Econometrica, Vol. 67, No. 6 (November, 1999), 1255–1294

STATE DEPENDENCE, SERIAL CORRELATION AND HETEROGENEITY IN INTERTEMPORAL LABOR FORCE PARTICIPATION OF MARRIED WOMEN

BY DEAN R. HYSLOP[1]

A dynamic search framework is developed to analyze the intertemporal labor force participation behavior of married women, using longitudinal data to allow for a rich dynamic structure. The sensitivity to alternative distributional assumptions is evaluated using linear probability and probit models. The dynamic probit models are estimated using maximum simulated likelihood (MSL) estimation, to overcome the computational difficulties inherent in maximum likelihood estimation of models with nontrivial error structures. The results find that participation decisions are characterized by significant state dependence, unobserved heterogeneity, and negative serial correlation in the error component. The hypothesis that fertility decisions are exogenous to women's participation decisions is rejected when dynamics are ignored; however, there is no evidence against this hypothesis in dynamic model specifications. Women's participation response is stronger to permanent than current nonlabor income, reflecting unobserved taste factors.

KEYWORDS: Intertemporal labor force participation, simulation estimation, state dependence, heterogeneity.

1. INTRODUCTION

THE INTERTEMPORAL LABOR SUPPLY BEHAVIOR of women remains the least studied area of labor supply research.[2] In contrast to male labor supply, the analysis of intertemporal female labor supply is complicated by the importance of considering behavior at both the extensive (participation) and the intensive (hours) margins.[3] Progress in this area has been hampered by the computational burden associated with classical maximum likelihood estimation (MLE) of nonlinear limited dependent variable models for intertemporal choice decisions. Although certain restrictive models can be estimated, high order integration is required to evaluate the likelihood function of general specifications. However,

[1] I thank Orley Ashenfelter, Michael Boozer, Kenneth Chay, Paul Devereux, Jim Heckman, Bo Honoré, Guido Imbens, Paul Ruud, Thomas Sargent, Duncan Thomas, an editor, three anonymous referees, and seminar participants at Chicago, ITAM, Pompeu Fabra, Princeton, UCLA, and Wisconsin for many helpful discussions and comments. I am also grateful to Orley Ashenfelter, Henry Farber, and Kevin Hallock for providing access to their CPU cycles for the Monte Carlo simulation experiments reported in this paper. The financial support of a UCLA Foundation Grant is gratefully acknowledged. All remaining errors in the paper are my responsibility.

[2] See for example MaCurdy (1981) and Altonji (1986) for analyses of intertemporal male labor supply; Heckman and MaCurdy (1980, 1982) and Mroz (1987) for female labor supply; and the surveys by Killingsworth (1983), Killingsworth and Heckman (1986), and Pencavel (1986).

[3] Although female labor force participation has been steadily increasing, the average annual participation rate of women remains substantially lower than males. For example, for a sample of married couples from the Panel Study of Income Dynamics (PSID), the participation rates for men and women are approximately 93 percent and 70 percent respectively, while the changes in annual participation status are approximately 3 percent and 13 percent respectively.

recent advances in simulation techniques for the estimation of multivariate limited dependent variable models have significantly reduced these computational difficulties.[4]

In this paper I analyze the intertemporal participation behavior of married women using a seven year longitudinal sample from the Panel Study of Income Dynamics (PSID). The participation decision is a natural starting point for analyzing female intertemporal labor supply for two reasons. Analyses by Cogan (1981) and Mroz (1987) provide evidence that the restrictions on the participation and hours decisions implied by a simple Tobit specification for labor supply are violated by significant fixed costs of labor supply. More general selection-correction models are required to analyze both margins of the female labor supply decision adequately. Also, in a recent survey of labor supply literature, Heckman (1993) concludes that labor supply response is strongest at the participation margin.

The analysis focuses on the relationships between the participation decisions and both fertility decisions and women's nonlabor income. Using cross-sectional data, Mroz (1987) concludes that, conditional on participation, fertility and nonwife income are both exogenous to women's hours of work decisions. In contrast, using panel data and a Tobit specification, Jakubson (1988) rejects the exogeneity of fertility. In a comprehensive review of fertility and labor supply issues, Browning (1992) emphasizes the importance of controlling for the dynamic structure of labor supply decisions in evaluating the interaction between fertility and labor supply decisions. The analysis treats husbands' labor supply outcomes as exogenous to their wives decisions. Husbands' labor earnings are used as a proxy for nonlabor income, and a simple stationary specification for nonlabor income is adopted in order to distinguish the direct effect of nonlabor income on participation decisions from possible endogenous taste or expectations effects.

A salient feature of annual participation behavior is the high degree of serial persistence in individual participation decisions. In a series of papers, Heckman (1978, 1981a, 1981c) discusses alternative sources of this serial persistence. Heckman distinguishes state dependence, whereby an individual's propensity to participate is changed *because* of past participation, from serial persistence due to persistent individual heterogeneity which may cause participation propensities to differ irrespective of past participation. Several sources of state dependence have been considered in the literature, such as intertemporally nonseparable preferences for leisure (e.g., Hotz, Kydland, and Sedlacek (1988)), human capital accumulation (e.g., Heckman (1981a)), and search costs which differ across participation states (e.g., Eckstein and Wolpin (1990)). In this paper, a simple search model of optimizing behavior under uncertainty is used to derive the common dynamic first-order Markov model for intertemporal participation decisions, in which an individual's current participation only depends structurally on their previous participation state. The empirical analysis

[4]Gourieroux and Monfort (1993), Hajivassiliou (1993), Hajivassiliou and Ruud (1994a), and Keane (1993) provide comprehensive surveys of the simulation literature.

also controls for observed and unobserved permanent (heterogeneity) or transitory (serial correlation) individual differences in the propensity to participate, and provides a rich dynamic structure.

The robustness of the identification of alternative sources of persistence to various econometric model specifications is examined using both linear probability and probit specifications. Although most econometric analysis of discrete choice models focuses on nonlinear specifications, the linear probability model has several attractions for applied researchers. In particular, it provides semiparametric identification and is relatively flexible in handling unobserved heterogeneity. A random effects specification is used in the probit model to control for unobserved heterogeneity which may be correlated with a family's fertility and income. The probit models are estimated using the method of maximum simulated likelihood (MSL) estimation.[5]

The results from both the linear and nonlinear models find that the participation decisions are characterized by substantial unobserved heterogeneity and positive state dependence. In specifications that allow for state dependence, both sets of analyses also find statistically significant negative serial correlation in the transitory errors. The consistency of results across the linear and nonlinear specifications provides confidence in the identification strategies employed. Controlling for dynamic factors is found to be important for the substantive issues of interest. Ignoring dynamic factors, the exogeneity of fertility to participation decisions is rejected. However, in the dynamic specifications that allow for serially correlated latent effects and/or state dependence, there is no evidence against the exogeneity hypothesis. The direct income effect on participation is negative but small, with an elasticity of about -0.04. The effect of permanent nonlabor income, which also reflects unobserved taste and expectations effects, is found to be stronger, with an elasticity of about -0.2. Although MSL estimation is known to be inconsistent for a finite number of replications, Monte Carlo evidence suggests that any bias associated with the MSL estimates presented is probably modest.

The paper is organized as follows. In the next section, a theoretical search model is presented and used to derive a dynamic specification for intertemporal participation decisions in the presence of search costs. Section 3 discusses the data set used in the analysis and presents a description of the relationship between intertemporal participation behavior and various demographic characteristics. In Section 4, we discuss the alternative specifications to be adopted, various econometric and identification issues, and simulated estimation tech-

[5]In many areas where simulation techniques could be profitably employed, the question of whether the high degree of observed serial persistence in individual outcomes reflects state dependence, or the effects of unobserved heterogeneity or serial correlation in latent factors affecting the decision, is important for an understanding of the behavioral relationships underlying the decision variable. To date, the literature on simulation techniques for limited dependent variables models has concentrated on specifications that allow for serially correlated error structures but no state dependence. Two exceptions are Hajivassiliou and McFadden (1998), who allow for state dependence in an analysis of debt repayments, and Mühleisen (1992), who models individual unemployment patterns allowing for state dependence.

niques for general multivariate discrete choice models. Section 5 presents the empirical results and simulations from the alternative specifications and Section 6 concludes.

2. A FRAMEWORK FOR INTERTEMPORAL PARTICIPATION DECISIONS

In this section I present a simple dynamic programming model of search behavior under uncertainty, in which search costs associated with labor market entry and labor market opportunities differ according to the individual's participation state.[6] This framework generates the implication that the labor-force participation decisions of married women depend on whether or not their market wage offer exceeds their reservation wage, which in turn may depend on their past participation state.

The assumptions of the model are as follows. Hours of work are constant across jobs, so the labor supply decision concerns whether or not to participate in each period. In order to receive a job offer, nonparticipants must search, which incurs a cost. There is no on-the-job search, but a current participant receives a wage offer next period without searching. The utility function is intertemporally separable, husband's labor supply choice is exogenous, and current period flow utility is defined over the joint family consumption C_t, and her own leisure l_t (alternatively, her participation h_t). The expected present value of discounted utility over an infinite lifetime[7] is

$$U_t = \sum_{s=0}^{\infty} \frac{1}{(1+\rho)^s} E_t u(C_{t+s}, h_{t+s}, Z_{t+s}) \tag{1}$$

where $u(\cdot)$ is period flow utility, Z_t is a vector of characteristics of the family in period t, which captures observed and unobserved heterogeneity both across families and over time, and ρ is the rate of time preference. Assuming neither borrowing nor lending occurs,[8] (1) is maximized on a period-by-period basis, subject to the budget constraint

$$C_t = y_t + w_t h_t - \gamma_1(1 - h_{t-1}), \tag{2}$$

where the price of consumption is normalized to 1, y_t is nonlabor income, w_t is the wage, and γ_1 is the cost of search.

[6]See, for example, Eckstein and Wolpin (1989a, 1989b, and 1990), Blundell, Ham, and Meghir (1995), and Burdett and Mortensen (1978). In contrast to Blundell et al. and Burdett and Mortensen, the nonparticipation and search states (i.e. participating but not employed) are not distinguished. I also abstract from any fixed costs associated with working. If such costs do not depend on a person's previous labor market state, these could be subsumed into the wage rate in this framework and will not affect the decision choice.

[7]Burdett and Mortensen (1978) discuss this assumption. Except for individuals near to retirement, it is likely to give a reasonable approximation to the finite horizon case. Alternatively, the rate of time preference may also capture the probability of retirement in each period.

[8]See, for example, Eckstein and Wolpin (1989a, 1990) in labor supply models, and Rust and Phelan (1997) in a model of retirement behavior for similar assumptions.

Assuming that both wages and nonlabor income are stationary, together with the infinite horizon and static budget constraints, implies that the value function is stationary. The value function at the beginning of period t given the participation state variable h_{t-1}, is described by $V(h_{t-1}, Z_t) = \max(V^0(h_{t-1}, Z_t), V^1(h_{t-1}, Z_t))$, where the superscripts 0 and 1 denote period t nonparticipation and participation states respectively. The dynamic program for a period $(t-1)$ nonparticipant is $V(0, Z_t) = \max(V^0(0, Z_t), V^1(0, Z_t))$. Comparing $V^0(0, Z_t)$ with $V^1(0, Z_t)$ implies that a nonparticipant's reservation wage w^*_{0t}, is defined by $V^0(0, Z_t) = V^1(0, Z_t | w^*_{0t})$, or

$$\text{(3a)} \qquad u(y_t + w^*_{0t} - \gamma_1, 1, Z_t) + \frac{1}{1+\rho} E_t V(1, Z_{t+1}) = u(y_t - \gamma_1, 0, Z_t) + \frac{1}{1+\rho} E_t V(0, Z_{t+1})$$

and participation occurs in period t if $w_t > w^*_{0t}$. Similarly, the dynamic program for a period $(t-1)$ participant is $V(1, Z_t) = \max(V^0(1, Z_t), V^1(1, Z_t))$. A participant's reservation wage, w^*_{1t} is defined by $V^0(1, Z_t) = V^1(1, Z_t | w^*_{1t})$ or

$$\text{(3b)} \qquad u(y_t + w^*_{1t}, 1, Z_t) + \frac{1}{1+\rho} E_t V(1, Z_{t+1}) = u(y_t, 0, Z_t) + \frac{1}{1+\rho} E_t V(0, Z_{t+1})$$

and participation occurs in period t if $w_t > w^*_{1t}$.

A comparison of the reservation wage expressions for nonparticipants and participants, (3a) and (3b), implies

$$u(y_t + w^*_{1t}, 1, Z_t) - u(y_t + w^*_{0t} - \gamma_1, 1, Z_t) = u(y_t, 0, Z_t) - u(y_t - \gamma_1, 0, Z_t).$$

Taylor series expansions of the left and right hand sides around $y_t + w^*_{0t}$ and y_t respectively, gives

$$w^*_{1t} \approx w^*_{0t} - \gamma_1 \left(1 - \frac{u_1(y_t, 0, Z_t)}{u_1(y_t + w^*_{0t}, 1, Z_t)} \right) = w^*_{0t} - \gamma$$

where $u_1(\cdot)$ is the marginal utility of consumption, and

$$\gamma = \gamma_1 \left(1 - \frac{u_1(y_t, 0, Z_t)}{u_1(y_t + w^*_{0t}, 1, Z)} \right).$$

Therefore, conditional on current and expected future realizations of the taste shifters, the period t participation decision can be characterized by

$$\text{(4)} \qquad h_t = 1(w_t > w^*_{0t} - \gamma h_{t-1})$$

where $1(\cdot)$ denotes an indicator function that is equal to 1 if the expression is true and 0 otherwise. In the presence of search costs ($\gamma_1 > 0$), the reservation wage for participants will be greater or less than for nonparticipants as $u_1(y_t, 0, Z_t)$ is greater or less than $u_1(y_t + w_{0t}^*, 1, Z_t)$. Assuming concave utility, $u_1(y_t, 0, Z_t) > u_1(y_t + w_{0t}^*, 0, Z_t)$. However, if the marginal utility of consumption is greater when working ($u_{12} > 0$), then $u_1(y_t + w_{0t}^*, 0, Z_t) < u_1(y_t + w_{0t}^*, 1, Z_t)$, and there are two offsetting effects in comparing $u_1(y_t, 0, Z_t)$ and $u_1(y_t + w_{0t}^*, 1, Z_t)$, so that γ may be positive or negative.

3. DATA

The data used in the analysis are from the 1986 panel of the Panel Study of Income Dynamics (PSID) and pertain to the seven calendar years 1979–85, corresponding to waves 12–19 of the PSID. The sample consists of 1812 continuously married couples, aged between 18 and 60 in 1980, and the husband is a labor force participant in each of the sample years.[9] The sample includes both the random Census subsample of families and the nonrandom Survey of Economic Opportunities (SEO) subsample of families, which accounts for approximately one third of the PSID sample.

Table I presents summary statistics on a selection of variables of interest in the sample. Annual earnings are expressed in constant (1987) dollars, computed as nominal earnings deflated by the consumer price index. Column (1) describes the characteristics for the whole sample. The distribution of years worked during the period for the full sample indicates there is significant persistence in the observed annual participation decisions of married women. For example, to take an extreme case, if individual participation outcomes were independent draws from a binomial distribution with fixed probability of 0.7 (the average participation rate during the period), then about 8 percent of the sample would be expected to work each year, and almost no one would not work at all (0.02 percent). In contrast, the sample relative frequencies are 48 percent and 11 percent respectively.

The observed frequency distributions of the number of years worked and the associated participation sequences suggest there may be significant differences in the work propensity of women. One source of this heterogeneity is differences in observable characteristics such as age, race, education, nonlabor income, and the number of children.[10] For this reason, columns (2)–(6) present the charac-

[9] An individual is defined as a participant if they report both positive annual hours worked and annual earnings.

[10] The number of years of education is imputed from the following categorical scheme: 1 = '0–5 grades' (2.5 years); 2 = '6–8 grades' (7 years); 3 = '9–11 grades' (10 years); 4 = '12 grades' (12 years); 5 = '12 grades plus nonacademic training' (13 years); 6 = 'some college' (14 years); 7 = 'college degree, not advanced' (16 years); 8 = 'college and advanced degree' (18 years). Education is measured as the highest level reported in the 1980–86 surveys. The simple relationship between the participation rate and this measure of years of education is approximately linear over the sample. The labor earnings of the husband are used as a proxy for nonlabor income; as an empirical matter this seems reasonable—e.g. Cutler and Katz (1991) report that nonlabor income accounts for only 10 percent of total family income on average; also the reliability of nonlabor income in survey data is very low.

TABLE I

SAMPLE CHARACTERISTICS

	Full Sample (1)	Employed 7 Years (2)	Employed 0 Years (3)	Single Transition from Work (4)	Single Transition to Work (5)	Multiple Transitions (6)
Age	34.34	34.52	39.66	34.35	33.12	32.08
(1980)	(.23)	(.31)	(.81)	(.89)	(.67)	(.44)
Education[a]	12.90	13.26	11.86	12.85	12.90	12.67
	(.05)	(.08)	(.17)	(.21)	(.16)	(.11)
Race (1 = Black)	0.22	0.25	0.24	0.16	0.15	0.20
	(.01)	(.01)	(.03)	(.03)	(.03)	(.02)
No. Children[b]	0.25	0.20	0.23	0.33	0.25	0.32
aged 0–2 years	(.01)	(.01)	(.02)	(.03)	(.02)	(.02)
No. Children[b]	0.30	0.24	0.26	0.28	0.41	0.40
aged 3–5 years	(.01)	(.01)	(.03)	(.03)	(.03)	(.02)
No. Children[b]	1.00	0.96	0.97	0.60	1.32	1.08
aged 6–17 years	(.02)	(.03)	(.07)	(.08)	(.07)	(.05)
Husband's Earnings[b]	29.59	27.90	35.17	31.46	33.64	28.22
(1987 $1000)	(.47)	(.64)	(1.93)	(1.56)	(1.97)	(.72)
Participation[b]	0.70	1	0	0.46	0.55	0.57
	(.01)			(.02)	(.02)	(.01)
No. years worked[c]						
zero	10.6	—	100	—	—	—
one	6.1	—	—	24.7	15.3	11.1
two	5.4	—	—	19.2	14.8	10.4
three	5.7	—	—	14.4	12.5	14.1
four	6.7	—	—	9.6	12.5	20.0
five	8.8	—	—	13.0	19.3	24.9
six	8.6	—	—	19.2	25.6	19.5
seven	48.2	100	—	—	—	—
Sample size	1812	873	192	146	176	425

Notes: Standard errors in parentheses. Sample selection criteria: continuously married couples, aged 18–60 in 1980, with positive husband's annual earnings and hours worked each year.

[a]Years of Education are imputed from the following categorical scheme: 1 = '0–5 grades' (2.5 years); 2 = '6–8' (7 years); 3 = '9–11' (10 years); 4 = '12' (12 years); 5 = '12 plus non-academic training' (13 years); 6 = 'some college' (14 years); 7 = 'college degree, not advanced' (16 years); 8 = 'college advanced degree' (18 years). Education is measured as the highest level reported in the 1980–86 surveys.

[b]Sample Averages: child variables based on 8 observations; participation and male earnings based on 7 observations.

[c]Column percentages.

teristics for various subsamples based on the observed annual participation outcomes of women during the sample period. Given the large number of possible participation sequences (128) over a 7-period panel, I choose a small group of selection criteria for these subsamples which provide a useful descriptive analysis of the sample differences. The subsample in column (2) consists of women who work in each year; in column (3), women who never work during the sample period; in column (4), women who experience a single transition from employment to nonemployment—that is, participation sequences '1000000',..., '1111110';[11] in column (5), women who experience a single transition from nonemployment to employment; and in column (6), women who experience more than a single transition in their participation status.

The differences in the characteristics across the subsamples in Table I can be summarized as follows. Women who are employed in each year (column (2)) are better educated, are more likely to be black, have fewer dependent children (especially young children under 6 years), and their husbands' earnings are slightly lower than average. In column (3), women who are never employed are older, less educated, and their husbands' earnings are higher than average. Interestingly, this sample does not have more dependent children; in fact, families in this sample have slightly fewer young children, reflecting the older age of the sample. Women who make a single transition from employment to nonemployment (column (4)) are less likely to be black and have fewer dependent children but are more likely to have infant children (aged 0–2 years). In contrast, women who experience a single transition from nonemployment to employment (column (5)) have significantly more (noninfant) children, and their husbands have above average earnings. Finally, women who experience multiple employment transitions (column (6)) are younger, have more dependent children of all ages, and their husbands have below-average earnings.

The patterns in Table I are broadly consistent with commonly held notions of the determinants of female participation. There appears to be a positive income effect exerted by husbands' earnings on women's nonmarket time. The presence of children, especially young children, tends to reduce the participation of women, although women who never work in the sample do not have more children than average (column (3)). The differences in the numbers of very young and older children between the single-transition subsamples in columns (4) and (5) are consistent with the notion that women leave employment to have children and re-enter employment as their children approach school age. Nevertheless, there is apparently some other source of heterogeneity between these two samples, as there is no significant age difference between women in these samples, suggesting that the composition of these samples is determined by more than simply fertility considerations. Finally, column (6) suggests that the presence of children, together with low male earnings, increases the probability of frequent employment transitions of women.

4. EMPIRICAL SPECIFICATION AND ESTIMATION ISSUES

Expression (4) implies that a woman's current participation decision will depend on human capital factors via w_{it}, taste shifters via w^*_{0it}, and, in the presence of search costs, on her previous participation decision. A reduced form specification for (4) is adopted that depends on human capital and demographic characteristics of the individual. Attention focuses on identifying the effects of

[11] A '1' in the tth position of the sequence denotes participation in year t, while a '0' denotes nonparticipation.

fertility and nonlabor income, which is complicated by two factors. First, if the realizations of fertility and nonlabor income also reflect tastes for work, then these variables will be correlated with unobserved heterogeneity and endogenous with respect to participation decisions.[12] Second, state dependence in participation implies that the current decision to participate will depend on the expectations of future outcomes as well as current outcomes. In this section, I first discuss the empirical specification adopted, and how these two issues are dealt with. Model identification and estimation under a variety of distributional and econometric assumptions is then outlined.

4.1. *Empirical Specification and Identification*

The empirical specification for modelling intertemporal participation decisions involves the following dynamic reduced form specification for equation (4):

$$h_{it} = 1(X_{it}'\beta + \gamma h_{it-1} + u_{it} > 0) \qquad (i = 1, \dots, N; t = 1, \dots, T-1) \tag{5}$$

where X_{it} is a vector of observed human capital, demographic and family structure variables that may affect the participation outcome, u_{it} captures the effects of unobserved factors, and β and γ are parameters. Specifically, X_{it} contains time dummy variables and a set of demographic variables: a quadratic in age; a race dummy variable for whether the individual is black; years of education; fertility variables—the numbers of children aged 0–2, 3–5, and 6–17 years; and nonlabor income.[13] The unobserved term, u_{it}, is assumed to have the following structure:

$$u_{it} = \alpha_i + \varepsilon_{it}$$

where α_i is an individual-specific component, which captures time invariant unobserved human capital and taste factors; and ε_{it} is a possibly serially correlated error term, which captures factors such as transitory wage movements. Throughout the analysis, ε_{it} is assumed to be independent of X_{it}. However, if fertility and nonlabor income reflect tastes for work, then α_i will be correlated with X_{it}; I will discuss the handling of this issue shortly.

Model identification in the dynamic binary response model (5) follows by analogy to that of the dynamic continuous linear model. In the absence of time-varying regressors, identification of the state dependence effect relies strongly on functional form restrictions (Chamberlain (1984, pp. 1278–1279)). More robust identification, which relies on the dynamic response to changes in

[12] This is econometric endogeneity of fertility, in the sense that fertility may be exogenous to individuals but not to the econometrician, who only observes future fertility outcomes as they occur.

[13] The PSID survey records the demographic characteristics at the time of the survey, which is, on average, during March; while labor force participation and earnings refer to the previous calendar year. I assume all information in survey year t refers to year $(t-1)$. However, I also include the number of children aged 0–2 years in survey year $(t-1)$ as a covariate for year $(t-1)$ participation, in order to control for possible matching problems. A simple exploratory analysis suggested that similar controls for older children is not necessary.

the regressors, is obtained in the presence of time varying covariates. In the absence of state dependence, a transitory change in X_{it} causes (at most) a transitory change in the binary outcome h_{it}; while, in the presence of state dependence, a transitory change in X_{it} may have a persistent effect on h_{it}. If first-order state dependence is the only source of dynamics in the model, $P(h_{it}=1|h_{it-1}, X_{it}, \alpha_i)$ does not depend on lagged X_{it}; while, if the errors are serially correlated, $P(h_{it}=1|h_{it-1}, X_{it}, \alpha_i)$ will depend on lagged values of X_{it}. However, there is an important distinction between the binary and continuous response models for identifying state dependence due to the effect of the index threshold in binary response models (see Heckman (1981a)). While a change in a regressor *always* has an effect on the outcome in a continuous response setting, in a binary response model an effect will only occur if the latent index passes some threshold. This difference has two contrasting effects for identification. First, some minimum threshold-level of change in the regressor(s) must occur before an effect is observed. In the absence of state dependence, the effect on the index is continuous and transitory; while, in the presence of state dependence, a threshold-level change in the outcome will have a discrete effect on the index.

Two alternative estimation approaches are used to evaluate the identification of the models. First, linear probability models are considered, which control for arbitrary correlation between the unobserved heterogeneity (α_i) and the regressors, and can be used to eliminate the incidental parameters associated with the unobserved heterogeneity. Second, random effects probit models, which parameterize the distributions of α_i and ε_{it}, are estimated. There are advantages and disadvantages associated with each of these methods. The linear models are robust to the form of unobserved heterogeneity, and moreover avoid the problem of the initial conditions of the dynamic process. In contrast, the probit models rely more strongly on the functional form assumptions made, and are significantly more demanding computationally. However, the linear probability model does not constrain the predicted probabilities to the unit interval, and the nonlinear models are likely to provide a better fit. In addition, to the extent the assumptions are correct, the identification is stronger in nonlinear models.

In the presence of state dependence, expectations of future outcomes may also affect current participation decisions. In order to achieve a tractable empirical specification, the following assumptions are made with respect to the expectations of fertility and nonlabor income. First, a robust prediction is surely that expectations' effects decline into the future; as a simplification, I assume that only expectations of one-period-ahead realizations affect the current period participation decision. Second, I assume perfect foresight with respect to lifecycle fertility decisions, and include an indicator variable for whether a birth occurs next period.[14] Third, I adopt a simple stationary stochastic process for the nonlabor income process, in which expected future income is permanent

[14] The perfect foresight assumption is extreme, and rejected by Hotz and Miller (1988). However, if expectations matter, then we might expect that a future birth will affect current participation.

income.[15] Therefore, if transitory income is uncorrelated with tastes, then it will only have a direct "income" effect on participation, while the total effect of permanent income on participation will consist of this direct effect, an "expectations" effect, and a "tastes" effect. Permanent nonlabor income is estimated by the sample average, and transitory income is measured as deviations from the sample average.

If observed fertility and/or income is correlated with unobserved tastes, then α_i will be correlated with X_{it}. To handle this issue, two alternative approaches are adopted. A standard fixed-effects specification is used in the linear probability models. In addition, a correlated random-effects (CRE) specification is adopted in both the linear probability and probit frameworks:

$$\alpha_i = \sum_{s=0}^{T} (\delta_{1s}.(\#\text{Kids0–2})_{is} + \delta_{2s}.(\#\text{Kids3–5})_{is} + \delta_{3s}.(\#\text{Kids6–17})_{is}) + \sum_{s=0}^{T-1} \delta_{4s}.y_{mis} + \eta_i \tag{6}$$

where y_{mis} is i's transitory nonlabor income in year s.[16] In the probit specifications, η_i is independent of X_i, and $\eta_i|X_i \sim N(0, \sigma_\eta^2)$.[17] Under the above assumptions concerning fertility expectations, a test of the exogeneity of fertility with respect to participation decisions is provided by testing $\delta_{1s} = \delta_{2s} = \delta_{3s} = 0$ for all s. Similarly, under the assumptions concerning nonlabor income expectations, a test of $\delta_{4s} = 0$ provides a test of the exogeneity of current nonlabor income with respect to participation.

4.2. *Linear Probability Models*

Consider the following linear probability model specification for (5):

$$h_{it} = \gamma h_{it-1} + X_{it}'\beta + \alpha_i + \varepsilon_{it} \qquad (i = 1, \ldots, N; t = 1, \ldots, T-1). \tag{7}$$

First, suppose that ε_{it} is serially uncorrelated. The standard fixed-effects approach is to first difference equation (7) to eliminate α_i:

$$\Delta h_{it} = \gamma \Delta h_{it-1} + \Delta X_{it}'\,\beta + \Delta\varepsilon_{it}. \tag{8}$$

[15] A distinction between the effects of permanent and transitory income effects dates back to Mincer (1962). In a labor supply model, permanent income is the direct determinant of labor supply via the budget constraint, but Mincer argued that transitory income may affect labor supply decisions in the presence of credit market constraints. In the search framework adopted here, current (transitory) income is the direct determinant of labor supply via the budget constraint, while permanent income may affect labor supply indirectly via tastes.

[16] Due to the collinearity between y_{mp} and $(y_{m0}, \ldots, y_{mT-1})$, the restriction $\delta_{4T-1} = -\sum_{s=0}^{T-2} \delta_{4s}$ is imposed.

[17] In a panel data Tobit framework, Jakubson (1988) examines the issue of the endogeneity of fertility outcomes and labor supply decisions of women using both a CRE and a fixed effects, specification. Although the Tobit specification has been shown to be seriously misspecified (Mroz (1987)), and the assumptions required for nonlinear CRE are strong, Jakubson's results are robust to the shortcomings of both the CRE and fixed effects specifications.

Equation (8) can be consistently estimated by instrumenting for Δh_{it-1} in (8) using h_{it-2} or previous lags and noncontemporaneous realizations of the covariates.[18]

An alternative approach is to adopt a correlated random effects specification for α_i. Using (6), substitute for α_i in (7):

$$(7') \qquad h_{it} = \gamma h_{it-1} + X_{it}'(\beta + \delta_t) + \sum_{s \neq t} X_{is}'\delta_s + \eta_i + \varepsilon_{it},$$

where only the fertility and nonlabor δ coefficients are nonzero. Equation (7′) can be consistently estimated in levels by instrumenting for h_{it-1} using Δh_{it-1} or previous lags.[19]

If ε_{it} is serially correlated then, in general, Δh_{it-1} will be correlated with ε_{it} and h_{it-2} will be correlated with $\Delta \varepsilon_{it}$, so neither of these approaches will yield consistent estimates, although consistent estimation is possible using only out-of-period regressors as instruments if X_{it} is exogenous with respect to ε_{it}. However, if ε_{it} follows an AR(1) process:

$$\varepsilon_{it} = \rho \varepsilon_{it-1} + v_{it}, \quad -1 < \rho < 1, \quad v_{it} \sim (0, \sigma_v^2),$$

then consistent estimation is possible using a modified "levels" or first-difference approach. Equation (7) can first be partial-differenced to eliminate the serial correlation in the errors:

$$(7'') \qquad h_{it} = (\rho + \gamma) h_{it-1} - \rho\gamma h_{it-2} + X_{it}'\beta - X_{it-1}'\rho\beta + (1-\rho)\alpha_i + v_{it}.$$

Equation (7″) can then be consistently estimated by instrumenting for h_{it-1} and h_{it-2} using Δh_{it-1} and Δh_{it-2}. Alternatively, first-difference (6″) to eliminate the unobserved heterogeneity, giving

$$(8') \qquad \Delta h_{it} = (\rho + \gamma)\,\Delta h_{it-1} - \rho\gamma \Delta h_{it-2} + \Delta X_{it}\,\beta - \Delta X_{it-1}\,\rho\beta + \Delta v_{it}, \qquad (t = 3, \ldots, T-1),$$

where Δh_{it-1} is correlated with Δu_{it}, but now h_{it-2} is a valid instrument. The parameters (γ, β, ρ) can be estimated using a two-step procedure: first, estimate unrestricted reduced form coefficients on the regressors in either equation (7″) or (8′), and then estimate (γ, β, ρ) using minimum distance techniques to impose the restrictions on the reduced form coefficients.

[18] There is a substantial literature concerned with efficient estimation of linear dynamic model specifications. For example, Arellano and Bond (1991) extend the instrument set in period t to include $h_{i0}, \ldots, h_{it-2}$; while Ahn and Schmidt (1995) suggest including a levels equation, which imposes covariance restrictions.

[19] This "levels-approach" requires the additional assumption of stationarity: $E(\alpha_i h_{it})$ is constant over time. Out-of-period realizations of the regressors will also be valid instruments, if the restrictions from the correlated random effects specification are imposed.

4.3. *Nonlinear Models*

The nonlinear random effects approach to estimating (5) requires that the distributional properties of α_i and ε_{it}, their statistical relationship to the regressors, and also the initial conditions for the dynamic process be specified. Common approaches to the initial conditions are either to assume that h_{i0} is exogenous and can be treated as fixed (e.g., Heckman (1978, 1981a, 1981c)) or that the process generating h_{it} is in equilibrium at the beginning of the sample period (e.g., Card and Sullivan (1988)). For many economic series neither of these assumptions is particularly appealing. An alternative approach explored by Heckman (1981b), is to use a flexible reduced form approach to approximate the initial conditions. In a latent linear framework, a natural approach to the initial period equation is to adopt a linear specification for h_{i0} in terms of the initial period regressors (X_{i0}), and allow the initial period error (u_{i0}) to be arbitrarily correlated with other period errors ($\alpha_i + \varepsilon_{it}$).

I adopt a probit specification for the model. The random effects are assumed to be normally distributed, conditional on a linear function of the regressors. In particular,

$$h_{i0} = 1(X_{i0}\beta_0 + u_{i0} > 0), \tag{9a}$$

$$h_{it} = 1(X_{it}\beta + \gamma h_{it-1} + u_{it} > 0), \quad \text{and} \tag{9b}$$

$$u_{it} = \alpha_i + \varepsilon_{it} \qquad (i = 1,\dots,N; \text{ and } t = 1,\dots,T-1).$$

where $u_{i0} \sim N(0, \sigma_0^2)$; α_i is a correlated random effect, specified in (6); ε_{it} is an AR(1) error component: $\varepsilon_{it} = \rho\varepsilon_{it-1} + v_{it}$, $v_{it} \sim N(0, \sigma_v^2)$, orthogonal to η_i; $\text{corr}(u_{i0}, u_{it}) = \rho_t$, $t = 1,\dots,T-1$; and, for identification, $\sigma_\eta^2 + \sigma_\varepsilon^2 = 1$ and $\sigma_0^2 = 1$.

Under certain restrictive assumptions the model can be estimated by MLE. For example, in the absence of state dependence, models with single factor error structures can be estimated by MLE using Gaussian quadrature procedures (Butler and Moffitt (1982), and Heckman (1981a)). However, estimation of the unrestricted dynamic model in equations (9a) and (9b) using MLE methods requires the evaluation of T-dimensional integrals of Normal density functions. For $T > 3$ the computational burden renders such estimation infeasible.[20] Recently developed simulation-based estimation techniques provide a feasible method to overcome this problem.[21]

Simulation-based estimation methods for LDV models generally take one of two approaches: direct simulation of the likelihood function, or indirect likelihood simulation, by simulating an expression for the score of the likelihood. Let the log-likelihood function for the unknown parameter vector θ, given the random sample of observations $(z_1,\dots,z_N)$, be $\ln 1_N(\theta) = \Sigma_{i=1}^N \ln 1(\theta; z_i)$. Let $\{\xi_i\} = \{\xi_{i1},\dots,\xi_{iR}\}$ be a sequence of primitive simulators, independent of the

[20] See Hajivassiliou and Ruud (1994a, pp. 2399–2400) for an example of this.

[21] Lerman and Manski (1981) provides the earliest econometric work on simulation estimators.

parameters of the model:[22] $\tilde{1}(\theta; z_i, \xi_i) = (1/R)\Sigma_{r=1}^{R}\tilde{1}(\theta; z_i, \xi_{ir})$, where $\tilde{1}(\theta; z_i, \xi_{ir})$ is an unbiased simulator for $1(\theta; z_i)$, and R is the number of replications. The maximum simulated likelihood (MSL) estimator for θ is then

$$\hat{\theta}_{MSL} = \underset{\theta}{\operatorname{argmax}} \ln \tilde{1}_N(\theta)$$

where $\ln \tilde{1}_N(\theta) = \Sigma_{n=1}^{N} \ln \tilde{1}(\theta, z_i, \xi_i)$.

MSL estimation involves first obtaining an unbiased simulator for the likelihood function and then maximizing the resulting log simulated likelihood function instead of the actual log-likelihood. The simulator used here is the smooth recursive conditioning (SRC) simulator, which is continuous in the parameters, bounded away from zero and one, and has been found to be very accurate.[23] Although an unbiased simulator for the likelihood is straightforward to obtain with a finite number of replications, the log (simulated) likelihood function will be biased, so that the MSL estimator obtained will be inconsistent for finite R. However, MSL is consistent if the number of replications $R \to \infty$ as the sample size $N \to \infty$, and is asymptotically efficient if $R/\sqrt{N} \to \infty$ (Hajivassiliou and Ruud (1994b)).

Two alternative simulation estimators have been proposed. McFadden (1989) proposed a method of simulated moments (MSM) estimator in which the score of the likelihood is first expressed as a moment condition, this moment condition is then simulated, and the MSM estimator solves for the root of this simulated moment condition.[24] In addition, Hajivassiliou and McFadden (1998) have proposed a method of simulated scores (MSS) estimator which solves for the root of the simulated score function directly. The principal advantage of MSM and MSS is that consistent estimators can be obtained using a finite number of replications if the score function, or the moment condition, can be simulated without bias. However, although MSL estimation is inconsistent for a finite number of replications, it does offer some advantages over MSM and MSS in practice. The MSL method is comparatively simple to implement: in comparison to classical MLE, the only difficulty is the need to simulate the choice probabilities, which is a straightforward task for discrete choice models based on multivariate normal random variables (see Appendix 1). In contrast, MSM and

[22] The asymptotic theory developed for the simulation estimators, and also estimator convergence, requires that the same fixed values of the primitive random draws, ξ_i, be used at each iteration of the estimation. In practice, $\xi_{ir} \sim U(0,1)$, which is transformed to provide a random draw from the appropriate distribution that depends on the model parameters (see Appendix 1).

[23] The SRC simulator is also referred to in the literature as the Geweke, Hajivassiliou, and Keane (GHK) simulator. For a detailed discussion of the SRC simulator and its properties, see Börsch-Supan and Hajivassiliou (1993), Keane (1993), and Keane (1994). Appendix 1 contains a brief description of the SRC simulator.

[24] McFadden (1989) and Pakes and Pollard (1989) establish the statistical properties of these simulation estimators.

MSS often require significant manipulation of the score function in order to implement the simulation and estimation procedures. Second, as long as a smooth simulator for the choice probabilities is employed that bounds the likelihood function away from zero and one, the simulated likelihood function, like the true likelihood function, is stable so that MSL estimation is computationally robust. In contrast, MSM estimation has often been found to be numerically unstable.[25] Third, efficient MSM estimation requires a procedure for obtaining the optimal weights $w_i = 1_{i\theta}(\theta^0)/1_i(\theta^0)$. This imposes an additional computational burden on the procedure and, if the optimal weights are also obtained by simulation, additional simulation noise. In fact, McFadden and Ruud (1994) recommend against simulating both the weights and the unobserved choice probabilities, as the resulting increase in simulation noise tends to outweigh any efficiency gain. Similarly, although MSS implicitly simulates the optimal weights for MSM, unbiased simulation of the score requires independent simulations for the numerator and the denominator of the score, which increase the computational burden relative to MSL.

For these reasons, the estimation strategy adopted here is to use MSL estimation with the SRC simulator, while Appendix 2 describes a simulation experiment to assess the properties of the MSL estimator used here.

5. RESULTS

In this section I present the results for a variety of empirical specifications of the labor force participation models discussed in Section 4. I first present estimates from linear probability models corresponding to specifications (7) and (8). Following this several nonlinear probit model specifications, corresponding to equations (9a) and (9b), are estimated: first, static models that allow for unobserved heterogeneity but no state dependence or serially correlated errors and may be estimated using classical methods;[26] and then specifications that also include serially correlated errors and structural state dependence. Finally, an evaluation of these alternative specifications and comparison of the predictions is considered.

[25] For example, McFadden and Ruud (1994, p. 22) state "Whereas both ML and MSL will avoid fitted values near zero, the MSM equations can promote them because the numerator in the weight may go to zero while the denominator is bounded away from zero. A solution...may exist on the boundary of the parameter space merely because all the weights can be driven to zero there." Also, Mühleisen (1991) reports difficulties in obtaining MSM to converge; and, in cases in which convergence was obtained, the resulting covariance matrix of the parameters was greatly overestimated.

[26] To the extent that observable differences can explain the observed serial persistence in participation decisions, very simple static models, which can be estimated using cross-sectional data, will be sufficient to explain participation decisions. If there exists unobserved heterogeneity, then static models that allow for such unobserved heterogeneity can be estimated with longitudinal data with two-year panel datasets using classical estimation methods.

5.1. *Linear Probability Models*

In order to evaluate the identification of the model in the absence of functional form restrictions, various dynamic linear probability specifications are first considered. These specifications are estimated both in first-differences and in levels.[27] A summary of the results, estimated using four years of data, is presented in Table II. The first row presents GLS estimates of the lagged dependent variable coefficient for the model in first-differences and levels: the estimates are -0.35 and 0.67, respectively. The former estimate will be biased downwards due to negative correlation between Δh_{it-1} and the error due to first

TABLE II

LINEAR PROBABILITY MODELS OF MARRIED WOMEN'S PARTICIPATION

	First Difference Specification				Levels Specification			
	Instruments	γ	ρ	Test Statistic	Instruments	γ	ρ	Test Statistic
(1)	—	-0.348	—	—	—	0.668	—	—
		(.02)				(.01)		
(2)	$\Delta X_{is}, \forall s$	-0.181	—	1.50(a)	$X_{is}, \forall s$	0.483	—	1.23(a)
		(.06)		(.06)		(.05)		(.25)
(3)	$\Delta X_{is}, \forall s$	0.257	—	14.34(a)	$X_{is}, \forall s$	0.337	—	13.41(a)
	h_{it-2}	(.03)		(.00)	Δh_{it-1}	(.02)		(.00)
(4)	h_{it-2}	0.274	—	—	Δh_{it-1}	0.306	—	—
		(.03)				(.03)		
(5)	$h_{it-s}, \forall s > 1$	0.338	—	26.96(b)	$\Delta h_{it-s}, \forall s > 0$	0.399	—	29.36(b)
		(.03)		(.00)		(.03)		(.00)
(6)	h_{it-2}	0.647	-0.194	10.73(c)	Δh_{it-1}	0.563	-0.166	11.15(c)
		(.09)	(.04)	(.06)	Δh_{it-2}	(.13)	(.10)	(.05)

Notes: All specifications include unrestricted time effects, a quadratic in age, race, years of education, permanent and transitory nonlabor income, current realizations of the number of children aged 0–2, 3–5, and 6–17, lagged realizations of the number of children aged 0–2, and a dummy variable for a birth next year. Arbitrary cross-equation correlation and cross-sectional heteroscedasticity-corrected estimated standard errors are in parentheses, except p-values for test statistics. The model is:

$$h_{it} = \gamma h_{it-1} + X_{it}'\beta + \alpha_i + \varepsilon_{it}.$$

Specifications in rows (1)–(5) assume ε_{it} is serially uncorrelated; specifications in row (6) assume $\varepsilon_{it} = \rho\varepsilon_{it-1} + v_{it}$. The estimates in row (6) are based on 2-step minimum distance estimation, using unrestricted first stage coefficient estimates.

(a) First-stage F statistic for the explanatory power of the instruments, conditional on the included exogenous variables; averaged over the period equations.

(b) Sargan over-identification statistic, with 3 degrees of freedom.

(c) Second-stage goodness-of-fit statistic, with 5 degrees of freedom.

[27] In fact, the levels-specification allows for an arbitrary error correlation structure. Also, correlated random effects models corresponding to the specifications in rows (1) and (4) were estimated, in which out-of-period covariates may affect current period participation. The null hypothesis of no correlation between the unobserved heterogeneity and the fertility outcomes and/or husband's earnings was easily accepted in both cases.

differencing, while the latter estimate will be biased upwards in the presence of unobserved heterogeneity. Row (2) contains results using out-of-period realizations of the covariates as instruments for the lagged dependent variable. Assuming that the regressors are exogenous with respect to the transitory error component, these will be valid instruments and enable consistent estimates of the lagged dependent variable effect. The lagged dependent variable coefficients are closer to zero than the corresponding GLS estimates in row 1, but they differ substantially from each other: -0.18 for the first-differences specification and 0.48 for the levels specification. Also, the F statistics for the power of the instruments from the first-stage regressions indicate that these are weak instruments, and thus biased towards the least-squares estimates (e.g., Bound, Jaeger, and Baker (1995)).

If there is no serial correlation in the transitory errors, lagged values of h are valid instruments for Δh_{it-1} and lagged values of Δh are valid instruments for h_{it-1}. In the specifications in row (3) of Table II, h_{it-2} is added to the vector of instruments for Δh_{it-1}, and Δh_{it-1} to the vector of instruments for h_{it-1}. The first-differences and levels estimates of the lagged dependent variable coefficients in this specification are 0.26 and 0.34, and the F statistics from the first-stage regressions indicate these instruments have substantial explanatory power. In row (4), when the regressors are dropped from the instrument sets, the estimated lagged dependent variable coefficients converge further, to 0.27 and 0.31. Following Arellano and Bond (1991), the specifications in row (5) include all valid lagged participation effects in the instrument sets (for example, in the first-differences specification, $h_{i0},\ldots,h_{it-2}$ are valid instruments for Δh_{it-1}). As well as increasing the efficiency of the estimates, this enables the specification of the dynamic model to be tested. The estimated coefficients from this specification, presented in row (5), are each significantly higher than those in row (4); also, the over-identification test results reject the first-order state dependence specification.

The final specification presented in Table II relaxes the assumption that the transitory errors are uncorrelated, and allows the errors to follow a stationary AR(1) process. The results, presented in row (5), for the first-difference and levels specifications are quite similar: the lagged dependent variable coefficients increase dramatically to 0.65 and 0.56 respectively, and are not significantly different from each other, while the estimates of the AR(1) serial correlation parameter are both negative and approximately -0.2. The goodness-of-fit statistics of the second-stage for these specifications (10.7 and 11.2 respectively) suggest borderline acceptance of the specification at the 5% level.

Table III presents the estimated regressor coefficients from the specifications presented in rows 3–5 of Table II. The estimates are broadly similar across the specifications. The levels specification estimates find that permanent nonlabor income has a significantly stronger negative effect than transitory income on participation, suggesting significant expectations and/or tastes effects associated

TABLE III

LINEAR PROBABILITY MODELS OF MARRIED WOMEN'S PARTICIPATION

	First-Differences			Levels		
	(1)	(2)	(3)	(4)	(5)	(6)
y_{mp}	—	—	—	−0.076	−0.068	−0.064
				(.01)	(.01)	(.01)
y_{mt}	−0.034	−0.026	−0.030	−0.021	−0.024	−0.023
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
#Kids0–2$_{t-1}$	−0.044	−0.050	−0.028	−0.048	−0.045	−0.034
	(.01)	(.01)	(.02)	(.01)	(.01)	(.02)
#Kids0–2$_t$	−0.034	−0.034	−0.047	−0.077	−0.070	−0.055
	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)
#Kids3–5$_t$	−0.031	−0.040	−0.024	−0.062	−0.053	−0.022
	(.02)	(.02)	(.02)	(.01)	(.01)	(.01)
#Kids6–17$_t$	−0.010	−0.008	−0.027	−0.010	0.010	−0.005
	(.01)	(.01)	(.01)	(.01)	(.01)	(.004)
Birth$_{t+1}$	0.038	0.045	0.030	0.004	0.010	0.003
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)
h_{t-1}	0.274	0.338	0.647	0.306	0.399	0.563
	(.03)	(.03)	(.09)	(.03)	(.03)	(.13)
ρ	—	—	−0.194	—	—	−0.166
			(.04)			(.10)
Instruments	h_{it-2}	h_{it-s} $\forall s>1$	h_{it-2}	Δh_{it-1}	Δh_{it-s} $\forall s>0$	Δh_{it-1} Δh_{it-2}

Notes: All specifications also include unrestricted time effects, a quadratic in age, race, and years of education. Arbitrary cross-equation correlation and cross-sectional heteroscedasticity-corrected estimated standard errors are in parentheses. The model is:

$$h_{it} = \gamma h_{it-1} + X_{it}'\beta + \alpha_i + \varepsilon_{it}.$$

Columns (1), (2), (4), and (5) assume ε_{it} is serially uncorrelated.
Columns (3) and (6) assume $\varepsilon_{it} = \rho\varepsilon_{it-1} + v_{it}$.

with permanent income.[28] The transitory income coefficient implies a participation elasticity of between -0.03 and -0.04: a transitory 10 percent increase in income will reduce the contemporaneous participation by between 0.2 and 0.3 percentage points. This is consistent with earlier results in the literature—for example, Mincer (1962) and Shaw (1992). The results also show that pre-school children have substantially stronger effects on participation outcomes than school-aged children. The effect of a future birth is estimated to have a positive

[28] Measurement error in reported earnings will likely cause the estimated income effects to be downward biased, and such bias is likely to be larger in the case of transitory earnings. For example, if only transitory (not permanent) earnings is measured with error, with a reliability of 0.8 (as estimated by Bound and Krueger (1991), assuming classical measurement error), and 50 percent of the variance in log male earnings is permanent (Hyslop (1994), using a similar sample to that used here), then the transitory income effect will be downward biased by 40 percent. However, this is likely to be an upper bound on the attenuation bias, due to mean reversion in the measurement error in earnings: taking account of mean reversion, Bound and Krueger estimate the reliability of male earnings to be at least 0.95, which implies the attenuation bias here will be less than 10 percent.

and significant effect on current participation in first-differences, while there is no significant effect in levels.

5.2. *Static Probit Models*

Table IV contains the results for models that focus on demographic characteristics of the women and ignore possible dynamic effects of their past employment outcomes on their current employment decisions. These models do not include the future birth variable, and are estimated over the full seven year period, 1979–85. For each model, White's (1982) quasi-maximum likelihood estimated asymptotic standard errors of the estimates are presented in parentheses below the estimates.

In column (1), the results from a simple probit model are presented. Qualitatively, the results are similar to the linear probability models. Each of the

TABLE IV

STATIC PROBIT MODELS OF MARRIED WOMEN'S PARTICIPATION OUTCOMES

	Simple Probit (1)	Random Effects (2)	Random Effects (3)	Correlated Random Effects (4)
y_{mp}	−0.325	−0.283	−0.314	−0.341
	(.05)	(.06)	(.05)	(.05)
y_{mt}	−0.111	−0.101	−0.106	−0.099
	(.03)	(.03)	(.03)	(.03)
#Kids0–2_{t-1}	−0.070	−0.059	−0.057	−0.058
	(.03)	(.02)	(.02)	(.02)
#Kids0–2_t	−0.434	−0.333	−0.354	−0.300
	(.04)	(.03)	(.03)	(.03)
#Kids3–5_t	−0.371	−0.281	−0.293	−0.247
	(.04)	(.03)	(.03)	(.03)
#Kids6–17_t	−0.079	−0.096	−0.097	−0.084
	(.02)	(.02)	(.02)	(.03)
Var(η_i)[(a)]	—	0.784	0.759	0.804
		(.01)	(.01)	(.02)
Log-likelihood	−7130.52	−4913.16	4916.05[(b)]	−4888.38
Goodness-of-fit[(c)]	10725.10	631.04	—	—
Wald Statistics for H_0: $CRE = 0$				
y_{mt}	—	—	—	48.50 (.00)
#Kids0–2	—	—	—	32.36 (.00)
#Kids3–5	—	—	—	12.77 (.12)
#Kids6–17	—	—	—	21.74 (.01)

Notes: All specifications also include unrestricted time effects, a quadratic in age, race, and years of education. Estimated (QML) standard errors in parentheses, except p-values for Wald Statistics. Number of observations: $N = 1812$ individuals, observed over $T = 7$ years. All models include time dummies; simple probit model assumes iid errors across i and t. Variance normalizations: Var(u_{it}) = 1. The models in columns (2) and (4) are estimated by MLE using Gaussian quadrature with 20 quadrature points; the model in column (3) is estimated by MSL using $R = 20$ simulation replications per observation. The CRE model expresses α_i as a linear function of y_{mt}, #Kids0–2, #Kids3–5, and #Kids6–17.

[(a)] Var(η_i) is expressed as a fraction of the total error variance; Var(α_i) in the RE models.

[(b)] The MSL objective value = −5023.47.

[(c)] Pearson goodness-of-fit statistics computed from 48 groupings of actual and predicted employment sequences in Table VI (see text for details).

fertility variables has a significantly negative effect on women's participation decisions, and younger children have stronger effects than older. An additional child aged less than 2 reduces the probability of participation by 17 percent, while each child aged 3–5 and 6–17 reduces the probability of participation by 12 percent and 2 percent respectively. Also, the participation effect of permanent nonlabor income is substantially greater than that of transitory income.

A Pearson goodness-of-fit (GOF) statistic for the simple probit model is presented in the second-to-last row of column (1). This statistic is computed by comparing the predicted and actual frequencies of 48 cell groupings of the participation sequences in the 7-year dataset (conditional on the explanatory variables for each individual).[29] The goodness-of-fit statistic for the model is 10725, indicating the intertemporal explanatory power of this model is extremely low. In fact, the predicted participation sequence frequencies from this simple cross-sectional probit model are strikingly close to those predicted by a simple binomial model with fixed probability of participation equal to 0.7 (see Table VI).

Column (2) contains the results from a probit model which allows for individual-specific random effects. This model is estimated by MLE using a Gaussian quadrature procedure with 20 quadrature points. This specification finds that 78 percent of the latent error variance is due to unobserved heterogeneity. Accounting for this unobserved heterogeneity, the estimated effects of young children decline by about 25 percent, while that of children aged 6–17 increases by 25 percent. Allowing for unobserved heterogeneity results in a dramatic improvement in the fit of the model, as measured by either the log-likelihood or the goodness-of-fit statistic. Nonetheless, the goodness-of-fit

[29] The Pearson goodness-of-fit statistics for the models presented are computed as

$$GOF = \sum_{s=1}^{S} \frac{(n_s - \hat{n}_s)^2}{\hat{n}_s}$$

where n_s and $\hat{n}_s$ are, respectively, the observed and predicted frequency of the sth cell, and S is the number of cells. There are a priori reasons to expect many of the individual sequences to have very small cell sizes, adversely affecting the finite sample approximation to the asymptotic distribution of the goodness-of-fit statistics computed from all 128 cells. For example, Table I and the patterns of participation choices presented in Heckman and Willis (1977) suggest that sequences involving several transitions will occur with very low probability. For this reason, each sequence with no more than two transitions is treated separately, while sequences with more than two transitions are grouped by the number of years of participation. For example, the sequence 1100111 forms a cell, while sequences with 5 years of participation and more than 2 transitions, such as 1010111 and 0111011 etc., are grouped together. Grouping sequences in this way reduces the number of cells over which goodness-of-fit is computed from 128 to 48. The goodness-of-fit statistics are intended as an informal summary of the fit of the model, not as a formal diagnostic, and have not been corrected for the estimation of the (k) parameters in the model. Suitable corrections to the statistics are outlined in Heckman (1984) and Moore (1977). Moore (1977) also discusses a result, due to Chernoff and Lehmann, that the distribution in this case has critical values that fall between those of the $\chi^2_{(S-k-1)}$ and $\chi^2_{(S-1)}$ distributions.

statistic indicates this model still provides a poor fit to the observed participation sequences ($GOF = 631$).

For purposes of comparison, the random effects model is re-estimated using MSL with 20 simulation replications, and the results are presented in column (3) of Table IV.[30] The estimates are generally similar to those in column (2). Although the maximized log-simulated-likelihood evaluated using 20 replications (-5023.47) is substantially lower than the true log-likelihood value evaluated at the MLE estimates in column (2) (-4913.16), the log-likelihood evaluated at the MSL estimates (-4916.05) implies that MSL estimation is reasonably accurate.

The probit specifications so far have assumed that the fertility variables are exogenous with respect to the participation decisions of married women, in the sense that contemporaneous outcomes are sufficient to capture the effects of fertility on participation choice, and also that noncontemporaneous nonlabor income effects are adequately summarized by sample average nonlabor income. If these assumptions are incorrect, the resulting coefficient estimates will be biased and inconsistent. We now consider the correlated random effects (CRE) specification for α_i, given in equation (6). The MLE results of this model are presented in column (4) of Table IV. The results provide evidence against the null hypothesis that the random effect is uncorrelated with the fertility and earnings variables. The likelihood ratio test of the simple versus correlated random effects model rejects the simple model (LR statistic $= 54.5$, 30 degrees of freedom, p-value < 0.01). In addition, separate Wald-statistics for the correlation between the unobserved heterogeneity and the three fertility variables and income rejects the hypothesis of no correlation in each case (p-values < 0.07). Allowing for correlation between the regressors and the unobserved heterogeneity reduces the estimated direct-effects of the contemporaneous fertility variables by approximately 10 percent. However, although there is evidence of a significant effect of non-contemporaneous income on current participation, the effects of transitory nonlabor income are largely unchanged from column (2).

5.3. *Dynamic Probit Models*

I now turn to the dynamic probit specifications of the intertemporal participation model. The results are presented in Table V. Columns (1) and (2) contain the results for the uncorrelated and correlated random effects specifications respectively, plus a stationary first-order autoregressive (AR(1)) error component; columns (3) and (4) contain the results for the corresponding specifications which also include first-order state dependence (SD(1)). Consider first the

[30] The results of a preliminary analysis of MSL estimation using three periods of data suggest that the simulation bias in the MSL estimates is comparatively small when $R \geq 10$ (Hyslop (1995), Appendix Table A1). However, as the number of replications required to accurately simulate the log likelihood function with little bias increases with the length of panel, $R = 20$ replications are used.

TABLE V

DYNAMIC PROBIT MODELS OF MARRIED WOMEN'S PARTICIPATION OUTCOMES

	RE + AR(1) (1)	CRE + AR(1) (2)	RE, AR(1) + SD(1) (3)	CRE, AR(1) + SD(1) (4)
y_{mp}	−0.316	−0.332	−0.272	−0.285
	(.05)	(.05)	(.05)	(.05)
y_{mt}	−0.097	−0.097	−0.140	−0.140
	(.03)	(.03)	(.04)	(.04)
$\#\text{Kids0–2}_{t-1}$	−0.055	−0.052	−0.129	−0.115
	(.02)	(.02)	(.04)	(.04)
$\#\text{Kids0–2}_t$	−0.311	−0.272	−0.296	−0.252
	(.03)	(.03)	(.04)	(.05)
$\#\text{Kids3–5}_t$	−0.270	−0.234	−0.174	−0.135
	(.03)	(.03)	(.04)	(.05)
$\#\text{Kids6–17}_t$	−0.089	−0.077	−0.048	−0.054
	(.02)	(.02)	(.02)	(.04)
h_{t-1}	—	—	1.063	1.042
			(.09)	(.09)
		Covariance Parameters		
$\text{Var}(\eta_i)$	0.559	0.546	0.479	0.485
	(.04)	(.04)	(.04)	(.04)
AR(1)Coeff, ρ	0.687	0.696	−0.219	−0.213
	(.03)	(.04)	(.04)	(.04)
$\text{Corr}(u_{i79}, u_{it})$	—	—	0.482	0.494
			(.03)	(.03)
Log-likelihood[a]	−4681.54	−4663.71	−4655.36	−4643.52
Goodness-of-fit[b]	57.62	—	62.57	—
Wald statistics				
H_0: $\text{Corr}(u_{i0}, u_{it}) = \rho_0$	—	—	8.80 (.12)	—
H_0: CRE y_{mt}	—	8.22 (.22)	—	2.92 (.82)
H_0: CRE #Kids0–2	—	9.65 (.29)	—	3.39 (.91)
H_0: CRE #Kids3–5	—	9.37 (.31)	—	3.84 (.87)
H_0: CRE #Kids6–17	—	8.04 (.43)	—	3.34 (.91)

Note: All specifications also include unrestricted time effects, a quadratic in age, race, and years of education. Estimated (QML) standard errors are in parentheses, except p-values for Wald-statistics. Number of observations: $N = 1812$ individuals, observed over $T = 7$ years. All models are estimated by MSL using $R = 20$ simulation replications. Variance normalizations: $\text{Var}(u_{i0}) = 1$ and $\text{Var}(u_{it}) = 1$. The CRE models express α_i as a linear function of y_{mt}, #Kids0–2, #Kids3–5, and #Kids6–17.

[a] Simulated log-likelihood value using 10,000 replications.

[b] Pearson goodness-of-fit statistics computed from groupings of employment sequences in Table VI (see text for details).

results in column (1) for the specification which includes an AR(1) error component. In comparison to the results for the static random effects models in Table IV, these results show the addition of a transitory component of error has a significant effect on the model. The estimated AR(1) coefficient is 0.69 and the variance of the random effect is now estimated to be 56 percent of the total error variance, and both of these parameters are precisely estimated. The covariate effects are each individually significant and are generally close to those in the static random effects model. There is a significant increase in the maximum log-simulated-likelihood value relative to the RE model. Also the goodness-of-fit statistic associated with the predicted sequences from this model (57.6) suggests the model provides a reasonable statistical fit to the actual participation outcomes (e.g., using the Chemoff-Lehmann rule of thumb, this Pearson statistic has a p-value of approximately 0.1).

Column (2) of Table V presents results for the model specification with correlated random effects. The hypothesis that the fertility and earnings variables are exogenous to the participation decisions of married women (i.e. the uncorrelated random effects specification is valid) is accepted both in aggregate using a likelihood ratio statistic comparing the models in columns (1) and (2), and for each variable separately using Wald statistics. However, the estimated effects of children on the participation decision are reduced by 10–15 percent after controlling for correlation with the unobserved heterogeneity.

Next, consider the specifications with first order state dependence. Allowing for expectations effects on current participation, the results find a small and statistically insignificant negative effect of a birth next year (t value $= 0.3$). For this reason, and to enable the model to be estimated for participation over the full seven-year period, the specification presented in column (3) does *not* include the future birth dummy variable. This specification models the initial conditions of the process using a reduced form probit specification for the first period participation outcome in terms of the initial period covariates, and restricts the correlations between the errors in this equation and the other periods, u_{i0} and u_{it}, to be equal.[31] The results show a large first-order state dependence effect: the coefficient estimate (1.06) converts to an average probability derivative of 0.37. Including state dependence also has dramatic effects on the estimated unobserved heterogeneity and serial correlation parameters: the fraction of error variance due to the random effect falls to 0.48, and the estimated AR(1) coefficient is -0.22. Although these results imply there is substantial structural state dependence in participation decisions, the overall fit of the model does not change dramatically. For example, there does not appear to be any gains in the model's ability to predict the observed participation sequences. In fact, the Pearson goodness-of-fit statistic for the model in column (3) implies it provides a slightly worse fit than the model in column (1).

Column (4) of Table V contains the results for the model when the random effects specification is relaxed to allow for correlation with the fertility and nonlabor income variables. The results for this specification again provide no evidence against the exogeneity hypothesis, and the estimates are very close to

[31] The Wald-statistic for the restrictions is 8.8 (5 degrees of freedom, p-value $= 0.12$), suggesting restrictions are reasonable. Also, the unrestricted correlations lie in the range [0.38, 0.51], compared to the restricted estimate of 0.48, and the results from the restricted and unrestricted specifications are almost identical.

those in column (3) assuming fertility and earnings are exogenous with respect to this component of tastes.

Finally, one problem with using MSL estimation is that it is inconsistent for a finite number of simulation replications. In order to assess the extent of estimation bias in the MSL estimators here, I performed a series of Monte Carlo simulation experiments. The experimental framework and results are described in Appendix 2. These experiments found that when the true data generating process (DGP) contains either positive state dependence or positive serial correlation, the MSL estimate of the state dependence is strongly biased upward, while there is downward bias in the estimated serial correlation and unobserved heterogeneity parameters. However, across a range of other DGPs, there is little evidence of bias in the MSL estimates. These results give support to the reliability of the MSL estimates discussed in this section.

The analysis provides the following conclusions. First, both the linear and nonlinear models find that intertemporal participation decisions are characterized by a substantial amount of positive state dependence and unobserved permanent heterogeneity, and a negatively correlated transitory error component. Second, the dynamics associated with state dependence have important implications for the relationship between participation and fertility decisions. The probit results show that, after controlling for serially correlated errors or state dependence, there is no evidence that fertility decisions are correlated with unobserved tastes for work. Third, the results across the alternative specifications are not consistent concerning the importance of fertility expectations on participation decisions. There is no evidence of an effect in the linear probability levels and probit specifications, while the linear probability first-differences specifications find the existence of a theoretically perverse positive effect of a future birth on current participation.

5.4. *Predicted Participation Sequences*

In order to provide a descriptive assessment of the fit of the models estimated in this paper, Table VI compares the frequencies of the participation sequences predicted by various models estimated. Table VI presents the actual frequencies for each of the 128 employment sequences, together with the predicted frequencies from a dynamic linear probability model estimated in first-differences,[32] the stacked cross-sectional probit model (specification in Table IV, column (1)), the random effects probit model (Table IV, column (2)), the probit model with AR(1) errors (Table V, column (1)), and the probit model with state dependence (Table V, column (3)). The sequences are ordered in groups according to the number of participation outcomes in the sequence.

[32] For greater comparability with the probit models, the linear probability specification corresponds to that presented in Table II, column 4, except that the variable for future births is omitted and the model is estimated for 5 years of participation. Similar predictions result using alternative specifications.

TABLE VIa

COMPARISON OF PREDICTED WITH OBSERVED PARTICIPATION SEQUENCE FREQUENCIES

		Predicted Frequency				
Sequence	Observed Frequency (1)	Linear Probability (2)	Simple Probit (3)	RE Probit (4)	AR(1) Probit (5)	SD Probit (6)
0000000	192	316.24	4.62	202.05	186.76	187.65
0000001	27	26.14	4.01	21.17	27.07	23.85
0000010	12	26.12	3.91	21.51	15.44	14.46
0000100	5	13.66	3.15	13.68	8.78	9.22
0001000	9	5.89	2.94	11.07	7.40	8.16
0010000	10	3.74	3.01	12.20	8.29	9.12
0100000	11	0.00	3.25	14.04	9.57	8.47
1000000	36	120.96	4.00	19.84	28.96	36.92
	110	196.51	24.27	113.51	105.51	110.20
0000011	26	4.05	5.73	9.86	27.30	24.46
0000101	4	2.25	4.41	5.93	3.70	3.14
0000110	5	2.18	4.33	6.17	7.58	6.23
0001001	2	1.01	3.81	4.38	2.30	1.78
0001010	1	0.99	3.78	4.62	1.84	1.45
0001100	4	0.66	3.10	3.22	4.10	3.41
0010001	1	0.53	3.75	4.40	2.20	2.10
0010010	1	0.53	3.72	4.68	1.42	1.06
0010100	0	0.36	3.08	3.28	1.23	1.02
0011000	3	0.21	2.92	2.93	4.03	3.57
0100001	1	0.00	4.02	4.76	2.40	1.88
0100010	1	0.00	3.97	4.99	1.39	1.00
0100100	1	0.00	3.28	3.46	0.95	0.60
0101000	1	0.00	3.11	3.10	1.17	0.86
0110000	6	0.00	3.33	3.72	4.94	3.63
1000001	3	9.75	4.86	6.44	7.75	7.23
1000010	1	9.00	4.77	6.65	4.12	3.93
1000100	3	4.39	3.93	4.58	2.58	2.49
1001000	2	0.96	3.70	4.08	2.52	2.16
1010000	4	0.00	4.00	4.96	4.08	3.64
1100000	28	0.00	4.62	6.49	20.34	27.49
	98	36.87	82.22	102.70	107.94	103.13
0000111	22	0.48	8.46	6.25	24.61	23.88
0001011	5	0.22	6.97	4.31	4.08	3.39
0001101	2	0.15	5.48	2.74	2.01	1.66
0001110	2	0.14	5.45	2.93	5.00	4.69
0010011	3	0.11	6.67	4.03	3.12	2.52
0010101	1	0.07	5.28	2.59	0.60	0.50
0010110	2	0.07	5.28	2.80	1.23	0.87
0011001	1	0.04	4.69	2.09	1.42	1.10
0011010	0	0.04	4.71	2.28	1.10	0.79
0011100	0	0.03	3.97	1.68	3.10	2.88
0100011	4	0.00	6.89	4.11	3.09	2.30
0100101	0	0.00	5.45	2.62	0.48	0.28

TABLE VIb

COMPARISON OF PREDICTED WITH OBSERVED PARTICIPATION SEQUENCE FREQUENCIES

		Predicted Frequency				
Sequence	Observed Frequency (1)	Linear Probability (2)	Simple Probit (3)	RE Probit (4)	AR(1) Probit (5)	SD Probit (6)
0100110	0	0.00	5.44	2.81	0.93	0.51
0101001	0	0.00	4.87	2.13	0.44	0.26
0101010	0	0.00	4.88	2.30	0.32	0.19
0101100	1	0.00	4.11	1.69	0.74	0.43
0110001	0	0.00	5.03	2.35	1.52	1.14
0110010	1	0.00	5.05	2.54	0.94	0.51
0110100	2	0.00	4.28	1.87	0.82	0.48
0111000	3	0.01	4.13	1.82	3.35	2.71
1000011	12	1.28	8.10	5.37	10.07	8.48
1000101	0	0.63	6.40	3.38	1.43	1.07
1000110	2	0.56	6.36	3.59	2.62	1.97
1001001	0	0.12	5.69	2.73	1.04	0.59
1001010	0	0.12	5.68	2.92	0.73	0.45
1001100	0	0.07	4.79	2.15	1.64	1.03
1010001	0	0.00	5.91	3.04	1.42	1.09
1010010	0	0.00	5.91	3.26	0.82	0.50
1010100	2	0.00	5.01	2.40	0.70	0.48
1011000	3	0.00	4.83	2.32	2.32	1.66
1100001	4	0.01	6.67	3.78	6.92	7.40
1100010	5	0.01	6.62	3.95	3.57	3.62
1100100	1	0.01	5.59	2.88	2.45	2.13
1101000	4	0.01	5.42	2.81	3.04	3.07
1110000	21	0.03	6.11	3.78	18.19	19.99
	103	4.21	196.18	104.30	115.86	104.62
0001111	22	0.04	13.13	5.43	22.99	24.76
0010111	1	0.02	12.33	4.81	4.75	4.40
0011011	2	0.01	10.49	3.53	2.73	2.50
0011101	1	0.01	8.48	2.32	1.68	1.93
0011110	6	0.01	8.62	2.58	4.40	4.68
0100111	8	0.00	12.28	4.58	3.71	2.42
0101011	3	0.00	10.53	3.42	0.83	0.55
0101101	0	0.00	8.53	2.25	0.42	0.26
0101110	1	0.00	8.65	2.47	0.97	0.69
0110011	1	0.04	10.52	3.50	2.30	1.58
0110101	1	0.03	8.56	2.32	0.44	0.30
0110110	1	0.04	8.72	2.57	0.86	0.49
0111001	1	0.09	7.85	2.02	1.35	1.07
0111010	0	0.10	8.02	2.25	0.99	0.70
0111100	5	0.16	6.94	1.73	2.99	2.41
1000111	16	0.11	14.00	5.80	11.68	9.08
1001011	1	0.02	11.97	4.28	2.11	1.25
1001101	1	0.01	9.69	2.81	1.02	0.60
1001110	1	0.01	9.79	3.05	2.28	1.61
1010011	2	0.00	12.01	4.43	2.28	1.45
1010101	1	0.00	9.77	2.93	0.42	0.28

TABLE VIc

COMPARISON OF PREDICTED WITH OBSERVED PARTICIPATION SEQUENCE FREQUENCIES

		Predicted Frequency				
Sequence	Observed Frequency (1)	Linear Probability (2)	Simple Probit (3)	RE Probit (4)	AR(1) Probit (5)	SD Probit (6)
1010110	0	0.00	9.92	3.22	0.77	0.46
1011001	1	0.00	8.96	2.54	1.03	0.62
1011010	2	0.00	9.12	2.80	0.72	0.41
1011100	2	0.00	7.90	2.17	2.01	1.49
1100011	11	0.07	13.10	5.25	10.30	10.10
1100101	0	0.05	10.61	3.40	1.55	1.17
1100110	0	0.04	10.72	3.67	2.67	2.00
1101001	0	0.06	9.78	2.95	1.43	1.10
1101010	3	0.05	9.90	3.21	0.93	0.75
1101100	3	0.05	8.57	2.48	2.13	1.66
1110001	4	0.16	10.57	3.67	7.64	7.43
1110010	2	0.16	10.70	3.96	4.09	3.06
1110100	4	0.16	9.31	3.08	3.53	2.87
1111000	14	0.25	9.29	3.30	17.00	15.83
	121	1.75	349.33	114.78	127.00	111.96
0011111	34	0.00	24.80	8.44	26.28	35.86
0101111	6	0.00	24.22	7.62	5.47	4.40
0110111	7	1.02	23.61	7.28	3.79	3.24
0111011	1	1.23	20.92	5.67	2.76	2.88
0111101	3	1.26	17.38	3.81	1.83	2.07
0111110	4	2.80	18.02	4.32	4.70	4.69
1001111	13	0.00	26.79	9.28	14.56	10.00
1010111	5	0.00	26.27	8.99	3.89	2.82
1011011	4	0.00	23.29	6.99	2.24	1.55
1011101	3	0.00	19.38	4.70	1.32	1.19
1011110	0	0.00	20.02	5.29	3.27	2.73
1100111	17	0.65	27.66	9.94	14.41	12.09
1101011	7	0.55	24.62	7.79	3.12	2.73
1101101	0	0.35	20.46	5.19	1.48	1.25
1101110	2	0.39	21.06	5.79	3.14	3.10
1110011	9	1.60	25.63	8.93	13.70	11.70
1110101	5	1.09	21.38	6.00	2.44	2.16
1110110	4	1.27	22.10	6.72	4.20	3.25
1111001	8	1.54	20.37	5.70	9.40	7.59
1111010	8	1.68	21.07	6.38	5.85	4.54
1111100	19	1.94	18.82	5.26	19.46	15.59
	159	17.37	467.87	140.09	147.31	135.43

First, consider the linear probability model's predicted frequencies, presented in column (2). These frequencies are computed conditional on the first two years' participation outcomes. In addition, the predicted probabilities are constrained to lie between 0 and 1 in each year, which tends to improve the predicted frequencies of the sequences. The linear probability model with state

TABLE VId

COMPARISON OF PREDICTED WITH OBSERVED PARTICIPATION SEQUENCE FREQUENCIES

Sequence	Observed Frequency (1)	Predicted Frequency: Linear Probability (2)	Simple Probit (3)	RE Probit (4)	AR(1) Probit (5)	SD Probit (6)
0111111	45	112.21	61.50	27.82	36.69	51.53
1011111	28	0.00	67.14	33.33	27.02	25.83
1101111	14	8.73	68.95	34.66	24.92	27.64
1110111	17	24.70	70.02	36.96	26.73	29.57
1111011	15	21.98	64.69	30.75	23.29	25.01
1111101	9	16.12	55.26	20.99	16.21	17.79
1111110	28	33.78	58.17	24.77	39.32	37.17
	156	217.52	445.73	209.28	194.18	214.54
1111111	873	1021.50	241.78	825.23	827.44	844.47

Notes: Column (2) frequencies predicted by the linear probability model estimated in first-differences (see text for details); column (3) predicted by simple probit model (Table IV, column (1)); column (4) by random effects probit model (Table IV, column (2)); column (5) by the probit model with random effect and AR(1) errors (Table V, column (1)); and column (6) by the probit model with state dependence (Table V, column (3)).

dependence does a poor job of predicting participation probabilities in the unit interval. For example, the fraction of the sample with predicted probabilities greater than 1 ranges from 41 percent in year 3 to 48 percent in year 7, and the fraction with predicted probabilities less than 0 ranges from 18 percent in year 3 to 15 percent in year 7. As a consequence, the linear probability model greatly overpredicts the frequency of the sequences with no change in participation status over the period.

Column (3) contains the predicted frequencies from the simple probit model. Conditional on the number of years worked during the sample, the predicted frequency distribution is almost uniform, while the frequency distribution across the number of years worked is close to that implied by a simple binomial model with fixed probability of participation equal to 0.7. Thus the observed exogenous variables contribute relatively little to explaining the pattern of intertemporal participation behavior of women in this model. The predictions from the random effects probit model are presented in column (4). This model predicts the frequency of the number of years in each participation state adequately. However, it overpredicts the aggregate frequency of 6 years of participation and underpredicts the frequency of 7 years, suggesting that the unobserved heterogeneity may have thicker tails than the normal distribution and/or be skewed to the left. Also, conditional on the number of years worked, the random effects model overpredicts sequences with transitions and underpredicts continuous sequences of 0's and 1's. This is expected if there are either transitory unobserved factors or state dependence that affect participation, as this model only allows for permanent person-specific differences.

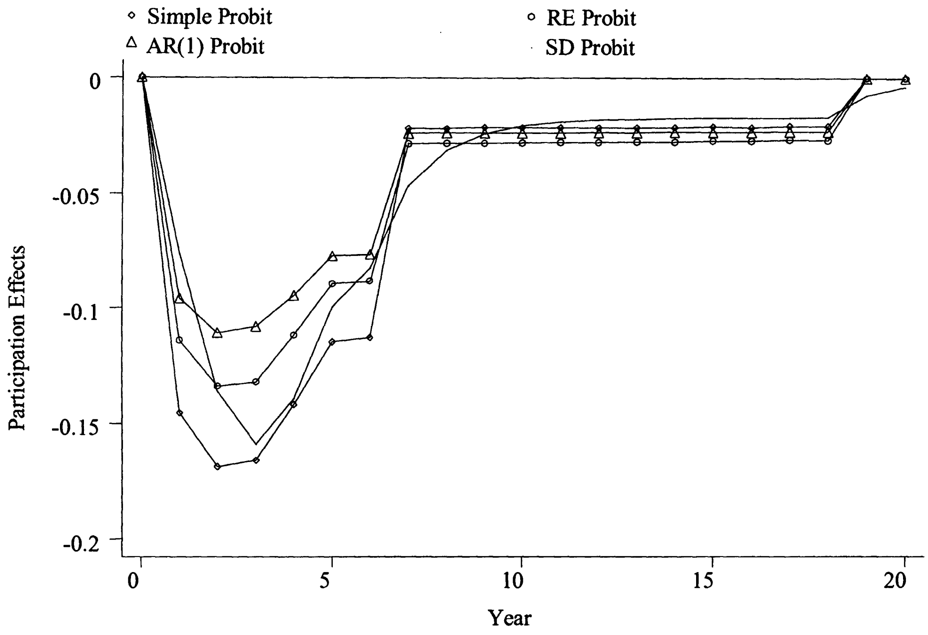
The final two columns in Table VI contain the predicted frequencies from the probit models that include an AR(1) error component (column (5)) and also state dependence (column (6)). As with the random effects probit model, these models predict the distribution of the number of years worked reasonably well, although not obviously better than in column (4). More importantly, conditional on the number of years worked, the predicted frequency distributions of sequences also closely matches the observed frequencies. In particular, the higher frequencies of sequences with few transitions are well identified by each of these models. The predictions from these models are very similar and fit the observed frequencies quite well, as is suggested by their respective goodness-of-fit statistics. Thus, although the effect of first-order state dependence is estimated to be strong, the additional predictive power from the model with state dependence relative to the model with a random effect plus AR(1) error components structure appears to be relatively small.

5.5. *Simulated Responses to Fertility and Nonlabor Income Changes*

Finally, to illustrate the participation responses to fertility and income changes, we present the following simulations using a linear probability specification and alternative probit models. We simulate the intertemporal participation response over 20 years to two events: a birth and a 10 percent increase in permanent nonlabor income. In each case, the base sample has characteristics in year 0 equal to the average over the sample period, and the event occurs in year 1.[33]

The simulated responses to a birth in year 1 are presented in Figure 1. The first panel in Figure 1 presents simulations based on the full sample using the simple probit model, the random effects probit model, the probit model with AR(1) errors, and the probit model with state dependence. The differences in the simulated responses are quite noticeable when the child is young—for example, the peak response to an additional child aged 0–2 ranges from about -0.11 using the probit model with AR(1) errors to -0.17 using the simple probit model. However, these differences decline as the child ages, and are nearly indistinguishable once it has reached school age—the effect of an additional child aged 3–5 is between 8 and 12 percentage points; and once the child has reached school age, the participation effect is between 2 and 3 percentage points. The principal difference between the model with state dependence and the other models is that the simulated dynamics are smoother over time due to the lagged dependent variable effect. In the second panel of Figure 1, I compare the response predicted by the probit model with state dependence to that predicted by the linear probability model with state dependence. The average responses are remarkably similar from these two models.

[33] Transitory income is assumed to be zero in each year. The sample average number of children in each age range is rounded to the nearest integer, and the age of each child within this range is randomly (and independently within families) allocated from a Uniform distribution. There is no discernible difference to the results if the initial year numbers of children are used instead of the sample averages.



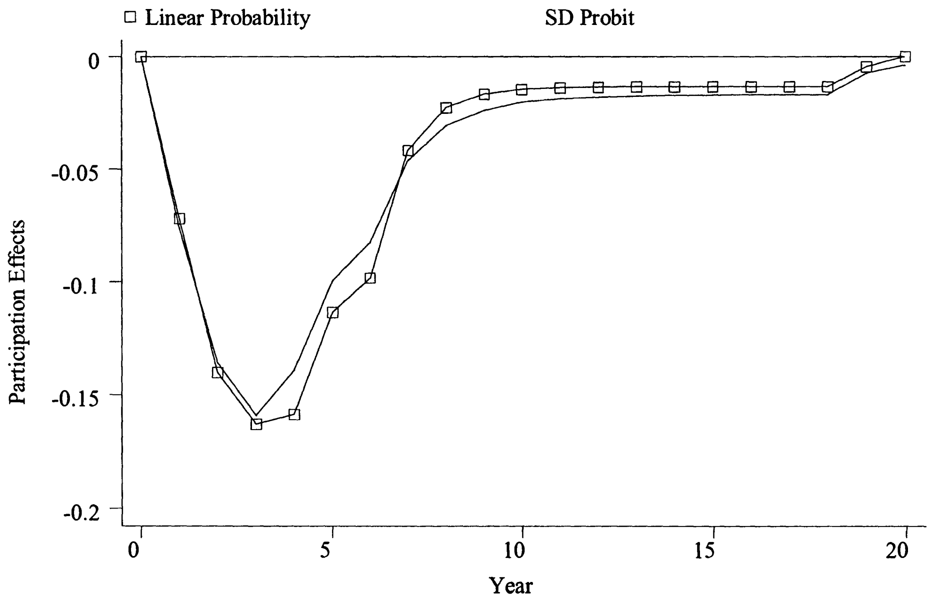


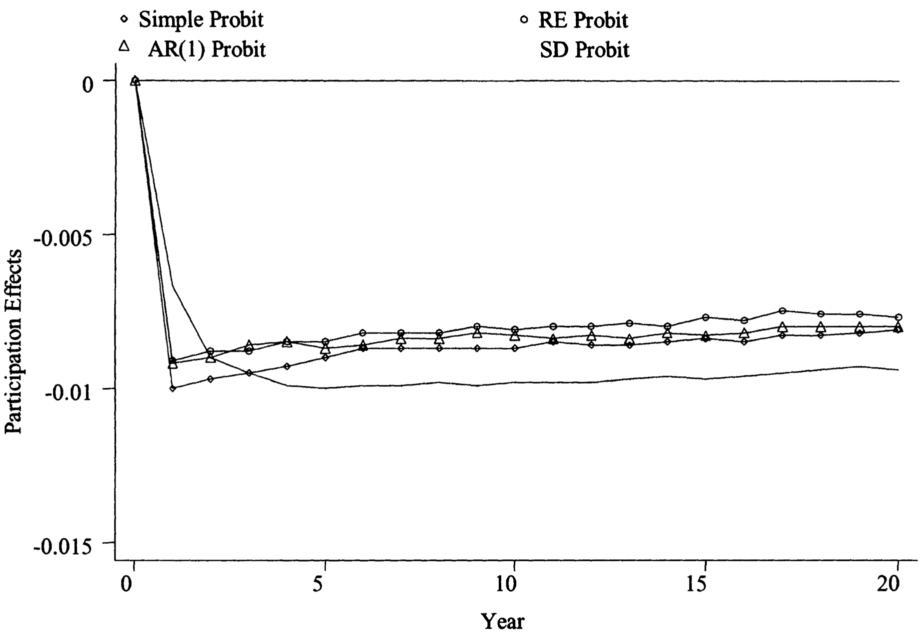
FIGURE 1.—Response to a birth in year 1.

The simulated effects of a permanent increase in nonlabor income on participation are presented in Figure 2. As in Figure 1, the first panel presents the simulations from the four probit models, while the second panel compares the linear probability and probit models with state dependence. The implications from each of these models are broadly similar: a 10 percent permanent increase in nonlabor income is predicted to reduce women's participation by between 0.8 and 1 percentage points on average. There appears to be a somewhat stronger effect in the models with state dependence although the presence of the lagged dependent variable in these models means that the full effect of the change takes 5 years to be realized. Also, the slight decrease in the simulated effects over time of the probit models reflects the effects of children aging over time in the base sample.

Although the linear probability and probit models give quite similar results in terms of sample average predictions, the probit specification enables differential responses across the distribution of characteristics, albeit in a very restrictive fashion. To illustrate one such differential response, Figure 3 presents simulation responses to a birth and an increase in permanent income for "low" educated (at most high school education) and "high" educated (more than high school education) women using the probit model with state dependence. The average-responses to a birth and to an increase in nonlabor income are compared in Figures 3(a) and 3(b) respectively. The responses of low-educated women are slightly greater than those of high-educated women in each case. For example, following a birth, low educated women are approximately 1 percent less likely to participate in the presence of a child aged 0–2 years than are high educated women. The differential response is relatively greater to a permanent increase in nonlabor income: the model predicts low educated women would reduce their participation by about 0.11 percentage points in response to a 10 percent increase in nonlabor income, compared to a reduction of 9 percentage points by high educated women.

6. CONCLUSION

This paper has used dynamic specifications of the employment participation decisions of married women to investigate the exogeneity of fertility variables to the participation decision, and the participation response to their nonlabor income. The empirical framework allows for three components that generate serial persistence in participation decisions: a permanent individual effect to control for unobserved heterogeneity, a serially correlated transitory error component, and a state dependence component to control for the effects of previous participation outcomes on the current participation decision. Both linear probability and probit specifications show that each of these components is statistically significant in characterizing the dynamics of women's participation decisions. In particular, there is a strong estimated effect of state dependence and unobserved heterogeneity. One peculiarity in the empirical results is that the transitory error component is found to be negatively serially correlated,



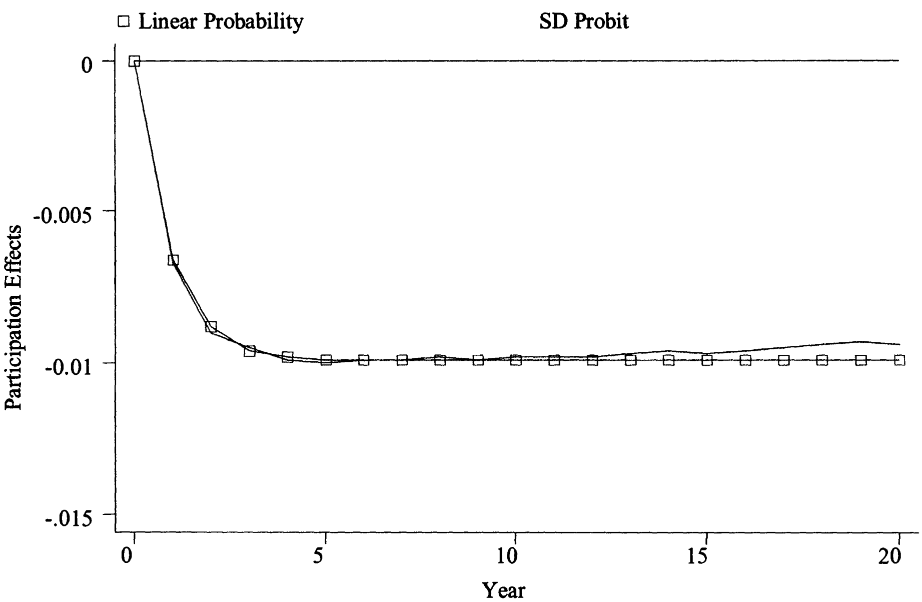


FIGURE 2.—Response to a 10 percent increase in permanent income.

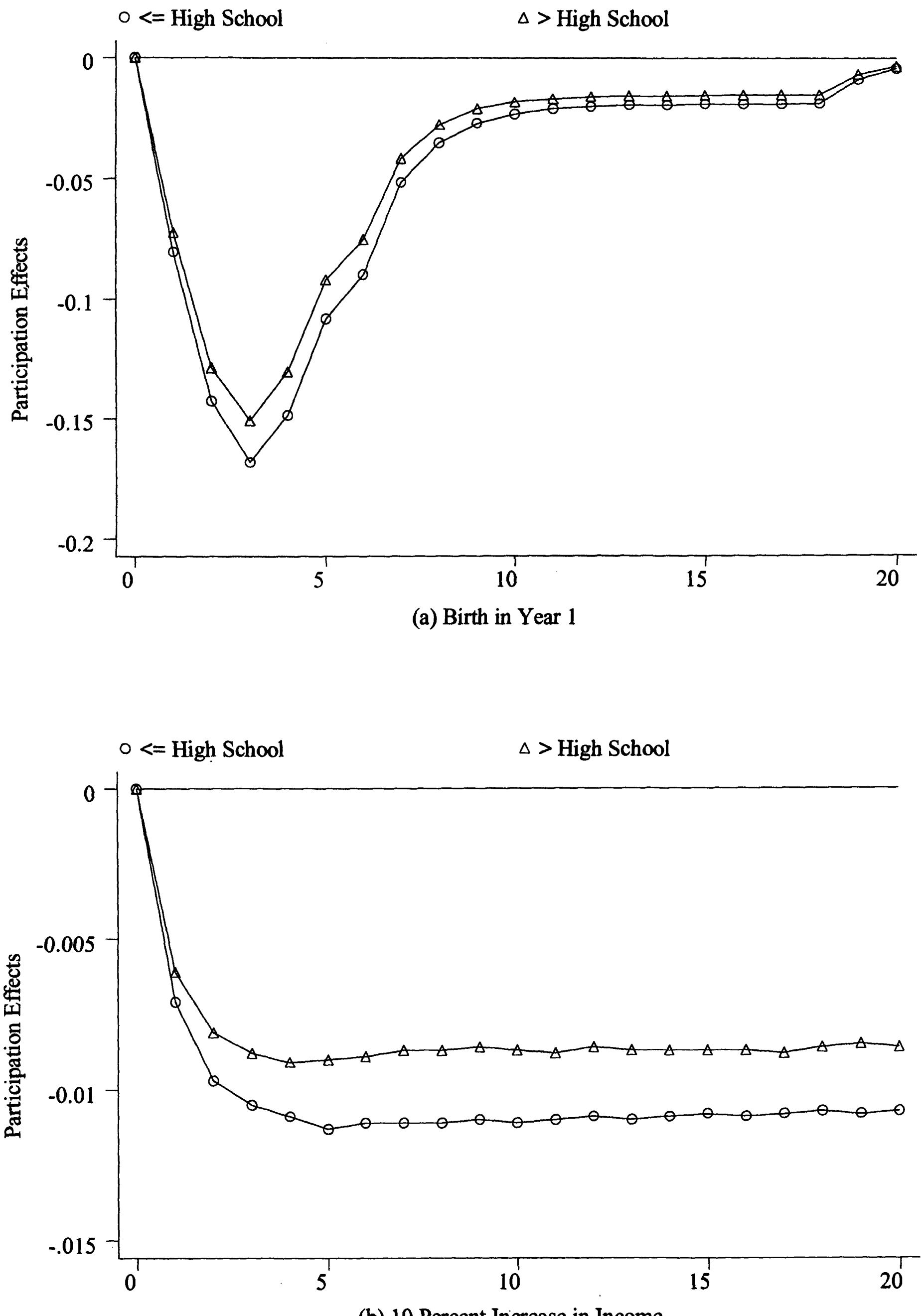


FIGURE 3.—Participation responses by education level.

although the magnitude of the correlation is relatively small. A suitable interpretation for this is not obvious. One possibility, beyond the scope of this paper, is that the form of state dependence is misspecified, and the AR(1) error component is acting as a fitting parameter in the model. For example, individuals' human capital, which affects their wage offers and depends on their past participation decisions, will imply a more general form of state dependence.

Substantively, the analysis finds that fertility is correlated with women's unobserved tastes for work, and is not exogenous with respect to their participation decisions, if the dynamic structure of participation decisions is ignored. However, when the dynamics are modelled, there is little evidence against the exogeneity assumption. Second, the effect of permanent nonlabor income on participation decisions is significantly stronger than that of transitory income, implying a small direct income effect, and a significant correlation between permanent income and tastes for work and/or an expectations effect of future income. The participation elasticities with respect to permanent and transitory income are approximately -0.2 and -0.04 respectively.

The predicted participation response to either a birth or an increase in nonlabor income are remarkably similar across the range of linear probability and probit models estimated. In contrast, the predicted participation sequences vary markedly across the various specifications. In particular, the dynamic linear probability and static probit models predict the observed patterns extremely poorly, while the probit models allowing for dynamic effects predict the observed patterns adequately.

Dept. of Economics, UCLA, 405 Hilgard Ave., Los Angeles, CA 90095-1477, U.S.A.

Manuscript received March, 1996; final revision received January, 1999.

APPENDIX 1: THE SMOOTH RECURSIVE CONDITIONING (SRC) SIMULATOR FOR MSL

Consider the binary dependent variable model defined by

$$y_{it} = 1(y_{it}^* > 0), \quad \text{and}$$

$$y_{it}^* = \mu_{it} + u_{it} \qquad (i = 1,\ldots,N;\ t = 1,\ldots,T),$$

where $\mu_{it} = X_{it}'\beta$ and $u_i = (u_{i1},\ldots,u_{iT})' \sim N(0,\Omega)$. The latent model can be expressed in vector form:

$$y_i^* = \mu_i + \Gamma v_i \qquad (i = 1,\ldots,N),$$

where $u_i = \Gamma v_i$; $v_i \sim N(0,I)$; and Γ is the lower-diagonal Cholesky decomposition of Ω—i.e., $\Gamma\Gamma' = \Omega$.

Let $D(\theta, y_i) = \{y_i^*\colon 1(y_{it}^* > 0) = y_{it},\ t = 1,\ldots,T\}$ be the subspace of R^T over which the latent vector, y_i^*, is consistent with the sequence of observed outcomes, where θ is the vector of parameters for the model: θ contains β and the parameters which characterize the covariance matrix Ω. The indicator for the observed sequence of participation decisions is

$$1(y_i^* \in D(\theta, y_i)) = 1(\mu_i + \Gamma v_i \in D(\theta, y_i)).$$

This can be written as the product of indicators of recursively defined events

$$1(y_i^* \in D(\theta, y_i)) = 1(v_{i1} \in D_1(v_{i1})) \prod_{t=2}^{T} 1(v_{it} \in D_t(v_{i,<t}))$$

where the subscript "$<t$" denotes the subvector whose elements are the first $(t-1)$ elements of the vector; $D_1(v_{i1}) = \{v_{i1}: \gamma_{1i1} < v_{i1} < \gamma_{2i1}\}$ and $D_t(v_{i,<t}) = \{v_{it}: \gamma_{1it} < v_{it} < \gamma_{2it}\}$, $t = 2, \dots, T$. Also,

$$\gamma_{1it} = \frac{a_{it} - \mu_{it} - \Gamma_{t,<t} \nu_{i,<t}}{\Gamma_{tt}};$$

$$\gamma_{2it} = \frac{b_{it} - \mu_{it} - \Gamma_{t,<t} \nu_{i,<t}}{\Gamma_{tt}};$$

$$a_{it} = \begin{cases} 0 & \text{if } y_{it} = 1, \\ -\infty & \text{if } y_{it} = 0; \quad \text{and} \end{cases}$$

$$b_{it} = \begin{cases} \infty & \text{if } y_{it} = 1, \\ 0 & \text{if } y_{it} = 0. \end{cases}$$

The likelihood contribution for the ith observation then is

$$1_i = P(y_i^* \in D(y_i)) = P(v_{i1} \in D_1(v_{i1})) \prod_{t=2}^{T} P(v_{it} \in D_t(v_{i,<t}))$$

$$= \prod_{t=1}^{T} \{\Phi(\gamma_{2it}) - \Phi(\gamma_{1it})\}$$

where $\Phi(\cdot)$ is the standard Normal CDF.

The SRC simulator simulates this contribution unbiasedly by randomly drawing $\tilde{\nu}_{it}$'s from $D_t(\cdot)$ as follows. If ξ_{it} is a draw from the $U[0,1]$ distribution, which is fixed throughout the estimation procedure, then

$$\tilde{\nu}_{it} = \Phi^{-1}(\xi_{it}\Phi(\tilde{\gamma}_{2it}) + (1 - \xi_{it})\Phi(\tilde{\gamma}_{1it}))$$

where $\tilde{\gamma}_{1it}$ and $\tilde{\gamma}_{2it}$ are the simulated counterparts to γ_{1it} and γ_{2it}. Note that, at each iteration of the estimation procedure, these simulations are made conditional on the current vector of parameter values and the covariates.

APPENDIX 2: MONTE CARLO EXPERIMENTS

This Appendix describes a small Monte Carlo simulation experiment to assess the properties of the MSL estimators in multivariate discrete choice models. A more detailed discussion of the framework and results can be found in Hyslop (1995). The data generating process (DGP) is a panel data probit model which includes a single observed exogenous variable, unobserved heterogeneity, a stationary AR(1) error component, and a stationary first-order Markov process state dependence:

(A.1a) $\quad y_{it} = 1(\beta_{10} + \beta_{11} Z_{it} + \gamma y_{it-1} + u_{it} > 0) \qquad (t = 1, \dots, T-1);$

(A.1b) $\quad u_{it} = \alpha_i + \varepsilon_{it}; \qquad \varepsilon_{it} = \rho \varepsilon_{it-1} + v_{it};$

where $\alpha_i \sim$ iid $N(0, \sigma_\alpha^2)$ and $v_{it} \sim$ iid $N(0, (1-\rho^2)(1-\sigma_\alpha^2))$. For DGPs with state dependence $(\gamma \neq 0)$, the stochastic process is in equilibrium in the initial sample period. One feature of the experiments is that, in order to replicate the exogenous variation in the empirical data, the exogenous variable, Z_{it}, is a composite variable generated from the data used in the empirical analysis. Specifically, $Z_{it} = X_{it}' \hat{\beta}$, where X_{it} is a vector of demographic variables, and $\hat{\beta}$ is a vector of parameters: for DGPs with (without) state dependence, $\hat{\beta}$ is a vector of parameter estimates from a probit model with (without) state dependence. (Hyslop (1995) provides details of the actual specifications used for these purposes.) Table A1 presents summary statistics on the Z_{it} variables used in the simulation experiments.

The model specification estimated is a probit model:

(A.2a) $$y_{i0} = 1(\beta_{00} + \beta_{01} Z_{i0} + u_{i0} > 0), \quad \text{and}$$

(A.2b) $$y_{it} = 1(\beta_{10} + \beta_{11} Z_{it} + \gamma y_{it-1} + u_{it} > 0) \qquad (i = 1, \dots, N, \text{ and } t = 1, \dots, T-1),$$

where $u_{it} = \alpha_i + \varepsilon_{it}$; $\varepsilon_{it} = \rho \varepsilon_{it-1} + \nu_{it}$, $t = 1, \dots, T-1$. For identification purposes I normalize $\text{Var}(u_{i0}) = 1$, and $\text{Var}(u_{it}) = \text{Var}(\alpha_i + \varepsilon_{it}) = \sigma_\alpha^2 + \sigma_\varepsilon^2 = 1$. For reasons of parsimony, the correlations between u_{i0} and u_{it} are restricted to be equal for all t; the effects of this restriction are partially investigated below.

The experiments vary the values of $(\gamma, \sigma_\alpha^2, \rho)$ in (A.1a) and (A.1b), while the other parameters are held fixed: $\beta_{10} = 0$ and $\beta_{11} = 1$. Three DGPs with no state dependence and varying degrees of heterogeneity and serial correlation are used: $(\gamma, \sigma_\alpha^2, \rho) = (0, 0.5, 0)$, $(0, 0.1, 0.7)$, and $(0, 0.5, 0.5)$. In addition four DGPs with state dependence are considered: $(\gamma, \sigma_\alpha^2, \rho) = (1, 0.1, 0)$, $(1, 0.5, 0)$, $(0.5, 0.5, 0.5)$ and $(1, 0.5, -0.5)$. The sampling frame consists of $N = 1000$ observations over $T = 7$ periods. Each DGP is simulated 25 times, and estimated by MSL using the SRC simulator with 20 replications per observation.

Table A2 summarizes the simulation results. The parameters of interest in equations (A.2a) and (A.2b) are the structural model coefficients and the error components parameters $(\beta_{10}, \beta_{11}, \gamma, \sigma_\alpha^2 \rho)$, while β_{00}, β_{01}, and $\rho_0 = \text{corr}(u_{i0}, u_{it})$ are nuisance parameters. For each DGP, the mean and median estimated parameter values and also the estimated standard error of the estimates are presented.

First, the results for DGP I, which contains only unobserved heterogeneity, show evidence of significant downward bias in σ_α^2, ρ, and β_{10} and upward bias in γ. The bias in each parameter is on the order of the standard error of the parameter estimate. For DGP II ($\gamma = 0$, $\sigma_\alpha^2 = 0.1$, $\rho = 0.7$), the results show downward bias in the estimates of ρ and β_{10}, and upward bias in γ and σ_α^2; again the bias is on the order of the standard errors of the estimates. Also, for this DGP, the parameters are much less precisely estimated, and the means and medians of γ and ρ are quite different, reflecting some skewness in their respective distributions. The results for the third DGP, with substantial permanent and serially correlated transitory components of error ($\sigma_\alpha^2 = \rho = 0.5$), show a substantial amount of bias in each of the parameters except for β_{11}. In particular, there is substantial upwards and downwards bias in γ and ρ respectively.

TABLE A1

SUMMARY STATISTICS OF THE EXOGENOUS VARIABLES USED IN THE MONTE CARLO EXPERIMENTS

Year	Average Participation	Exogenous Variables in Models: Without Lagged Dependent Variable	With Lagged Dependent Variable
0	0.726	0.655 (.39)	−0.150 (.29)
1	0.709	0.582 (.38)	−0.148 (.29)
2	0.714	0.551 (.38)	−0.142 (.30)
3	0.702	0.542 (.38)	−0.128 (.30)
4	0.711	0.598 (.38)	−0.052 (.31)
5	0.751	0.711 (.39)	0.053 (.32)
6	0.738	0.663 (.40)	−0.060 (.33)

Notes: Standard deviations in parentheses. Number of observations = 1000.

TABLE A2

SUMMARY STATISTICS OF MONTE CARLO SIMULATIONS

	Parameter				
DGP	β_{10}	β_{11}	γ	σ_α^2	ρ
I:	**0**	**1**	**0**	**0.5**	**0**
Mean	−0.080 (.01)	0.986 (.01)	0.140 (.02)	0.455 (.01)	−0.097 (.02)
Median	−0.076 (.02)	0.981 (.02)	0.152 (.03)	0.457 (.01)	−0.084 (.02)
Standard Error	0.065	0.061	0.107	0.040	0.078
II:	**0**	**1**	**0**	**0.1**	**0.7**
Mean	−0.148 (.03)	0.993 (.02)	0.240 (.07)	0.184 (.02)	0.514 (.04)
Median	−0.075 (.02)	1.020 (.01)	0.119 (.02)	0.188 (.01)	0.588 (.03)
Standard Error	0.183	0.102	0.368	0.081	0.232
III:	**0**	**1**	**0**	**0.5**	**0.5**
Mean	−0.341 (.03)	0.920 (.02)	0.633 (.06)	0.406 (.01)	−0.010 (.04)
Median	−0.364 (.03)	0.918 (.01)	0.661 (.04)	0.409 (.01)	−0.052 (.02)
Standard Error	0.158	0.070	0.280	0.067	0.204
IV:	**0**	**1**	**1**	**0.1**	**0**
Mean	−0.049 (.01)	0.978 (.01)	1.077 (.02)	0.078 (.01)	−0.035 (.01)
Median	−0.050 (.02)	0.984 (.02)	1.048 (.03)	0.078 (.01)	−0.040 (.01)
Standard Error	0.063	0.070	0.083	0.025	0.047
V:	**0**	**1**	**1**	**0.5**	**0**
Mean	−0.087 (.02)	0.958 (.02)	1.151 (.03)	0.432 (.01)	−0.035 (.02)
Median	−0.095 (.03)	0.929 (.03)	1.150 (.04)	0.424 (.02)	−0.036 (.01)
Standard Error	0.100	0.097	0.150	0.057	0.075
VI:	**0**	**1**	**0.5**	**0.5**	**0.5**
Mean	−0.431 (.02)	0.896 (.02)	1.302 (.02)	0.317 (.01)	0.068 (.01)
Median	−0.432 (.02)	0.899 (.02)	1.315 (.03)	0.318 (.01)	0.065 (.01)
Standard Error	0.082	0.101	0.139	0.057	0.064
VII:	**0**	**1**	**1**	**0.5**	**−0.5**
Mean	0.078 (.01)	1.025 (.01)	0.872 (.02)	0.539 (.01)	−0.440 (.01)
Median	0.068 (.02)	1.015 (.01)	0.901 (.04)	0.537 (.02)	−0.443 (.02)
Standard Error	0.066	0.065	0.098	0.047	0.043

Notes: Bootstrap standard errors in parentheses, using 100 bootstrap replications. Standard errors of the parameter estimates are computed as sample standard deviations: these are comparable to the parameter standard errors estimated using the outer product of the scores. Each model is simulated 25 times and estimated using $R = 20$ replications per observation. The estimation is based on $N = 1000$ observations over a $T = 7$ period panel.

Consider next the four DGPs that include positive state dependence. The results for these experiments again generally find positive bias in the estimated γ, while the estimated σ_α^2 and ρ are negatively biased. However, with the exception of DGP VI, the results show that the bias is relatively modest, and there is very little bias in β_{11}. Interestingly, for DGP VII ($\gamma = 1$, $\sigma_\alpha^2 = 0.5$, $\rho = -0.5$), the direction of bias is reversed.

Finally, I investigate whether the estimated bias in the DGPs with positive state dependence or positive serial correlation is due to a small number of simulation replications, or some other source such as the restriction imposed on $\text{corr}(u_{i0}, u_{it})$, which would lead to MLE bias. In the absence of MLE bias, the score of the log-likelihood function evaluated at the true parameter values will on

TABLE A3

SUMMARY STATISTICS OF SCORE-LIKELIHOOD FUNCTION DGP IV

	Number of Replications, R			
	10	20	100	1000
χ^2-statistic (Score = 0)	95.156	87.425	33.506	9.448
(p-values, 8df)	(<0.001)	(<0.001)	(<0.001)	(0.306)

Notes: P-values, in parentheses. Estimation is based on 1000 observations over a $T=7$ year panel. The model ignores the initial conditions, and is simulated and estimated using equations (A.2a) and (A.2b). Results based on 100 simulations.

average be zero; alternatively, if there is MLE bias, the score will be systematically different from zero. This implies that if the observed bias is only simulation bias then, as the number of simulations increases, the score evaluated at the true parameter values will converge to zero. To examine this proposition, the log-likelihood function for DGP VI with known initial conditions (see Hyslop (1995) for details) was simulated 100 times using 10, 20, 100, and 1000 SRC replications per observation, and its score evaluated at the true parameter values. Table A3 presents χ^2 statistics for the joint hypothesis of a null mean score. These results find that, for less than 100 replications, the hypothesis can be rejected. However, there is no evidence against this hypothesis when 1000 replications are used. These results imply that the bias observed in the Monte Carlo experiments presented here is due to simulation bias.

These results suggest that the performance of the MSL estimator with 20 replications is adequate if there is not positive state dependence, serial correlation, and heterogeneity. However, for processes that have substantial positive serial persistence in the underlying process, the bias is significant. In these cases, the number of replications required to reduce the bias may be prohibitively high.

REFERENCES

AHN, SEUNG C., AND PETER SCHMIDT (1995): "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5–27.

ALTONJI, JOSEPH G. (1986): "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy*, 94, S176–S215.

ARELLANO, MANUEL, AND STEPHEN BOND (1991): "Some Tests of the Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277–297.

BLUNDELL, RICHARD, JOHN HAM, AND COSTAS MEGHIR (1995): "Unemployment, Discouraged Workers and Female Labor Supply," Manuscript, Department of Economics, University of Pittsburgh.

BÖRSCH-SUPAN, AXEL, AND VASSILIS A. HAJIVASSILIOU (1993): "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models," *Journal of Econometrics*, 58, 347–368.

BOUND, JOHN, DAVID A. JAEGER, AND REGINA M. BAKER (1995): "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443–450.

BOUND, JOHN, AND ALAN B. KRUEGER (1991): "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, 9, 1–24.

BROWNING, MARTIN (1992): "Children and Household Economic Behavior," *Journal of Economic Literature*, 30, 1434–1475.

BURDETT, KENNETH, AND DALE T. MORTENSEN (1978): "Labor Supply Under Uncertainty," *Research in Labor Econometrics*, 2, 109–157.

BUTLER, J. S., AND ROBERT MOFFITT (1982): "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model," *Econometrica*, 50, 761–764.

CARD, DAVID, AND DANIEL SULLIVAN (1988): "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," *Econometrica*, 56, 497–530.

CHAMBERLAIN, GARY (1984): "Panel Data," Chapter 22 in *Handbook of Econometrics*, Vol. 2, ed. by Z. Griliches and M. D. Intrilligator. Amsterdam: Elsevier Science Publishers B.V.

COGAN, JOHN F. (1981): "Fixed Costs and Labor Supply," *Econometrica*, 49, 945–963.

CUTLER, DAVID M., AND LAWRENCE F. KATZ (1991): "Macroeconomic Performance and the Disadvantaged," *Brookings Papers on Economic Activity*, 2, 1–61.

ECKSTEIN, ZVI, AND KENNETH I. WOLPIN (1989a): "Dynamic Labor Force Participation of Married Women and Endogenous Work Experience," *Review of Economic Studies*, 56, 375–390.

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

——— (1990): "On the Estimation of Labor Force Participation, Job Search, and Job Matching Models using Panel Data," Chapter 4 in *Advances in the Theory and Measurement of Unemployment*, ed. by Yoram Weiss and Gideon Fishelson. New York: Macmillan.

GOURIEROUX, CHRISTIAN, AND ALAIN MONFORT (1993): "Simulation-based Inference: A Survey with Special Reference to Panel Data Models," *Journal of Econometrics*, 59, 5–33.

HAJIVASSILIOU, V. A. (1993): "Simulation Estimation Method for Limited Dependent Variable Models," Chapter 19 in *Handbook of Statistics*, Vol. 11, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod. Amsterdam: Elsevier Science Publishers B.V.

HAJIVASSILIOU, VASSILIS, AND DANIEL L. MCFADDEN (1998): "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica*, 66, 863–896.

HAJIVASSILIOU, VASSILIS, AND PAUL RUUD (1994a): "Classical Estimation Methods for LDV Models Using Simulation," Chapter 40 in *Handbook of Econometrics*, Vol. 4, ed. by R. Engle and D. McFadden. Amsterdam: Elsevier Science Publishers B.V.

——— (1994b): "Simulation to the Max," Manuscript, Department of Economics, UC Berkeley.

HECKMAN, JAMES J. (1978): "Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence Against the Hypothesis of Spurious State Dependence," *Annales de l'INSEE*, No. 30-31, 227–269.

——— (1981a): "Statistical Models for Discrete Panel Data," Chapter 3 in *Structural Analysis of Discrete Data*, ed. by Charles Manski and Daniel McFadden. Cambridge, MA: MIT Press.

——— (1981b): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," Chapter 4 in *Structural Analysis of Discrete Data*, ed. by Charles Manski and Daniel McFadden. Cambridge, MA: MIT Press.

——— (1981c): "Heterogeneity and State Dependence," in *Studies in Labor Markets*, ed. by Sherwin Rosen. Chicago: University of Chicago Press.

——— (1984): "The χ^2 Goodness-of-Fit Statistic for Models with Parameters Estimated from Microdata," *Econometrica*, 52, 1543–1547.

——— (1993): "What has been Learned about Labor Supply in the Past Twenty Years?" *American Economic Review*, 83, 116–121.

HECKMAN, JAMES J., AND THOMAS E. MACURDY (1980): "A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47, 47–74.

——— (1982): "Corrigendum on A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 49, 659–660.

HECKMAN, JAMES J., AND ROBERT J. WILLIS (1977): "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," *Journal of Political Economy*, 85, 27–58.

HOTZ, V. JOSEPH, FINN E. KYDLAND, AND GUILHERME L. SEDLACEK (1988): "Intertemporal Preferences and Labor Supply," *Econometrica*, 56, 335–360.

HOTZ, V. JOSEPH, AND ROBERT A. MILLER (1988): "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply," *Econometrica*, 56, 91–118.

HYSLOP, DEAN R. (1994): "The Covariance Structure of Intrafamily Earnings, Rising Inequality and Family Labor Supply," Industrial Relations Section Working Paper, #339, Princeton University.

——— (1995): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Participation Behavior: Monte Carlo Evidence and Empirical Results for Married Women," Industrial Relations Section Working Paper #347, Princeton University.

JAKUBSON, GEORGE (1988): "The Sensitivity of Labor-Supply Parameter Estimates to Unobserved Individual Effects: Fixed- and Random-Effects Estimates in a Nonlinear Model Using Panel Data," *Journal of Labor Economics*, 6, 302–329.

KEANE MICHAEL P. (1993): "Simulation Estimation for Panel Data Models with Limited Dependent Variables," Chapter 20 in *Handbook of Statistics*, Vol. 11, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod. Amsterdam: Elsevier Science Publishers B.V.

——— (1994): "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62, 95–116.

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

KILLINGSWORTH, MARK R., AND JAMES J. HECKMAN (1986): "Female Labor Supply: A Survey," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard. Amsterdam: North Holland.

LERMAN, STEVEN R., AND CHARLES F. MANSKI (1981): "On the Use of Simulated Frequencies to Approximate Choice Probabilities," Chapter 7 in *Structural Analysis of Discrete Data*, ed. by Charles Manski and Daniel McFadden. Cambridge, MA: MIT Press.

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

MCFADDEN, DANIEL (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 57, 995–1026.

MCFADDEN, DANIEL, AND PAUL A. RUUD (1994): "Estimation by Simulation," Manuscript, UC Berkeley.

MINCER, JACOB (1962): "Labor-force Participation of Married Women: a Study of Labor Supply," in *Aspects of Labor Economics*, ed. by H. G. Lewis. Princeton: Princeton University Press.

MOORE, DAVID (1977): "Generalized Inverses, Wald's Method and the Construction of Chi-Squared Tests of Fit," *Journal of the American Statistical Association*, 72, 131–137.

MROZ, THOMAS M. (1987): "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica*, 55, 765–799.

MÜHLEISEN, MARTIN (1991): "On the Use of Simulated Estimators for Panel Models with Limited Dependent Variables," Manuscript, University of Munich.

——— (1992): "Simulation Estimation of State Dependence Effects in Unemployment," University of Munich, Discussion Paper 92-21.

PAKES, ARIEL, AND DAVID POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

PENCAVEL, JOHN (1986): "Labor Supply of Men: A Survey," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard. Amsterdam: North Holland.

RUST, JOHN, AND CHRISTOPHER PHELAN (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65, 781–831.

SHAW, KATHRYN L. (1992): "The Life-Cycle Labor Supply of Married Women and its Implications for Household Income Inequality," *Economic Inquiry*, 30, 659–672.

WHITE, HALBERT (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.