B_1 l_{nt} - \sum_{s=0}^{\rho} \delta_s l_{nt} l_{n,t-s}, \quad (6.9)$$

where λ_t^N is derived from the sequence of estimated time dummies in the consumption equation.

Absent aggregate shocks, the only uncertainty about labour supply choices in some future period is whether the individual will participate or not in that period. Conditional on participating, the number of hours worked at date $t+s$ depends on variables known at that date. In equations (6.8) and (6.9), the uncertainty about future participation decisions is captured by the participation probabilities and their derivatives. If variables such as $p_{knt}^{(s)}$ and $f_{1nt}^{(s)}$ were known by the econometrician, then the disturbance terms in (6.8) and (6.9) would be identically zero. Since these variables are unknown and must be estimated, the disturbances (h_{1nt}^N, h_{2nt}^N) arise from the estimation error in the conditional choice probabilities and individual-specific effects.

With aggregate shocks, however, the error terms (h_{1nt}^N, h_{2nt}^N) would be correlated because they also include a common component of the forecast error in predicting the future aggregate state $\lambda_{t+s}\omega_{t+s}$. To overcome this problem, we make use of the decomposition of individuals' shadow value of consumption and individual wages together with

Assumption 1 on the exogenous price processes described in Section 2. This step of our estimation requires a distributional assumption for the innovations $\{\pi_t\}_{t=0}^{\infty}$ to the growth rates of the aggregate prices defined in Assumption 1. We now assume this to be lognormally distributed with mean 0 and variance σ^2. Armed with this distributional assumption, we can generate a sequence of independent future aggregate prices for each individual and calculate the behavioural response of each individual to such simulated shocks. Since the future aggregate state has been simulated to be independent across individuals, the person-specific disturbances h_{1nt}^N and h_{2nt}^N defined by equations (6.8) and (6.9) will also be independent across individuals, up to the sampling error induced by the estimation of the conditional choice probabilities and individual effects.

The simulated shock sequence is generated by drawing a sequence of independent $N(0, 1)$ shocks $\pi_{nt}^{(s)}$ for each observation and defining a hypothetical future aggregate price series for each individual as

$$\lambda_{nt}^{(s,N)}\omega_{nt}^{(s,N)}=\lambda_t^N\omega_t^N \exp(\pi_{nt}^{(s)}\sigma), \tag{6.10}$$

for $s=1, \dots, \rho$. In (6.10), λ_t^N and ω_t^N are derived from the estimated time dummies in the consumption and wage equations, respectively, and σ is an unknown parameter to be estimated. The behavioural response of each individual to the simulated sequence of prices is calculated using the conditional choice probabilities $p_{knt}^{(s,N)}$, which are evaluated as a function of the simulated state vector

$$\psi_{knt}^{(s,N)}\equiv(z_{knt}^{(s)\prime}, v_n^N\eta_n^N\lambda_{nt}^{(s,N)}\omega_{nt}^{(s,N)})',$$

for $k=\{0, 1\}$. Since we have calculated the simulated response of individuals to independent draws of the aggregate shocks when computing the conditional choice probabilities $p_{knt}^{(s,N)}$ for $k=0, 1$ and $s=1, \dots, \rho$, it is easy to see that h_{1nt}^N and h_{2nt}^N are uncorrelated across individuals. They are also uncorrelated with any variable known at time t, that is, $E_t[h_{int}]=0$ for $i=1, 2$.

The remaining unknown parameters of the model consist of the parameters in the hours and participation equations, the discount factor β, and the standard deviation of the simulated shock process σ. Define the $(2K+\rho+1)$-dimensional vector $\Theta_3\equiv(B_0', B_1', \delta_0, \dots, \delta_\rho)'$. Then the unknown parameters of the hours and participation equations consist of $(\Theta_3', \beta, \sigma)'$. Only β and σ enter these equations nonlinearly. This fact (together with some difficulties in estimating β as a free parameter) motivates an estimator of Θ_3, conditional on different values for β and σ. For this purpose, define the $(2K+\rho+1)$-dimensional vectors Ψ_{1nt} and Ψ_{2nt} by

$$\Psi_{1nt}\equiv(0, -x_{nt}', -2l_{nt}, -l_{nt-1}, -l_{nt-2}, \dots, -l_{nt-\rho})'$$

and

$$\Psi_{2nt}\equiv(x_{nt}', x_{nt}'l_{nt}, l_{nt}^2, l_{nt}l_{nt-1}, l_{nt}l_{nt-2}, \dots, l_{nt}l_{nt-\rho})'.$$

Also define the 2×1 vector $y_{nt}^N=(y_{1nt}^N, y_{2nt}^N)'$ where

$$y_{1nt}^N\equiv\eta_n^N\lambda_t^N w_{nt}+\textstyle\sum_{s=1}^{\rho}\beta^s(1-p_{1nt}^{(s,N)})^{-1}\left(\dfrac{f_{1nt}^{\prime(s,N)}}{f_{1nt}^{(s,N)}}-\dfrac{f_{nt}^{\prime(s,N)}}{f_{nt}^{(s,N)}}\right)p_{1nt}^{(s,N)},$$

$$y_{2nt}^N\equiv\ln\left(\frac{p_{nt}^N}{1-p_{nt}^N}\right)-\eta_n^N\lambda_t^N w_{nt}l_{nt}-\textstyle\sum_{s=1}^{\rho}\beta^s\ln\left(\dfrac{1-p_{1nt}^{(s,N)}}{1-p_{0nt}^{(s,N)}}\right).$$

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

$$h^N_{int} = y^N_{int} - \Psi'_{int}\Theta_3, \qquad i = 1, 2. \tag{6.11}$$

Let $h^N_{nt} \equiv (h^N_{1nt}, h^N_{2nt})'$ and let T_3 denote the set of periods for which the hours and participation conditions are valid. Define the vector $h^{N'}_n \equiv (h^{N'}_{n1}, \ldots, h^{N'}_{n,T_3})'$ as the vector of idiosyncratic errors for a given individual over time and $y^N_n \equiv (y^{N'}_{n1}, \ldots, y^{N'}_{nT_3})'$. Similarly, define $\Psi_{nt} \equiv (\Psi'_{1nt}, \Psi'_{2nt})'$, $\Psi_n \equiv (\Psi'_{n1}, \ldots, \Psi'_{nT_3})'$ and $\Phi^N_n \equiv E_t[h^N_n h^{N'}_n]$. Notice that the matrix Φ^N_n is block diagonal with diagonal elements defined as $\Phi^N_{nt} \equiv E_t[h^N_{nt} h^{N'}_{nt}]$, and off-diagonal elements that are zero because $E_t[h^N_{nt} h^{N'}_{ns}] = 0$ for $s \neq t$, $s < t$. The 2×2 conditional heteroscedasticity matrix Φ^N_{nt} associated with the individual-specific errors h^N_{nt} is evaluated using a nonparametric estimator based on the estimated residuals h^N_{nt} derived from an initial consistent estimate of Θ_3. The optimal instrumental variables estimator for Θ_3 satisfies

$$\Theta_3^{(N)} = [1/N \textstyle\sum_{n=1}^N \Psi'_n(\Phi^N_n)^{-1}\Psi_n]^{-1}\,[1/N \textstyle\sum_{n=1}^N \Psi'_n(\Phi^N_n)^{-1} y^N_n]. \tag{6.12}$$

Appendix A.3 derives the asymptotic properties of a general class of estimators which can be used to show the consistency and asymptotic normality of the estimator $\Theta_3^{(N)}$.

6.3. *Empirical findings*

Table VI contains estimates of alternative specifications of the participation and hours equations. Column (1) reports estimates of the preference parameters that are based on nonparametric estimates of the individual effects η_n and $\eta_n v_n$, while the estimates in the remaining columns are based on standard fixed-effects estimators of η_n and $\eta_n v_n$, as described in Section 4.4. The subcolumns marked by (a) refer to single-equation estimates of the participation equation, those marked by (b) refer to single-equation estimates of the hours equation, and those marked by (c) refer to joint estimates of the participation and hours equations. The vector of demographic characteristics that enters the hours and participation equations x_{nt} is defined to include the variables KIDS_{nt}, $\text{KIDS}_{nt} \times \text{MAR}_{nt}$, AGE_{nt}, and AGE^2_{nt}, where KIDS_{nt} equals the total number of children in the household. The estimates reported in Table VI are conditional on given values of the discount factor β and the standard deviation of the innovations to the growth rates of the aggregate price process, σ. As a baseline case, β is set equal 0·90 and σ equal to the standard deviation of the innovations to the growth rates of the estimated aggregate prices.

The results in columns (1) and (2) show that alternative methods for computing the individual-specific effects yield roughly similar results; the coefficients involving the effects of current and past labour supply on current utility all have the same signs and are approximately of the same magnitude. While it is difficult to estimate the fixed costs of participation as captured by the parameters of $U_0(z_{nt})$, the specification in column (2) based on the time-averaged fixed effects yields more precise estimates of this component of the utility function. For this reason, we report results using the time-averaged fixed effects in what follows.

An important finding that emerges from all the specifications considered in Table VI is the substitutability of leisure choices in intertemporal preferences, current labour supply increasing the future marginal disutility of work through the interaction with future labour supply since $\delta_i < 0$ for $i = 1, \ldots, 3$. In the parlance of home production economics, human capital can be accumulated by spending less time in market work and increasing household

TABLE VI

Hours and participation†

Variable	Parameter	1. $\beta=0·9$, $\sigma=0·05$ (a)	(b)	(c)	2. $\beta=0·9$, $\sigma=0·05$ (a)	(b)	(c)	3. $\beta=0·9$, $\sigma=0$ (a)	(b)	(c)	4. $\beta=0·9$, $\sigma=0$ (c)
1	B_{00}	1257	—	−676	−14	—	−154	−44	—	−277	—
		(566)		(230)	(264)		(115)	(264)		(113)	
$KIDS_{nt}$	B_{01}	31	—	116	204	—	−43	173	—	120	—
		(222)		(110)	(132)		(72)	(132)		(41)	
$KIDS_{nt} \times MAR_{nt}$	B_{02}	−184	—	5·62	−187	—	42	−154	—	−77	—
		(208)		(106)	(129)		(71)	(129)		(25)	
AGE_{nt}	B_{03}	−37	—	42.2	31	—	13	29	—	18	—
		(33)		(16)	(13)		(6·6)	(13)		(6·5)	
AGE^2_{nt}	B_{04}	0·06	—	0·03	−0·03	—	−0·02	−0·03	—	−0·03	—
		(0·03)		(0·02)	(0·01)		(0·01)	(0·01)		(0·01)	
l_{nt}	B_{10}	−11·6	−10·4	−11	−5·4	−5·6	−5·5	−4·8	−5·1	−4·5	−2·7
		(0·51)	(0·04)	(0·25)	(0·24)	(0·02)	(0·12)	(0·24)	(0·02)	(0·11)	(0·12)
$KIDS_{nt} \times l_{nt}$	B_{11}	0·16	0·28	0·22	−0·55	−0·14	−0·40	−0·49	−0·14	0·3E−05	−0·51
		(0·17)	(0·02)	(0·10)	(0·10)	(0·01)	(0·07)	(0·10)	(0·01)	(0·1E−04)	(0·15)
$(KIDS_{nt} \times MAR_{nt})l_{nt}$	B_{12}	0·18	−0·12	0·07	0·60	0·34	0·52	0·54	0·34	−0·05	0·90
		(0·16)	(0·02)	(0·10)	(0·10)	(0·01)	(0·07)	(0·10)	(0·01)	(0·02)	(0·15)
$AGE_{nt} \times l_{nt}$	B_{13}	0·31	0·03	−0·60	−0·52	0·008	−0·20	−0·47	0·003	−0·24	0·19
		(0·49)	(0·003)	(0·26)	(0·17)	(0·001)	(0·01)	(0·17)	(0·001)	(0·01)	(0·005)
$AGE^2_{nt} \times l_{nt}$	B_{14}	0·002	0·002	0·002	0·002	0·002	0·002	0·002	0·001	0·002	0·002
		(0·05E−03)	(0·5E−04)	(0·3E−03)	(0·2E−03)	(0·2E−04)	(0·1E−03)	(0·2E−03)	(0·2E−04)	(0·1E−03)	(0·2E−03)
l^2_{nt}	δ_0	0·002	0·002	0·002	0·001	0·001	0·001	0·001	0·001	0·001	0·3E−04
		(0·1E−03)	(0·2E−04)	(0·6E−04)	(0·6E−04)	(0·1E−05)	(0·3E−04)	(0·6E−04)	(0·8E−05)	(0·3E−04)	(0·1E−03)
$l_{nt}l_{n,t-1}$	δ_1	−0·001	−0·001	−0·001	−0·6E−03	−0·5E−03	−0·6E−03	−0·5E−03	−0·5E−03	−0·6E−03	−0·7E−03
		(0·1E−03)	(0·2E−04)	(0·1E−04)	(0·5E−04)	(0·1E−05)	(0·4E−04)	(0·5E−04)	(0·1E−05)	(0·4E−04)	(0·1E−03)
$l_{nt}l_{n,t-2}$	δ_2	−0·7E−03	−0·7E−03	−0·7E−03	−0·4E−03	−0·4E−03	−0·3E−03	−0·3E−03	−0·4E−03	−0·3E−03	−0·6E−03
		(0·1E−03)	(0·2E−04)	(0·1E−04)	(0·6E−04)	(0·1E−05)	(0·4E−04)	(0·6E−04)	(0·9E−05)	(0·5E−04)	(0·1E−03)
$l_{nt}l_{n,t-3}$	δ_3	−0·7E−03	−0·8E−03	−0·6E−03	−0·3E−03	−0·3E−03	−0·3E−03	−0·2E−03	−0·3E−03	−0·2E−03	−0·5E−03
		(0·1E−03)	(0·2E−04)	(0·1E−04)	(0·5E−04)	(0·1E−05)	(0·4E−04)	(0·5E−05)	(0·9E−05)	(0·4E−04)	(0·1E−03)

† Standard errors in parentheses. a. Participation equation; b. Hours equation; c. Joint estimates.
Column 1: Nonparametric fixed effects. Columns 2, 3 and 4: Time-averaged fixed effects.

production to offset future household demands on nonmarket time. The only counterintuitive finding is that the coefficient on the quadratic term for current hours worked δ_0 is positive and significant. This result violates the concavity assumption of preferences with respect to the number of hours worked, and casts doubt on how well the model describes the determination of hours by women in the sample. Further research on this issue may shed light on the reasons. One possibility is that the quadratic parameterization is a misspecification of preferences. Another possibility is that individuals are constrained about the number of hours that they work, where the maximum number of hours varies across the population.

Our framework allows for the effects of past labour supply on both current wages and current utilities. Comparing our results with those of other papers that have similar features, our results on the effects of lagged labour supply on current wages are broadly similar to those of Eckstein and Wolpin (1989a) and Miller and Sanders (1997). Eckstein and Wolpin define experience as the number of years a woman has participated in the past, and consider an earnings function in which past experience, past experience squared, and schooling enter log-linearly. By contrast, the approach in Miller and Sanders and this study is to model experience in terms of hours worked at different periods in the past.[12] Both our study and Miller and Sanders (1997) also allow for time-invariant, person-specific productivity effects in wages as well as for the effects of such demographic variables as age and age times education (which Eckstein and Wolpin limit to schooling). All three studies find that past experience in the labour market increases wages and thus, has a positive effect on current labour supply.

Turning to the effects of past labour supply on current utility, we can compare our results with those obtained by Hotz, Kydland and Sedlacek (1988), Eichenbaum, Hansen and Singleton (1988), Eckstein and Wolpin (1989a) and Miller and Sanders (1997). Broadly speaking, studies that use annual data on hours worked typically find that current and past leisure choices are substitutes. These include Hotz, Kydland and Sedlacek (1988), Eckstein and Wolpin (1989a) and our study. In contrast to Eckstein and Wolpin's study, however, we find that the estimated coefficients on lagged labour supply $\delta_i<0$ for $i=1,\ldots,3$ are significant. Studies that use data on hours worked sampled over shorter intervals find that hours worked in the past reduces the marginal disutility of current work, or that women display habit persistence in labour supply choices. These include Eichenbaum, Hansen and Singleton (1988), who use quarterly aggregate data on consumption and labour supply, and Miller and Sanders (1997), who use monthly data on individual welfare and labour market participation decisions. Thus, it appears that the interval over which hours worked is measured affects the type of implied nonseparabilities in preferences, data sampled over shorter intervals leading to a finding of the complementarity of current and past leisure choices, and data sampled over longer intervals leading to a finding of substitutability.

Table VI also reports estimates of the determinants of the fixed costs of labour market participation, and the effects of demographic characteristics such as age, the number of children, and marital status on the participation and hours decisions. Unlike the parameters that determine the effects of past labour supply, however, many of the coefficients are not precisely estimated, and there are some sign reversals depending on whether the Euler equation is estimated separately or jointly with the participation equation. Nevertheless, some broad conclusions can still be drawn. For example, the fixed costs of participation decrease with age (but at a decreasing rate), implying that if wages were constant,

12. Miller and Sanders allow for work experience accumulated in five past periods whereas we allow for hours worked and labour market participation up to three years in the past to affect current wages.

an older woman would be more likely to participate in the labour market. These effects are roughly similar across the specifications in columns 1.c and 2.c as well as the specification in 3.c, which assumes that there are no aggregate shocks. For example, the estimates in column 2.c imply that the fixed costs of participation decline with age as $-154+13\text{AGE}_{nt}-0{\cdot}02\text{AGE}_{nt}^2$. However, conditional on participating, the estimates of B_{13} and B_{14} that use the participation equation (either alone or jointly with the Euler equation) imply an older woman would work a smaller number of hours. Using the estimates in column 2.c, the marginal disutility from work increases with age as $-5{\cdot}5-0{\cdot}2\text{AGE}_{nt}+0{\cdot}002\text{AGE}_{nt}^2$.

The effects of children on the participation decision are also difficult to pin down precisely. Among the different specifications reported in Table VI, the estimates in column 2.c (which allows for aggregate shocks and considers the Euler and participation equations jointly) imply that the effect of an additional child is to reduce participation for both unmarried and married women since $B_{01}<0$ and $B_{01}+B_{02}<0$. Since our framework allows for the separate effects of children in the household on the probability of participation and the marginal choice of hours, a distinction that has been typically ignored by the literature, we can also analyse the effects of additional children on the number of hours worked.[13] Based on the estimates in column (2) (as well as columns 3.a and 3.b), an increase in the number of children has different effects on the hours worked by unmarried vs. married women. Since the estimate of B_{11} is negative and significant, additional children reduce the time spent at work for unmarried women; for married women, the effect of an additional child is to increase hours worked.

Some additional specifications were estimated to determine the effects of varying the values of the subjective discount factor on the remaining parameter estimates. For alternative values of β in the grid $\{0{\cdot}6, 0{\cdot}65, \ldots, 1\}$, the estimates were not affected up to the fourth decimal place. Even given large amounts of data, the numerical routines have difficulty differentiating ways in which females would value their future labour supply profiles due to small changes in the discount factor. That the subjective discount factor is, for all practical purposes, unidentified is a feature that our model shares with many other estimated structural models of dynamic choice.[14] These results suggest that we should consider estimating models in which agents maximize the expected value of utility per unit time, rather than attempting to recover a parameter which purports to show how quickly people discount future events.

Our framework allows us to derive some measures of the importance of aggregate shocks. In the specification reported in column (2), σ is set equal to the estimated standard deviation of the growth rates of the aggregate price process while in column (3), estimates for the no aggregate shock case with $\sigma=0$ are reported. Comparing the values of the estimated coefficients in the columns marked (2) vs. (3) shows that most are within a standard deviation of each other, a notable exception being the coefficient on the number of current hours, B_{10}. Overall a regression towards the mean is observed, many coefficients shrinking in absolute value when the common shocks are ignored. This provides some vindication for incorporating aggregate shocks in our framework. On the other hand,

13. For example, in their dynamic models of female participation and labour supply, Heckman and MaCurdy (1980) and Hotz and Miller (1988) investigate empirically the effects of children on an index function that determines participation, and for those who participate, the number of hours worked. However, their specifications constrain the data from differentiating between the effects of the participation decision vs. the marginal hours choice. See Browning (1992) for a review of the effects of children on female labour supply.

14. Many dynamic programming models simply set the discount factor to an arbitrary number such as 0·9 or 0·99. Some studies acknowledge the computational difficulties that they encounter. See the surveys referred to earlier for further discussion.

magnifying the standard deviation of the aggregate shocks by a factor of five had practically no effect on the coefficient estimates, and consequently this specification is not reported.

Finally, the effects of the estimated selection bias from ignoring the unobserved factors that help determine participation in labour market activities are considered. The estimates in column (4) are obtained by using standard Euler equation techniques, thereby assuming the participation decision is exogenous. The results are quite startling. The effect of hours worked in the current period on utility is half its former size while the effect of age is positive and significant. The difference between the effect of an additional child on hours worked is three times the value in columns (2) through (3), as are the effects of labour market time two and three years past on current utility. These findings suggest that while the estimated coefficients are not sensitive to variation in the discount factor and the standard deviation for aggregate shocks, the estimated selection bias from ignoring the unobserved factors which help determine participation in labour market activities is quite substantial. These findings thus provide evidence against specifications that treat participation as exogenous.

7. CONCLUSION

Our empirical findings show that previous labour market experience is a determinant of both wages and utility. The estimated wage equation includes person-specific fixed effects as well as time dummies to capture changing wage conditions and cohort effects, requiring us to identify the effects of labour experience on wages from first differences. The effects of aging in preferences as well as wealth effects, captured by its marginal utility, are incorporated into the equations determining participation and labour supply. Our main conclusion is, therefore, that over and beyond the person-specific effects and demographic controls we have incorporated into our empirical study, experience on the job and non-separable preferences over time both contribute towards observed variation in female labour supply. In addition, demographic factors, such as age and children, appear to operate on both the extensive and intensive margins of labour supply.

Our earlier work on males, Altuğ and Miller (1990) explained how to rationalize the life-cycle approach to labour supply estimated in Heckman and MaCurdy (1980), MaCurdy (1981), and others within a dynamic structural model of the economy, by assuming the existence of a complete set of competitive markets. In a convex economy, this assumption yields the Pareto optimality of observed allocations. Because of the nonconvexity from the fixed costs of labour force participation and increasing returns in past labour supply, Pareto optimality cannot be achieved by fully decentralized competitive markets within our current framework. Intuitively, the reason is that an agent deciding whether to work or not would value her extra income by her strictly concave indirect utility function for consumption (induced by convex preferences for consumption), whereas the social planner for a large population averages the extra income over the whole population according to their assigned social weights; therefore, the additional income enters the social planning problem linearly when the agents are distributed over a continuum. If labour supply was indivisible, a similar problem would arise. In our setup, Pareto optimality could be achieved by a centrally-administered mechanism that supports the optimal labour supply allocations or by a system of lotteries such as that described in Hansen (1985), Rogerson (1988), or Card (1990). As a tool useful for the empirical analysis, we have assumed that the existing contractual and institutional arrangements in labour markets decentralize decision-making to bring about Pareto optimality. Whether existing arrangements perform this task is controversial, and a topic that we leave for future research.

APPENDIX

A.

Part A of this appendix has three parts. First, it shows that there are primitives to the social planner's problem that are consistent with Assumption 1. Second, it provides a proof of the alternative representation of the conditional valuation functions used in Section 5. Third, it defines a class of conditional choice probability (CCP) estimators, to which the estimator used in Section 6.3 belongs, and shows the consistency and asymptotic normality of these estimators.

A.1. Pareto optimality. We first show that Assumption 1 in Section 2.2 is consistent with the primitives of the social planner's problem described by equations (2.4) and (2.5). Our approach is adapted from Novales (1990), who specifies an exogenous consumption process that is consistent with his assumption about the shadow price of consumption in an equilibrium model of capital accumulation. In a similar manner, we seek conditions on endowments to justify an exogenously specified process for the shadow price of a unit of standardized human capital defined by $\{\lambda_t\omega_t\}_{t=0}^{\infty}$.

Recall that by Assumption 1, the random variable $\tau_t \equiv (\lambda_t\omega_t)/(\lambda_{t-1}\omega_{t-1})$ is assumed to be identically and independently distributed with distribution function G.

Proposition 0. *Given any well-defined probability distribution function $G(\cdot)$, there exists an endowment process $\{e_t\}_{t=0}^{\infty}$ such that $\{\lambda_t\omega_t\}_{t=0}^{\infty}$ is generated as part of the solution for the social planner's problem for that economy*

Proof. Recall that the optimal allocations $c_{nt}^{0}, l_{nt}^{*}, d_{nt}^{o}$, and l_{nt}^{o} described in the text are derived under the assumption that the aggregate price process $\{\lambda_t\omega_t\}_{t=0}^{\infty}$ evolves according to Assumption 1. Substituting the Frisch demand functions for consumption, the fraction of time spent at work, and labour market participation into the aggregate feasibility condition yields

$$\begin{aligned} e_t &= \int_0^1 w_{nt} l_{nt}^{o} d\mathscr{L}(n) - \int_0^1 c_{nt}^{o} d\mathscr{L}(n) \\ &= \int_0^1 v_n\omega_t\gamma(z_{nt}) d^{o}(z_{nt}, \eta_n v_n\lambda_t\omega_t, \varepsilon_{1nt}) l^{*}(z_{nt}, \eta_n v_n \lambda_t\omega_t) d\mathscr{L}(n) \\ &\quad - \int_0^1 c^{o}(x_{nt}, \eta_n\lambda_t, \varepsilon_{2nt}) d\mathscr{L}(n). \end{aligned} \tag{A.1}$$

Thus, $\{\lambda_t\omega_t\}_{t=0}^{\infty}$ is a sequence of shadow prices that is consistent with the sequence of optimal allocations $\{c_{nt}^{o}, l_{nt}^{o}\}_{t=0}^{\infty}$ if and only aggregate feasibility condition in (A.1) is satisfied. Accordingly, we define the dividend process $\{e_t\}_{t=0}^{\infty}$ to satisfy (A.1). ||

A.2. New representation for conditional valuation functions. The new representation for conditional valuation functions described in Proposition 1 forms the basis for the estimation of the optimal labour supply and participation conditions. In this section, we provide a proof of this proposition, noting that it applies when there are more than two discrete choices. For the sake of simplicity, we provide a proof when there are only two discrete choices.

Proposition 1. *Define*:

$$u_k(\psi_{nt}) \equiv \begin{cases} U_1(z_{nt}, 0) & \text{for } k=0; \\ U_0(z_{nt}) + U_1(z_{nt}, l_{nt}^{*}) + \eta_n\lambda_t w_{nt} l_{nt}^{*} & \text{for } k=0. \end{cases}$$

Let $p_{knt}^{(s)} \equiv p(\psi_{knt}^{(s)})$ for $k \in \{0, 1\}$ and $s = 1, \ldots, \rho$. For $k \in \{0, 1\}$, the conditional valuation functions can be expressed as

$$\begin{aligned} V_k(\psi_{nt}) = u_k(\psi_{nt}) + E_t\{\textstyle\sum_{s=1}^{\rho} \beta^{s}[u_0(\psi_{knt}^{(s)}) + \varphi_0(p_{knt}^{(s)}) + p_{knt}^{(s)}(q(p_{knt}^{(s)}) + \varphi_1(p_{knt}^{(s)}) - \varphi_0(p_{knt}^{(s)}))] \\ + \beta^{\rho+1}[V_0(\psi_{knt}^{(\rho+1)}) + \varphi_0(p_{knt}^{(\rho+1)}) + p_{knt}^{(\rho+1)}(q(p_{knt}^{(\rho+1)}) + \varphi_1(p_{knt}^{(\rho+1)}) - \varphi_0(p_{knt}^{(\rho+1)}))]\}. \end{aligned} \tag{A.2}$$

Proof. To prove this proposition, we note that the objective function for our problem can be written as

$$E_0\{\sum_{t=0}^{\infty}\sum_{k=0}^{1}\beta^t d_{knt}[u_k(z_{nt})+\varepsilon_{knt}]\}, \tag{A.3}$$

where ε_{knt} is assumed identically and independently distributed across (n, t), and $d_{knt}=1$ if the k-th choice is picked and 0 otherwise. Let d^o_{knt} denote the optimal decision rule. We define the conditional valuation for action k at time t as

$$V_k(\psi_{nt})=u_k(\psi_{nt})+E\{\sum_{r=t+1}^{\infty}\sum_{k'=0}^{1}\beta^s d^o_{knr}[u_{k'}(\psi_{nr})+\varepsilon_{k'nr}]\mid \psi_{nt}, d^o_{knt}=1\}. \tag{A.4}$$

Using Bellman's principle on (A.4), it follows that

$$V_k(\psi_{nt})=u_k(\psi_{nt})+\beta E_t\{p(\psi^{(1)}_{knt})[V_1(\psi^{(1)}_{knt})+\varphi_1(p(\psi^{(1)}_{knt}))]+(1-p(\psi^{(1)}_{knt}))[V_0(z^{(1)}_{knt})+\varphi_0(p(\psi^{(1)}_{knt}))]\}. \tag{A.5}$$

We can derive an alternative expression for $V_k(\psi_{nt})$ by using the definitions for $q(p(\psi^{(1)}_{nt}))$ as

$$V_k(\psi_{nt})=u_k(\psi_{nt})+\beta E_t\{V_0(\psi^{(1)}_{knt})+\varphi_0(p(\psi^{(1)}_{knt}))+p(\psi^{(1)}_{knt})[q(p(\psi^{(1)}_{knt}))+\varphi_1(p(\psi^{(1)}_{knt}))-\varphi_0(p(\psi^{(1)}_{knt}))]\}. \tag{A.6}$$

Equation (A.6) implies that

$$V_0(\psi^{(1)}_{knt})=u_0(\psi^{(1)}_{knt})+\beta E_{t+1}\{V_0(\psi^{(2)}_{knt})+\varphi_0(p(\psi^{(2)}_{knt}))+p(\psi^{(2)}_{knt})[q(p(\psi^{(2)}_{knt}))+\varphi_1(p(\psi^{(2)}_{knt}))-\varphi_0(p(\psi^{(2)}_{knt}))]\}.$$

Substituting for $V_0(\psi^{(1)}_{knt})$ using the previous expression and performing a ρ-step induction yields equation (A.2) of the proposition. ||

A.3. A class of CCP estimators. The estimator Θ^N_3 defined by equation (6.12) is an example of a CCP estimator, in which the estimated individual-specific effects $v^N_n\eta^N_n$ and the conditional choice probabilities p^N_{nt} and $p^{(s,N)}_{nt}$ for $s=1,\ldots,\rho$ enter as incidental parameters. This estimator falls within a class of CCP estimators that can be described as follows.

Let $h_n(\Theta, v_n, p_n)$ be a $\varphi\times 1$ vector function, and Z_n a $q\times\varphi$ matrix such that Θ_0 is the unique root of $E[Z_n h_n(\Theta, v_n, p_n)]$. For each $n\in\{1, 2, \ldots\}$ and $\Theta\in\Theta$, let v^N_n be a kernel estimator which converges uniformly to v_n, let $p^N_n(\Theta, v_n)$ be a kernel estimator which converges uniformly to $p_n(\Theta, v_n)$, and let Z^N_n converge uniformly to Z_n. We define $\Theta^{(N)}$ as any solution to

$$0=1/N\sum_{n=1}^{N}Z^N_n h_n(\Theta^{(N)}, v^N_n, p^N_n(\Theta^{(N)}, v^N_n)). \tag{A.7}$$

The proof of Proposition 2 below shows that $\Theta^{(N)}$ is asymptotically normal, but is not centred on zero. While an asymptotically unbiased estimator could be calculated following the procedure in Hotz and Miller (1993) of forming a linear combination of estimators which are based on different bandwidths for the incidental parameters, the limited empirical evidence available suggests that the asymptotic bias is unimportant.[15]

Proposition 2. *$\Theta^{(N)}$ converges in probability to Θ_0 and $\sqrt{N}(\Theta^{(N)}-\Theta_0)$ is asymptotically normal with mean $-E(v_n)/2$ and covariance matrix $(D'_0)^{-1}S_0D_0$, where v_n, D_0, and S_0 are defined by equations* (A.14), (A.20) *and* (A.21).

Proof. In the text, we assume that v^N_n and $p^N_n(\Theta, v_n)$ take the form of nonparametric kernel estimators of weighted or unweighted probability density functions of the form

$$v^N_n=\sum_{m=1,m\neq n}^{N}\phi_m\delta^{-q}J[\delta_N^{-1}(x_m-x_n)], \tag{A.8}$$

and

$$p^N_n(\Theta, v_n)=\sum_{m=1,m\neq n}^{N}d_m\delta^{-q}J[\delta_N^{-1}(k(z_m, \Theta, v_m)-k(z_n, \Theta, v_n))], \tag{A.9}$$

where $k(z_m, \Theta, v_m)$ is a mapping that defines the distance between the observations. The proof that $\Theta^{(N)}$ converges in probability to Θ_0 is standard, relying on the uniform convergence of the incidental parameters to their true values, so that the approximating sample moments obtained by substituting the incidental parameter estimates for their respective true values only affect the resulting structural parameter estimator by an $o_p(1)$ term.[16]

To establish the asymptotic normality of this estimator, its asymptotic mean, covariance, and bias, we first consider another estimator denoted $\tilde{\Theta}^{(N)}$ and show that this has the same asymptotic distributional properties

15. For evidence on the magnitude of this asymptotic bias, readers are referred to the Monte Carlo simulations in Powell, Stock and Stoker (1989) and the fertility application in Hotz and Miller (1993).

16. For example, see Hotz and Miller (1993) for a consistency proof of a very similar semiparametric estimator.

as $\Theta^{(N)}$. To simplify the notation, let $h_n \equiv h_n(\Theta_0, v_n, p_n)$ and $p_n \equiv p_n(\Theta_0, v_n)$ and

$$h_{0n} \equiv \left[\frac{\partial h_n(\Theta_0, v_n, p_n)}{\partial \Theta} + \left(\frac{\partial h_n(\Theta_0, v_n, p_n)}{\partial p_n}\right)\left(\frac{\partial p_n(\Theta_0, v_n)}{\partial \Theta}\right)\right],$$

$$h_{1n} \equiv \left[\frac{\partial h_n(\Theta_0, v_n, p_n)}{\partial v_n} + \left(\frac{\partial h_n(\Theta_0, v_n, p_n)}{\partial p_n}\right)\left(\frac{\partial p_n(\Theta_0, v_n)}{\partial v_n}\right)\right],$$

$$h_{2n} \equiv \left[\frac{\partial h_n(\Theta_0, v_n, p_n)}{\partial p_n}\right].$$

The estimator $\tilde{\Theta}^{(N)}$ satisfies the equation

$$-N^{-1}\sum_{n=1}^{N} Z_n[h_n + h_{0n}(\tilde{\Theta}^{(N)} - \Theta_0)] = N^{-1}\sum_{n=1}^{N} Z_n[h_{1n}(v_n^N - v_n) + h_{2n}(p_n^N(\Theta_0, v_n) - p_n(\Theta_0, v_n))]. \quad (A.10)$$

Define the quantities

$$v_{1mn}^N \equiv Z_n h_{1n}\left[\phi_m \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - v_n\right] + Z_m h_{1m}\left[\phi_n \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - v_m\right], \quad (A.11)$$

$$v_{2mn}^N \equiv Z_n h_{2n}\left[d_m \delta_N^{-q} J\left(\frac{k(z_m, \Theta_0, v_m) - k(z_n, \Theta_0, v_n)}{\delta_N}\right) - p_n\right]$$
$$+ Z_m h_{2m}\left[d_n \delta_N^{-q} J\left(\frac{k(z_m, \Theta_0, v_m) - k(z_n, \Theta_0, v_n)}{\delta_N}\right) - p_m\right], \quad (A.12)$$

$$v_{mn}^N \equiv v_{1mn}^N + v_{2mn}^N, \quad (A.13)$$

$$v_n \equiv Z_n f(x_n)[h_{1n}(v_n + \phi_n) + h_{2n}(p_n + d_n)] - Z_n h_{1n} v_n - Z_n h_{2n} p_n, \quad (A.14)$$

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

Expanding the second expression on the right-side of (A.10) using the definition of the nonparametric estimator for v_n yields

$$N^{-1}\sum_{n=1}^{N} Z_n h_{1n}(v_n^N - v_n) = N^{-1}\sum_{n=1}^{N} Z_n h_{1n}\left[\sum_{m=1, m\neq n}^{N} \phi_m \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - v_n\right]$$
$$= N^{-1}\sum_{n=1}^{N}\sum_{m=1, m\neq n}^{N} Z_n h_{1n}\left[\phi_m \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - v_n\right]$$
$$= N^{-1}(N-1)^{-1}\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} v_{1mn}^N. \quad (A.15)$$

Similarly, the third expression in (A.10) may be written as

$$N^{-1}\sum_{n=1}^{N} Z_n h_{2n}[p_n^N(\Theta_0, v_n) - p_n(\Theta_0, v_n)] = N^{-1}(N-1)^{-1}\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} v_{2mn}^N. \quad (A.16)$$

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

$$N^{-1}(N-1)^{-1}\sum_{n=1}^{N-1}\sum_{m=n+1}^{N} v_{mn}^N = \frac{E[v_{mn}^N]}{2} - (N-1)^{-1}\sum_{n=1}^{N-1}\{E[v_{mn}^N | n] - E[v_{mn}^N]\} + o_p(1). \quad (A.17)$$

The right-side of A.17) depends on N. To derive the asymptotic distribution of $\tilde{\Theta}^{(N)}$, Lemma 1 derives the appropriate limit for the right-side of (A.17) as

$$N^{-1/2}\frac{E[v_{mn}^N]}{2} + N^{-1/2}\sum_{n=1}^{N}\{E[v_{mn}^N | n] - E[v_{mn}^N]\} = N^{-1/2}\frac{E(v_n)}{2} + N^{-1/2}\sum_{n=1}^{N}\{v_n - E(v_n)\} + o_p(1). \quad (A.18)$$

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

$$-N^{-1/2}\sum_{n=1}^{N} Z_n[h_n + h_{0n}(\tilde{\Theta}^{(N)} - \Theta_0)] = N^{-1/2}\frac{E(v_n)}{2} + N^{-1/2}\sum_{n=1}^{N}\{v_n - E(v_n)\} + o_p(1). \quad (A.19)$$

The Central Limit Theorem implies that the right-side of (A.19) converges in the distribution to a normal random variable with mean $-E(v_n)/2$. Hence, $\sqrt{N}(\tilde{\Theta}^{(N)} - \Theta_0)$ converges to a normal random variable with mean

$-E(v_n)/2$ and covariance $(D_0')^{-1}S_0D_0^{-1}$ where

$$D_0 \equiv E[Z_n h_{0n}], \tag{A.20}$$

$$S_0 \equiv E[(Z_n h_n + v_n - E(v_n))(Z_h h_n + v_n - E(v_n))']. \tag{A.21}$$

We complete the proof of this proposition with Lemma 2 provided below, which implies that $\Theta^{(N)}$ and $\tilde{\Theta}_{(N)}$ have the same asymptotic distribution, that is, $\sqrt{N}(\Theta^{(N)} - \tilde{\Theta}^{(N)})$ is $o_p(1)$. ||

Lemma 1.

$$N^{-1/2}\frac{E[v_{mn}^N]}{2} + N^{-1/2}\sum_{n=1}^N \{E[v_{mn}^N|n] - E[v_{mn}^N]\} = N^{-1/2}\frac{E(v_n)}{2} + N^{-1/2}\sum_{n=1}^N \{v_n - E(v_n)\} + o_p(1). \tag{A.22}$$

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

$$v_{1mn}^N \equiv Z_n h_{1n}\phi_m \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - Z_n h_{1n} v_n + Z_m h_{1m}\phi_n \delta_N^{-q} J\left(\frac{x_m - x_n}{\delta_N}\right) - Z_m h_{1m} v_m. \tag{A.23}$$

Taking the first-term on the right-side of (A.23),

$$\begin{aligned}
E\left[Z_n h_{1n}\delta^{-q}\phi_m J\left(\frac{x_m - x_n}{\delta}\right)\middle| x_n\right] &= Z_n h_{1n}\int \delta^{-q} v(x) J\left(\frac{x - x_n}{\delta}\right) f(x)dx \\
&= Z_n h_{1n}\int v(x_n + \delta u) f(x_n + \delta u) J(u)du \\
&= \int Z_n h_{1n}\{v(x_n)f(x_n) + v(x_n + \delta u)f(x_n + \delta u) - v(x_n)f(x_n)\}J(u)du \\
&= Z_n h_{1n} v(x_n) f(x_n) + Z_n h_{1n} t_n(\delta)
\end{aligned}$$

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

$$\begin{aligned}
E[t(\delta)^2] &= E\left\{\phi_n^2\left[\int h_1(x_n + \delta u)f(x_n + \delta u) - h_1(x_n)f(x_n)J(u)du\right]^2\right\} \\
&= E\left\{\phi_n^2\left[\int_{x_n}^{x_n + \delta u} \frac{\partial(h_1 f)(x)}{\partial x} J(u)du\right]^2\right\} \\
&\leq E\left[\phi_n^2 \int \delta^2 u^2 \left\|\frac{\partial(h_1 f)(x)}{\partial x}\right\| J(u)du\right] \\
&= E\left[\phi_n^2 \delta^2 \left\|\frac{\partial(h_1 f)(x)}{\partial x}\right\| \sigma_u^2\right] \\
&= o_p(1).
\end{aligned}$$

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

$$\begin{aligned}
&N^{-1/2}\sum_{n=1}^N \left\{E\left[Z_n h_{1n}\delta^{-q}\phi_m J\left(\frac{x_m - x_n}{\delta}\right)\middle| x_n\right] - E\left[Z_n h_{1n}\delta^{-q}\phi_m J\left(\frac{x_m - x_n}{\delta}\right)\right]\right\} \\
&= N^{-1/2}\sum_{n=1}^N \{Z_n h_{1n} f(x_n) v_n - E[Z_n h_{1n} f(x_n) v_n]\} + o_p(1).
\end{aligned}$$

Similarly, considering the third term in (A.22)

$$\begin{aligned}
&N^{-1/2}\sum_{n=1}^N \left\{E\left[Z_m h_{1m}\delta^{-q}\phi_n J\left(\frac{x_m - x_n}{\delta}\right)\middle| x_n\right] - E\left[z_m h_{1m}\delta^{-q}\phi_n J\left(\frac{x_m - x_n}{\delta}\right)\right]\right\} \\
&= N^{-1/2}\sum_{n=1}^N \{Z_n h_{1n} f(x_n)\phi_n - E[Z_n h_{1n} f(x_n)\phi_n]\} + o_p(1).
\end{aligned}$$

It now follows that

$$N^{-1/2}\sum_{n=1}^{N-1}\{E[v_{1mn}^N|n]-E[v_{1mn}^N]\}=N^{-1/2}\sum_{n=1}^{N-1}\{Z_nh_{1n}f(x_n)(v_n+\phi_n)-Z_nh_{1n}v_n$$
$$-E[Z_nh_{1n}f(x_n)(v_n+\phi_n)+Z_nh_{1n}v_n]\}+o_p(1). \tag{A.24}$$

By a similar argument,

$$N^{-1/2}\sum_{n=1}^{N-1}\{E[v_{2mn}^N|n]-E[v_{2mn}^N]\}=N^{-1/2}\sum_{n=1}^{N-1}\{Z_nh_{2n}f(x_n)(p_n+d_n)-Z_nh_{2n}p_n$$
$$-E[Z_nh_{2n}f(x_n)(p_n+d_n)+Z_nh_{2n}p_n]\}+o_p(1). \tag{A.25}$$

||

Lemma 2. $\sqrt{N}\,(\Theta^{(N)}-\tilde{\Theta}^{(N)})$ *is* $o_p(1)$.

Proof. Expanding the right-side of (A.7) about Θ_0, the true value of the structural parameters, and $\{v_n, p_n\}_{n=1}^N$, the true value of the incidental parameters, we obtain

$$-1/N\sum_{n=1}^N Z_n^N[h_n+\tilde{h}_{0n}(\Theta^{(N)}-\Theta_0)]$$
$$=1/N\sum_{n=1}^N Z_n^N[\tilde{h}_{1n}(v_n^N-v_n)+\tilde{h}_{2n}(p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N))], \tag{A.26}$$

where the $\tilde{}$ superscript indicates that the associated partial derivative is evaluated at a point in the interior of the convex neighbourhood defined by the points (Θ_0, v_n, p_n) and $(\Theta^{(N)}, v_n^N, p_n^N)$. Subtracting (A.26) from (A.10), we obtain

$$1/N\sum_{n=1}^N(Z_n-Z_n^N)h_n-1/N\sum_{n=1}^N[Z_n\tilde{h}_{0n}(\Theta_0-\Theta^{(N)})-Z_nh_{0n}(\Theta_0-\tilde{\Theta}^{(N)})]$$
$$=1/N\sum_{n=1}^N\{[Z_n^N\tilde{h}_{1n}-Z_nh_{1n}](v_n^N-v_n)+Z_n^N\tilde{h}_{2n}[p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N)]$$
$$-Z_nh_{2n}[p^N(\Theta_0,v_n)-p(\Theta_0,v_n)]\}. \tag{A.27}$$

To investigate the asymptotic properties of the three expressions on the right-side of (A.27), we consider

$$1/N\sum_{n=1}^N[Z_n^N\tilde{h}_{0n}(\Theta_0-\Theta^{(N)})-Z_nh_{0n}(\Theta_0-\tilde{\Theta}^{(N)})]$$
$$=1/N\sum_{n=1}^N\{[Z_nh_{0n}(\Theta_0-\Theta^{(N)})-Z_nh_{0n}(\Theta_0-\tilde{\Theta}^{(N)})]+[Z_n^N\tilde{h}_{0n}-Z_nh_{0n}](\Theta_0-\Theta^{(N)})\}$$
$$=1/N\sum_{n=1}^N Z_nh_{0n}(\tilde{\Theta}^{(N)}-\Theta^{(N)})+o_p(1)(\Theta_0-\Theta^{(N)})$$
$$=E[Z_nh_{0n}](\tilde{\Theta}^{(N)}-\Theta^{(N)})+o_p(1)+o_p(1)(\Theta_0-\Theta^{(N)}) \tag{A.28}$$
$$=\{E[Z_nh_{0n}]+o_p(1)\}(\tilde{\Theta}^{(N)}-\Theta^{(N)})+o_p(1).$$

Considering the second expression in (A.27)

$$1/N\sum_{n=1}^N[Z_n^N\tilde{h}_{1n}-Z_nh_{1n}](v_n^N-v_n)=o_p(1)(1/N)\sum_{n=1}^N(v_n^N-v_n), \tag{A.29}$$

where the right-side of (A.27) follows from the fact that $Z_n^N\tilde{h}_{0n}$ converges in probability to Z_nh_{0n} uniformly in n. Appealing to the same U-statistic arguments used to justify the asymptotic normality of $N^{1/2}(\tilde{\Theta}^{(N)}-\Theta_0)$, one can show that $N^{-1/2}\sum_{n=1}^N(v_n^N-v_n)$ converges in distribution to a normal random variable which is $O_p(1)$. Therefore, (A.27) is $o_p(N^{1/2})$. Finally, the third expression may be written as

$$1/N\sum_{n=1}^N\{Z_n^N\tilde{h}_{2n}[p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N)]-Z_nh_{2n}[p_n^N(\Theta_0,v_n)-p_n(\Theta_0,v_n)]\}$$
$$=1/N\sum_{n=1}^N(Z_n^N\tilde{h}_{2n}-Z_nh_{2n})[p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N)]$$
$$+1/N\sum_{n=1}^N\{Z_nh_{2n}[p_n^N(\Theta^{(N)},v_n^N)-p_n^N(\Theta_0,v_n)]-Z_nh_{2n}[p_n(\Theta^{(N)},v_n^N)-p_n(\Theta_0,v_n)]\}. \tag{A.30}$$

Taking a first-order Taylor series expansion to $p_n^N(\Theta^{(N)},v^N)$ and $p_n(\Theta^{(N)},v^N)$ around (Θ_0, v_n), we obtain

$$1/N\sum_{n=1}^N(Z_n^N\tilde{h}_{2n}-Z_nh_{2n})[p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N)]$$
$$+1/N\sum_{n=1}^N Z_nh_{2n}[p_{1n}^N(\Theta',v_n')(\Theta^{(N)}-\Theta_0)+p_{2n}^N(\Theta',v_n')(v_n^N-v_n)]$$
$$-1/N\sum_{n=1}^N Z_nh_{2n}[p_{1n}^N(\Theta'',v_n'')(\Theta^{(N)}-\Theta_0)+p_{2n}^N(\Theta'',v_n'')(v_n^N-v_n)]$$
$$=1/N\sum_{n=1}^N(Z_n^N\tilde{h}_{2n}-Z_nh_{2n})[p_n^N(\Theta^{(N)},v_n^N)-p_n(\Theta^{(N)},v_n^N)]$$

$$+1/N\sum_{n=1}^{N} Z_n h_{2n}[p_{1n}^N(\Theta', v_n') - p_{1n}^N(\Theta'', v_n'')](\Theta^{(N)} - \Theta_0)$$
$$-1/N\sum_{n=1}^{N} Z_n h_{2n}[p_{2n}^N(\Theta', v_n') - p_{2n}^N(\Theta'', v_n'')](v_n^N - v_n), \tag{A.31}$$

where (Θ', v_n') and (Θ'', v_n'') are in the $o_p(1)$ neighbourhoods of (Θ_0, v_n) and the second equality on the right-side of (A.31) follows from a rearrangement of terms. But $(Z_n^N \hat{h}_{2n} - Z_n h_{2n})$ is uniformly $o_p(1)$, as is $Z_n h_{2n}[p_{1n}^N(\Theta', v_n') - p_{1n}^N(\Theta'', v_n'')]$ and $Z_n h_{2n}[p_{2n}^N(\Theta', v_n') - p_{2n}^N(\Theta'', v_n'')]$. Therefore, (A.31) can be simplified as

$$o_p(1)\,1/N\sum_{n=1}^{N}[p_n^N(\Theta^{(N)}, v_n^N) - p_n(\Theta^{(N)}, v_n^N) - (v_n^N - v_n)] + o_p(1)(\Theta^{(N)} - \Theta_0). \tag{A.32}$$

The first expression in (A.32) is $o_p(1)$ because the terms

$$N^{-1/2}\sum_{n=1}^{N}[p_n^N(\Theta^{(N)}, v_n^N) - p_n(\Theta^{(N)}, v_n^N) - (v_n^N - v_n)],$$

and

$$N^{-1/2}\sum_{n=1}^{N}(v_n^N - v_n),$$

are asymptotically normal. Using the results obtained for (A.28), (A.29) and (A.32) in (A.27), we thus establish that

$$0 = \{E[Z_n h_{0n})] + o_p(1)\}N^{1/2}(\tilde{\Theta}^{(N)} - \Theta^{(N)}) + o_p(1). \tag{A.33}$$

Noting that $E[Z_n h_{0n}]$ is invertible, we see from (A.33) that $\sqrt{N}(\tilde{\Theta}^{(N)} - \Theta^{(N)})$ is $o_p(1)$ as claimed. ||

B.

In part B of this Appendix, we describe in more detail the construction of our sample and the construction of the variables used in our study. We used data from the Family-Individual File of the Michigan Panel Study of Income Dynamics (PSID), Waves I through XIX. The Family-Individual File contains a separate record for each member of all households included in the survey in a given year. Each record is 25,236 characters, the first 23,984 characters containing (possibly duplicate) information about the household to which each individual belongs. The remaining part of the record contains information relating to various characteristics of the individual and, most important, to his or her status within the household and relationship to the household head.

One issue that must be addressed when using this dataset is the issue of respondents vs. nonrespondents. According to the definition used in the PSID, an individual is denoted a *main family nonresponse* in a given year if both the individual and his or her family are lost to the study in that year. Alternatively, an individual may be a *mover-out nonresponse* if he or she has left a family that is still included in the study in a given year. The individual may subsequently become response if he or she moves into a panel family or becomes a splitoff by forming a new panel family or household. Mover-out nonrespondents have some nonzero individual data in the year that they became nonresponse because they were part of a panel family in the year preceding the one when they became nonresponse. The nineteen-year Family-Individual Respondents File contains data on individuals (and families) that were respondents as of the 1986 interviewing year as well as individuals who became mover-out nonrespondents in that year. In our selection, we did not distinguish between respondents and individuals who had become mover-out nonrespondents during a given year.

We initially selected women who were respondents or mover-out nonrespondents as of the 1986 interviewing year by setting the individual-level variables "Relationship to Head" to head or wife, "Sex of the Individual" to female and the "Why Nonresponse" variable to the zero category, which denotes individuals who were still a member of a panel family. Out of an initial sample of 20,437 individuals included in the nineteen-year Family-Individual Respondents File of the PSID, this initial selection produced a sample of 5900 observations for the 1986 interviewing year. The corresponding number of observations for the interviewing years 1968 through 1985 are given by 2474, 2592, 2761, 2912, 3079, 3260, 3445, 3619, 3815, 3973, 4130, 4363, 4597, 4793, 4987, 5153, 5358 and 5652, respectively. Since individuals who had become nonrespondents as of 1986 either because they and their families were lost to the study or they were mover-out nonrespondents in years prior to the 1986 interviewing year are not included in the nineteen-year Family-Individual Respondents File, the number of observations increases with the interviewing years.

We lost additional observations due to missing data or inconsistent observations. We first describe the nature of the coding errors or missing data and then note the number of observations lost due each year due to such errors. The first type of coding errors occur with respect to the measure of annual hours and average hourly earnings, which are identical to the PSID variables of the same names. In the PSID data tapes, average hourly earnings for both husbands and wives are defined from the ratio of total labour income to total annual hours of work. We encountered cases (due to reporting or coding errors) for which annual hours were positive but average hourly earnings were zero or vice versa. There is also an issue about the way average hourly earnings was coded in 1968 vs. the remaining survey years. Taken by itself, this coding error leads to a loss of 980 person-years. In 1968, 9's were coded instead of 0's when the head or wife did not work for money and therefore had no hourly earnings. In the remaining years, average hourly earnings above 99·99 dollars were coded as 99·99 dollars. Taken by itself, this coding error results in a loss of 40 person-years.

The second type of coding errors occurs for the different measures of consumption in the PSID from which we construct our consumption measure. More precisely, our measure of food consumption expenditures for a given year is obtained by summing the values of annual food expenditures for meals at home, annual food expenditures for eating out, and the value of food stamps received for that year. We then measured consumption expenditures for year t by taking 0·25 of the value of this variable for year $t-1$ and 0·75 of its value for 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 food 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 in the 1968 survey year is 999 dollars while the relevant value for this variable in the subsequent years and for the value of food consumed at home and eating out is 9,999 dollars. Taken by itself, truncation of the different consumption variables results in a loss of 452 person-years. We also use variables describing various demographic characteristics of the women in our sample. First, we obtained the age of each women from the individual variables located in the latter part of the data records of the Family-Individual File. For this variable, a value of 99 indicates missing data. The age variable results in a loss of 74 person-years.

There are no separate individual variables describing the race of the individual or the region where they are currently residing. Hence, variables from the family portion of the data record must be used for this purpose. 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 1–4 correspond to the regions Northeast, North Central, 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 variable.

Third, we used the family variable "Race of the Household Head" to measure the race variable in our study. There is a family variable that records information about the race of the wife, but this variable was included in the PSID only for the interviewing years 1985 and 1986. Defining the race variable in our empirical study as the race of the household head should not create much measurement error because the women in our subsample are either household heads themselves or wives of such heads. For the interviewing years 1968–1970, the values of 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, this variable was coded such that values of 1–4 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, and a value of 9 denotes missing data. Missing data in this variable results in a 200 person-years.

We also used the family variables that indicate the educational attainment level of the household head or wife to measure the education variable. We did this because the variable "Completed Education" recorded in the individual part of the data record does not apply if the individual is a household head or wife. However, one difficulty in using the family level education variables is that if the individual was a wife of a PSID household head for the interviewing years 1969, 1970 or 1971, there is no information about her education attainment level because questions regarding the wife's completed education level were not asked for those years. A second difficulty is that the variables denoting the head's and wife's completed education level are not strictly comparable across the different waves of the PSID. Since 1975, information pertaining to advanced (graduate or professional) degrees as well as that pertaining to additional nonacademic training have been coded for this variable. Another noncomparability problem is that the question regarding difficulty in reading or writing was omitted from the coding of this variable after 1984. For both the head and wife, the coding of this variable is as follows: 1: 0–5 grades, 2: 6–8 grades, 3: 9–11 grades, 4: 12 grades and no further training, 5: 12 grades plus nonacademic training, 6: College but no degree, 7: College BA but no advanced degree, and 8: College and advanced or professional degree. For both the head's and wife's education variable, a value of 9 denotes missing data. Taken by itself, a loss of 2282 person-years can be attributed to the education variable.

The marital status of a woman in our subsample was determined from the marital status of the head. This variable was coded differently for the interviewing year 1968, on the one hand, and the remaining years on the other. For 1968, the values 1 through 5 denote the categories married, single, widowed, divorced, and separated, respectively. The value 8 denotes married but spouse absent, and 9 denotes missing data. After 1968, the sixth category is dropped. However, there is no loss of person-years that can be attributed to this variable. The number of individuals in a household and the total number of children within that household were also determined from the family level variables of the same name. In 1968, a code for missing data (equal to 99) was allowed for the first variable, but in other years, missing data were assigned. The second variable, which indicates the total number of children under 18 in the family regardless of their relationship to the head, was truncated above at the value of 9 for the interviewing years 1968 and 1971. After 1975, this variable denotes the actual number of children within the family unit.

Taken jointly, the existence of missing data or coding errors as described above reduced our effective sample for the calendar years 1967 through 1985 as follows: 2439, 2549, 2704, 2840, 2982, 3127, 3288, 3449, 3628, 3778, 3922, 4135, 4347, 4506, 4682, 4835, 5014, 5235, and 5444, respectively. The total loss of observations is that are jointly attributable to the different coding errors is less than the sum of observations lost due to each type of error. This is because some types of coding errors occur for the same observation.

We constructed some additional variables that were used as instruments. 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 the variable of the above variables showing the value of rent paid and rental value of free housing for a household. Finally, household income was measured from the PSID variable total family money income, which included taxable income of head and wife, total transfers of head and wife, taxable income of others in the family units, and their total transfer payments.[17] We also constructed variables that show the age distribution of children within the family. For the interviewing years 1975 to 1986, we were able to obtain the number of children in the family between the ages of 1–2, 3–5, and 6–13 from family-level variables which show the total number of children in these age groups who were currently in the family unit.

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 framework of Section 2. We accordingly measured real food consumption expenditures and real wages as the ratio of the nominal consumption expenditures and wages and the annual implicit price deflator for food consumption expenditures published in Table 7.12 of the National Income and Product Accounts. (See the U.S. Department of Commerce, Bureau of Economic Analysis publication *Business Statistics 1986*, a supplement to the *Survey of Current Business*.) On the other hand, we deflated variables such as the nominal value of home ownership or nominal family income by the implicit price deflator for total personal consumption expenditures. *Per capita* consumption of nondurables and services is measured using the relevant consumption annual consumption from the National Income and Product Accounts divided by the U.S. Census Bureau measure of the civilian population of all ages. Aggregate wages are measured as average weekly earnings per production or nonsupervisory worker on nonagricultural payrolls deflated by the consumption deflator for nondurables and services.

Acknowledgements. We have received helpful comments from three anonymous referees. Earlier versions of this paper have been presented at the Conference for the Empirical Applications of Structural Models, organized by the Econometric Society and SSRI, Madison, Wisconsin, May 1990; Specification Search and Robust Estimation in Micro Labor Markets, CentER Tilburg, June 1990; the Sixth World Congress of the Econometric Society, Barcelona, Spain, August 1990; Dynamics of European Labour Markets, Center for European Economic Research, Mannheim, Germany, March 1992; and at seminars at various universities. Grants from the National Science Foundation (SES-9108659 and SES-9108738) and the Minnesota Supercomputer Institute are gratefully acknowledged.

REFERENCES

AGUIRREGABIRIA, V. (1994), "Dynamic Models of Price and Inventory Decisions with Fixed Adjustment Costs" (Ph.D. Thesis, Universidad Complutense de Madrid).

ALTONJI, J., HAYASHI, F. and KOTLIKOFF, L. (1996), "Risk-sharing Between and Within Families", *Econometrica*, **64**, 261–294.

ALTUĞ, S. (1989), "Time-to-Build and Aggregate Fluctuations: Some New Evidence", *International Economic Review*, **30**, 889–920.

ALTUĞ, S. and MILLER, R. (1990), "Household Choices in Equilibrium", *Econometrica*, **58**, 543–570.

17. The issue of truncation from above also arises for the variables used to construct measures of the above variables. However, we did not eliminate any observations or person-years due to the existence of such upper limits because the fact that some of the variables used as instruments were truncated from above for certain years does not invalidate the use of these instruments.

BROWNING, M. (1992), "Children and Household Economic Behavior", *Journal of Economic Literature*, **30**, 1434–1476.

BROWNING, M. and MEGHIR, C. (1991), "The Effects of Male and Female Labor Supply on Commodity Demands", *Econometrica*, **59**, 925–952.

CARD, D. (1990), "Labour Supply With a Minimum Threshold", *Carnegie-Rochester Series on Public Policy*, **33**, 137–168.

COGAN, J. (1981), "Fixed Costs and Labour Supply", *Econometrica*, **49**, 945–964.

ECKSTEIN, Z. and WOLPIN, K. (1989a), "Dynamic Labor Force Participation of Women and Endogenous Work Experience", *Review of Economic Studies*, **56**, 375–390.

ECKSTEIN, Z. and WOLPIN, K. (1989b), "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey", *Journal of Human Resources*, **24**, 562–598.

EICHENBAUM, M., HANSEN, L. and SINGLETON, K. (1988), "A Time Series Analysis of Representative Agents Models of Consumption and Leisure Choice Under Uncertainty", *Quarterly Journal of Economics*, **103**, 51–78.

HANSEN, G. (1985), "Indivisible Labor and the Business Cycle", *Journal Of Monetary Economics*, **16**, 309–327.

HECKMAN, J. and MACURDY, T. (1980), "A Life Cycle Model of Female Labor Supply", *Review of Economic Studies*, **47**, 47–74.

HONORE, B. (1992), "Trimmed LAD and Least Squares Estimators of Truncated and Censored Regression Models with Fixed Effects", *Econometrica*, **60**, 533–566.

HOTZ, V., KYDLAND, F. and SEDLACEK, G. (1988), "Intertemporal Preferences and Labor Supply", *Econometrica*, **56**, 335–360.

HOTZ, V. and MILLER, R. (1988), "An Empirical Model of Life Cycle Fertility and Labor Supply", *Econometrica*, **56**, 91–118.

HOTZ, V. and MILLER, R. (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, **60**, 497–529.

HOTZ, V., MILLER, R., SANDERS, S. and SMITH, J. (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice", *Review of Economic Studies*, **61**, 265–289.

KILLINGSWORTH, M. (1983) *Labor Supply* (Cambridge: Cambridge University Press).

MACURDY, T. (1981), "An Empirical Model of Labor Supply in a Life-Cycle Setting", *Journal of Political Economy*, **89**, 1059–1085.

MACE, B. (1991), "Full Insurance in the Presence of Aggregate Uncertainty", *Journal of Political Economy*, **99**, 928–956.

MADDALA, G. (1983) *Limited Dependent and Qualitative Variables in Econometrics* (New York: Cambridge University Press).

MANKIW, N., ROTEMBERG, J. and SUMMERS, L. (1985), "Intertemporal Substitution in Macroeconomics", *Quarterly Journal of Economics*, **100**, 225–252.

MILLER, R. (1984), "Job Matching and Occupational Choice", *Journal of Political Economy*, **92**, 1086–1120.

MILLER, R. (1997), "Estimating Models of Dynamic Optimization with Microeconomic Data", in Pesaran, M. and Schmidt, P. (eds) *Handbook of Applied Econometrics 2, Microeconometrics* (London: Basil Blackwell), 247–299.

MILLER, R. and SANDERS, S. (1997), "Human Capital Development and Welfare Participation", *Carnegie-Rochester Conference Series on Public Policy*, **46**, 1–44.

MILLER, R. and SIEG, H. (1997), "A Microeconometric Comparison of Household Behavior Across Countries", *Journal of Business and Economic Statistics*, **15**, 237–253.

MROZ, T. (1987), "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumption", *Econometrica*, **55**, 765–800.

PAKES, A. (1986), "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", *Econometrica*, **54**, 755–784.

POWELL, J., STOCK, J. and STOKER, T. (1989), "Semiparametric Estimation of Index Coefficients", *Econometrica*, **57**, 1403–1430.

PRAKASA RAO, B. (1983) *Nonparametric Functional Estimation* (Orlando: Academic Press).

ROBINSON, P. (1987), "Asymptotically Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form", *Econometrica*, **55**, 875–892.

ROGERSON, R. (1988), "Indivisible Labor, Lotteries and Equilibrium", *Journal of Monetary Economics*, **21**, 3–16.

RUST, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, **55**, 999–1034.

RUST, J. (1994), "Structural Estimation of Markov Decision Processes", in Engle, R. and McFadden, D. (eds.) *Handbook of Econometrics* (Amsterdam: North Holland).

SHAW, K. (1989), "Life-Cycle Labor Supply with Human Capital Accumulation", *International Economic Review*, **30**, 431–456.

SMITH, J. (1980) *Female Labor Supply: Theory and Estimation* (Princeton: Princeton University Press).

TOWNSEND, R. (1994), "Risk and Insurance in Village India", *Econometrica*, **62**, 539–591.

WOLPIN, K. (1984), "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality", *Journal of Political Economy*, **92**, 852–874.

LINKED CITATIONS

You have printed the following article:

The Effect of Work Experience on Female Wages and Labour Supply
Sumru Altu#; Robert A. Miller
The Review of Economic Studies, Vol. 65, No. 1. (Jan., 1998), pp. 45-85.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199801%2965%3A1%3C45%3ATEOWEO%3E2.0.CO%3B2-B

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

[Footnotes]

[4] **The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey**
Zvi Eckstein; Kenneth I. Wolpin
The Journal of Human Resources, Vol. 24, No. 4. (Autumn, 1989), pp. 562-598.
Stable URL:
http://links.jstor.org/sici?sici=0022-166X%28198923%2924%3A4%3C562%3ATSAEOD%3E2.0.CO%3B2-R

[6] **Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects**
Bo E. Honoré
Econometrica, Vol. 60, No. 3. (May, 1992), pp. 533-565.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199205%2960%3A3%3C533%3ATLALSE%3E2.0.CO%3B2-2

[7] **Conditional Choice Probabilities and the Estimation of Dynamic Models**
V. Joseph Hotz; Robert A. Miller
The Review of Economic Studies, Vol. 60, No. 3. (Jul., 1993), pp. 497-529.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199307%2960%3A3%3C497%3ACCPATE%3E2.0.CO%3B2-B

[7] **A Simulation Estimator for Dynamic Models of Discrete Choice**
V. Joseph Hotz; Robert A. Miller; Seth Sanders; Jeffrey Smith
The Review of Economic Studies, Vol. 61, No. 2. (Apr., 1994), pp. 265-289.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199404%2961%3A2%3C265%3AASEFDM%3E2.0.CO%3B2-9

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

[13] **A Life Cycle Model of Female Labour Supply**
James J. Heckman; Thomas E. Macurdy
The Review of Economic Studies, Vol. 47, No. 1, Econometrics Issue. (Jan., 1980), pp. 47-74.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198001%2947%3A1%3C47%3AALCMOF%3E2.0.CO%3B2-%23

[13] **An Empirical Analysis of Life Cycle Fertility and Female Labor Supply**
V. Joseph Hotz; Robert A. Miller
Econometrica, Vol. 56, No. 1. (Jan., 1988), pp. 91-118.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198801%2956%3A1%3C91%3AAEAOLC%3E2.0.CO%3B2-O

[13] **Children and Household Economic Behavior**
Martin Browning
Journal of Economic Literature, Vol. 30, No. 3. (Sep., 1992), pp. 1434-1475.
Stable URL:
http://links.jstor.org/sici?sici=0022-0515%28199209%2930%3A3%3C1434%3ACAHEB%3E2.0.CO%3B2-B

[15] **Semiparametric Estimation of Index Coefficients**
James L. Powell; James H. Stock; Thomas M. Stoker
Econometrica, Vol. 57, No. 6. (Nov., 1989), pp. 1403-1430.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198911%2957%3A6%3C1403%3ASEOIC%3E2.0.CO%3B2-R

[15] **Conditional Choice Probabilities and the Estimation of Dynamic Models**
V. Joseph Hotz; Robert A. Miller
The Review of Economic Studies, Vol. 60, No. 3. (Jul., 1993), pp. 497-529.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199307%2960%3A3%3C497%3ACCPATE%3E2.0.CO%3B2-B

[16] **Conditional Choice Probabilities and the Estimation of Dynamic Models**
V. Joseph Hotz; Robert A. Miller
The Review of Economic Studies, Vol. 60, No. 3. (Jul., 1993), pp. 497-529.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199307%2960%3A3%3C497%3ACCPATE%3E2.0.CO%3B2-B

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

References

Household Choices in Equilibrium
Sumru Altug; Robert A. Miller
Econometrica, Vol. 58, No. 3. (May, 1990), pp. 543-570.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199005%2958%3A3%3C543%3AHCIE%3E2.0.CO%3B2-B

Children and Household Economic Behavior
Martin Browning
Journal of Economic Literature, Vol. 30, No. 3. (Sep., 1992), pp. 1434-1475.
Stable URL:
http://links.jstor.org/sici?sici=0022-0515%28199209%2930%3A3%3C1434%3ACAHEB%3E2.0.CO%3B2-B

The Effects of Male and Female Labor Supply on Commodity Demands
Martin Browning; Costas Meghir
Econometrica, Vol. 59, No. 4. (Jul., 1991), pp. 925-951.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199107%2959%3A4%3C925%3ATEOMAF%3E2.0.CO%3B2-8

Fixed Costs and Labor Supply
John F. Cogan
Econometrica, Vol. 49, No. 4. (Jul., 1981), pp. 945-963.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198107%2949%3A4%3C945%3AFCALS%3E2.0.CO%3B2-O

Dynamic Labour Force Participation of Married Women and Endogenous Work Experience
Zvi Eckstein; Kenneth I. Wolpin
The Review of Economic Studies, Vol. 56, No. 3. (Jul., 1989)
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198907%2956%3A3%3C%3ADLFPOM%3E2.0.CO%3B2-0

The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey
Zvi Eckstein; Kenneth I. Wolpin
The Journal of Human Resources, Vol. 24, No. 4. (Autumn, 1989), pp. 562-598.
Stable URL:
http://links.jstor.org/sici?sici=0022-166X%28198923%2924%3A4%3C562%3ATSAEOD%3E2.0.CO%3B2-R

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty
Martin S. Eichenbaum; Lars Peter Hansen; Kenneth J. Singleton
The Quarterly Journal of Economics, Vol. 103, No. 1. (Feb., 1988), pp. 51-78.
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28198802%29103%3A1%3C51%3AATSAOR%3E2.0.CO%3B2-V

A Life Cycle Model of Female Labour Supply
James J. Heckman; Thomas E. Macurdy
The Review of Economic Studies, Vol. 47, No. 1, Econometrics Issue. (Jan., 1980), pp. 47-74.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28198001%2947%3A1%3C47%3AALCMOF%3E2.0.CO%3B2-%23

Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects
Bo E. Honoré
Econometrica, Vol. 60, No. 3. (May, 1992), pp. 533-565.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199205%2960%3A3%3C533%3ATLALSE%3E2.0.CO%3B2-2

Intertemporal Preferences and Labor Supply
V. Joseph Hotz; Finn E. Kydland; Guilherme L. Sedlacek
Econometrica, Vol. 56, No. 2. (Mar., 1988), pp. 335-360.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198803%2956%3A2%3C335%3AIPALS%3E2.0.CO%3B2-W

An Empirical Analysis of Life Cycle Fertility and Female Labor Supply
V. Joseph Hotz; Robert A. Miller
Econometrica, Vol. 56, No. 1. (Jan., 1988), pp. 91-118.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198801%2956%3A1%3C91%3AAEAOLC%3E2.0.CO%3B2-O

Conditional Choice Probabilities and the Estimation of Dynamic Models
V. Joseph Hotz; Robert A. Miller
The Review of Economic Studies, Vol. 60, No. 3. (Jul., 1993), pp. 497-529.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199307%2960%3A3%3C497%3ACCPATE%3E2.0.CO%3B2-B

NOTE: *The reference numbering from the original has been maintained in this citation list.*

A Simulation Estimator for Dynamic Models of Discrete Choice
V. Joseph Hotz; Robert A. Miller; Seth Sanders; Jeffrey Smith
The Review of Economic Studies, Vol. 61, No. 2. (Apr., 1994), pp. 265-289.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28199404%2961%3A2%3C265%3AASEFDM%3E2.0.CO%3B2-9

An Empirical Model of Labor Supply in a Life-Cycle Setting
Thomas E. MaCurdy
The Journal of Political Economy, Vol. 89, No. 6. (Dec., 1981), pp. 1059-1085.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198112%2989%3A6%3C1059%3AAEMOLS%3E2.0.CO%3B2-2

Full Insurance in the Presence of Aggregate Uncertainty
Barbara J. Mace
The Journal of Political Economy, Vol. 99, No. 5. (Oct., 1991), pp. 928-956.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28199110%2999%3A5%3C928%3AFIITPO%3E2.0.CO%3B2-R

Intertemporal Substitution in Macroeconomics
N. Gregory Mankiw; Julio J. Rotemberg; Lawrence H. Summers
The Quarterly Journal of Economics, Vol. 100, No. 1. (Feb., 1985), pp. 225-251.
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28198502%29100%3A1%3C225%3AISIM%3E2.0.CO%3B2-2

Job Matching and Occupational Choice
Robert A. Miller
The Journal of Political Economy, Vol. 92, No. 6. (Dec., 1984), pp. 1086-1120.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198412%2992%3A6%3C1086%3AJMAOC%3E2.0.CO%3B2-4

A Microeconometric Comparison of Household Behavior between Countries
Robert A. Miller; Holger Sieg
Journal of Business & Economic Statistics, Vol. 15, No. 2, Structural Estimation in Applied Microeconomics. (Apr., 1997), pp. 237-253.
Stable URL:
http://links.jstor.org/sici?sici=0735-0015%28199704%2915%3A2%3C237%3AAMCOHB%3E2.0.CO%3B2-7

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions
Thomas A. Mroz
Econometrica, Vol. 55, No. 4. (Jul., 1987), pp. 765-799.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198707%2955%3A4%3C765%3ATSOAEM%3E2.0.CO%3B2-S

Patents as Options: Some Estimates of the Value of Holding European Patent Stocks
Ariel Pakes
Econometrica, Vol. 54, No. 4. (Jul., 1986), pp. 755-784.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198607%2954%3A4%3C755%3APAOSEO%3E2.0.CO%3B2-X

Semiparametric Estimation of Index Coefficients
James L. Powell; James H. Stock; Thomas M. Stoker
Econometrica, Vol. 57, No. 6. (Nov., 1989), pp. 1403-1430.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198911%2957%3A6%3C1403%3ASEOIC%3E2.0.CO%3B2-R

Asymptotically Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form
P. M. Robinson
Econometrica, Vol. 55, No. 4. (Jul., 1987), pp. 875-891.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198707%2955%3A4%3C875%3AAEEITP%3E2.0.CO%3B2-A

Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher
John Rust
Econometrica, Vol. 55, No. 5. (Sep., 1987), pp. 999-1033.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198709%2955%3A5%3C999%3AOROGBE%3E2.0.CO%3B2-N

Life-Cycle Labor Supply with Human Capital Accumulation
Kathryn L. Shaw
International Economic Review, Vol. 30, No. 2. (May, 1989), pp. 431-456.
Stable URL:
http://links.jstor.org/sici?sici=0020-6598%28198905%2930%3A2%3C431%3ALLSWHC%3E2.0.CO%3B2-C

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

Risk and Insurance in Village India
Robert M. Townsend
Econometrica, Vol. 62, No. 3. (May, 1994), pp. 539-591.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199405%2962%3A3%3C539%3ARAIIVI%3E2.0.CO%3B2-9

An Estimable Dynamic Stochastic Model of Fertility and Child Mortality
Kenneth I. Wolpin
The Journal of Political Economy, Vol. 92, No. 5. (Oct., 1984), pp. 852-874.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198410%2992%3A5%3C852%3AAEDSMO%3E2.0.CO%3B2-F

NOTE: *The reference numbering from the original has been maintained in this citation list.*