A Partial Survey of Recent Research on The Labor Supply of Women

By JAMES J. HECKMAN*

This paper is a progress report on some recent research on the labor supply of married women. In previous work on this topic two conceptually distinct interpretations of the coefficients of regressions of labor force participation status (measured as a rate for a group or by a dummy variable for an individual) on wage rates and unearned income have been presented. The first interpretation, which stems from Jacob Mincer's pioneering work and which is pursued in later work by Marvin Kosters, Glen Cain, and Orley Ashenfelter and the author, interprets estimated wage and income coefficients as estimates of substitution and income effects. The later work in this tradition interprets the income and substitution effects within the Hicks-Slutsky framework. The operating assumption in this literature is that the estimated wage and income coefficients from an analysis of participation admit the same theoretical interpretation as wage and income coefficients obtained from hours of work regressions. The economic model that yields such an interpretation is a life cycle model of labor supply.

A second interpretation, that has been the guiding principle of much recent work on labor supply, stresses the point that labor force participation at a point in time is a discrete decision. A woman either participates or does not participate, and estimates of participation equations are conceptually distinct from estimates of hours of work functions. Participation regressions describe "corner phenomena" and do not estimate "interior solution" Hicks-Slutsky income and substitution effects, although they certainly estimate parameters of the utility function of consumers. The modern statement of the second view has been developed in papers by Yoram Ben-Porath and H. Gregg Lewis which were important stimuli to later work on female labor supply by Reuben Gronau, the author (1974), J. Cogan, and Giora Hanoch.

While the development of the second approach is still underway, it is important to note that at the present time, analysts operating within the second tradition have ignored the focus on the *life cycle* that underlies the first tradition. Analysts operating within the second tradition assume that their one-period models of labor supply apply to the data used to test their models—typically hours of work or weeks of work in a survey year—and ignore the point that most consumers have ample opportunity to substitute time and goods over the life cycle, and to invest in human capital.

This paper presents some of the principle results of recent research on life cycle labor force participation that attempts to merge these two traditions. In order to place this work in a suitable perspective, it is necessary to review some of the previous literature. In this review, certain implicit assumptions in previous work are spelled out for what is believed to be the first time. Given space limitations, only informal arguments are presented. A more complete analysis is available in a companion paper (see the author, 1977a).

I. Mincer's Model and Subsequent Developments

The basic Mincer model is presented in a few lines of his seminal paper:

*University of Chicago. This work was partially supported by a U.S. Department of Labor grant to the National Bureau of Economic Research. I have greatly benefited from discussions with T. MaCurdy and J. Mincer. Any remaining errors are mine.

> In a broad view, the quantity of labor supplied to the market by a wife is the fraction of her married life during which she participates in the labor force. Abstracting from the temporal distribution of labor force activities over a woman's life, this fraction could be translated into a probability of being in the labor force in a given period of time for an individual, hence into a labor force rate for a large group of women. [p. 68]

Mincer goes on to develop a model of the timing of labor force activity over the life cycle. The implicit model that underlies the econometric models of Mincer, Cain, and Ashenfelter and the author is based on a lifetime utility model of consumer decision making. Since this model has never been explicitly developed, it seems useful to do so here to emphasize the principles involved.

Lifetime utility is a well-behaved function of lifetime consumption of goods X and leisure L. The maximum amount of leisure that can be consumed is T. Let ϵ be an unobserved "taste" or "household production" variable assumed independent of other variables. Lifetime utility is $U = U(X, L, \epsilon)$. The consumer faces a permanent real wage W net of all money costs of work. The price of goods is unity. Assume a zero interest rate. Given resources A, the consumer works sometime in his life if

$$M(A, T, \epsilon) = \frac{U_2(A, T, \epsilon)}{U_1(A, T, \epsilon)} \leq W \tag{1}$$

where U_j is the partial of U with respect to argument j (assumed here to be positive for all variables). The term on the left of the inequality is the marginal rate of substitution between goods and leisure at full leisure, i.e., the lifetime reservation wage.

If the consumer works, labor supply is determined by the solution to the equations

$$M(X, L, \epsilon) = \frac{U_2(X, L, \epsilon)}{U_1(X, L, \epsilon)} = W \tag{2a}$$

$$X + WL = WT + A \tag{2b}$$

The equations can be solved for the fraction of total time worked h as $h = (T - L)/T = h(W, A, \epsilon)$. Denoting h_j as the partial of h with respect to its jth argument, h_1 is the ordinary (uncompensated) Slutsky effect of permanent wage change on labor supply and h_2 is the income effect.

By construction, the simple theory is silent on the timing of labor supply over the life cycle. Ben-Porath and others interpret Mincer's paper as enriching the simple theory by assuming that, aside from "transitory" factors (for example, factors like children, transitory income variation, and the like), the timing of participation over the life cycle is "random." The probability of finding a randomly selected consumer of a group of consumers with permanent wage W and assets A in the labor force at any point in time is the population mean of h. *Assuming that all consumers work sometime in their life cycle,* regressions of a participation measure on variables W and A estimate the mean values of h_1 and h_2. These estimated coefficients may be interpreted as estimates of Hicks-Slutsky substitution and income effects.

It is not necessary to assume that observationally identical individuals have identical participation probabilities. Depending on the distribution of the unobserved variable ϵ in the population, there may be considerable dispersion in tastes, or unobserved permanent components of wage rates, leading to a distribution in population probabilities. The only crucial assumptions are 1) that the participation rate measures the fraction of lifetime labor supplied to the market, and 2) that the fraction of lifetime labor supplied to the market is generated by an *interior solution* to the lifetime utility-maximization problem. Assumption 1 is quite weak and is satisfied in a wide variety of situations while assumption 2 is much more demanding. The only requirement for assumption 1 to be valid is that the economic environment is stationary or, if it is not, that it be changing in a way that is known or can be determined.[1] In particular one must be able to control for cyclical and cohort effects.

[1]Thus Ben Porath's dichotomy between homogeneity and heterogeneity in his analysis of hypothesis

Ben-Porath has recently noted that many women do not participate in the labor force over substantial time intervals (ten years in his data). Later work by the author and Willis suggests that the proportion of married women who never participate is substantial in short intervals of data. However, a recent note by Mincer and Haim Ofek suggests that the fraction of married women who never participate (after leaving school) is only 4–5 percent when lifetime labor supply is measured (however, in their own data, 25 percent of women do not work after they are married). Given that *some* women never work, what is the implication of this finding for the interpretation of participation regressions?

The answer to this question is simple, although it is not in the literature. The probability that a randomly selected member of a group with permanent wage W and asset position A *ever* works is:

$$Pr(W \geq M(A, T, \epsilon)) = P(W, A)$$

where the probability is computed from the population distribution of unobservable variables. The expected probability that a randomly selected individual with wage W and assets A who ever works will be found working at a randomly selected point in time is $E_c(h) = E[h(W, A, \epsilon) | W \geq M(A, T, \epsilon)]$ where $E_c(h)$ is the fraction of lifetime hours that are worked by an ever-working individual. Recent work in labor supply demonstrates that $\partial E_c(h)/\partial W \neq h_1$ because the left-hand side term includes the effect of movement across taste distributions as the wage is raised, while the right-hand side term is the Slutsky effect holding tastes constant. The left-hand term includes the effect of an increase in wages on the mean unobservable "taste for work" for those who work while the right-hand side term does not.

The probability that a randomly selected individual from a population of people with permanent wage W and asset position A works at a point in time is

$$P(W, A)\, E_c(h) \tag{3}$$

If everyone works sometime in the life cycle, as Mincer maintains, $P(W, A) = 1$ for all values of W and A and $E_c(h) = E(h)$. Regression estimates of participation status on wage rates and assets estimate h_1 and h_2, the Hicks-Slutsky terms.

If the population is dichotomized into full-time workers and full-time nonworkers, $E_c(h) = 1$, and participation regressions estimate the parameters of $P(W, A)$, i.e., the parameters of M, the reservation wage function at full leisure, and the parameters of the distribution ϵ. (This is Ben-Porath's hypothesis II.)

In the general case, cross-section participation equations estimate the first partials of expression (3) which are *not* h_1 and h_2, the Hicks-Slutsky terms, and which are *not* P_1 and P_2, parameters derived from the marginal rate of substitution function M and the parameters of the distribution of ϵ. Estimates of the parameters of expression (3) yield the parameters of the aggregate labor supply function for women in terms of permanent wage rates and assets, since this expression combines the effect of permanent wage and asset changes on the entry of women into the market and on hours supplied by previously working women. The aggregate supply function can be used to predict the probability that a randomly selected woman with wage W and assets A will be in the market at a point in time, and it can be used to predict the fraction of her lifetime that a randomly selected woman with these traits will work. Thus the agreement in the estimates of the permanent wage elasticities for these two measures of labor supply reported by Cain (p. 100) is reassuring.

The simple theory just discussed is not entirely satisfactory for three reasons: it rests on certain implicit assumptions which when brought to light make it somewhat unpalatable; it does not provide a natural framework for integrating different dimensions of labor supply activity in a unified model; and it is easy to misuse it as a guide to dynamic models of labor supply.

The assumption about the source of ran-

I and hypothesis II in his paper is quite misleading. The real issue is "interior solution vs. corner solution" and not homogeneity vs. heterogeneity.

domness that generates the timing of the work decision is critical in interpreting estimates of the response of participation to wage rates and assets as Hicks-Slutsky effects, even if everyone works. One plausible source of variation in the timing of labor supply over the life cycle is differences in transitory components of wage rates known to the consumer but unknown to the observing economist.[2] This source is mentioned in Mincer's work. If consumers are rational, and if wages (or shadow costs) vary over the life cycle, even if all work sometime in their lifetime, regressions of participation status on permanent wages and assets *do not* estimate income and substitution effects.

To see this, break life up into T periods with wages $W_1, \ldots, W_T$. Denote the largest wage by W^1 and the smallest by W^T. The permanent wage can be defined as the average wage over the life cycle, as is customary. The decision to ever participate is made by comparing the lifetime reservation wage to the highest wage W^1 and *not* the permanent wage as in inequality (1). The number of periods a consumer works is not determined from equations (2a) and (2b). In fact, a consumer works k periods if

$$\text{(4a)}\quad M\left(A + \sum_{i=1}^{k-1} W^i, T - k + 1, \epsilon\right) \leq W^k$$

$$\text{(4b)}\quad M\left(A + \sum_{i=1}^{k} W^i, T - k, \epsilon\right) > W^{k+1}$$

While the permanent wage is a determinant of these inequalities, it does not play the same role in them as in equations (2a) and (2b). The marginal wage relevant to the participation decision, and the decision to work k periods, is not the average wage. The average wage for the periods in which the consumer works, which is relevant for evaluating "income effects," exceeds the marginal wage relevant for evaluating "substitution effects" in this model.

Given the joint distribution of wages, and ϵ (assumed independent of wages), inequalities (4a) and (4b) can be used to derive the probability that the consumer works $t = 0, \ldots, T$ periods as a function of permanent wage W and permanent asset position A; $P_t = P_t(W, A)$, $\Sigma_{t=0}^{T} P_t = 1$. The partial derivative of each P_t function with respect to W does *not* directly estimate a Hicks-Slutsky income and substitution effect. However, utilizing a direct extension of the statistical model of Richard Rosett and Forrest Nelson, estimates of the P_t functions can be used to piece together segments of the lifetime utility function at $T + 1$ intervals so that rough estimates of the consumer's indifference system may be obtained. These estimates can be used to estimate Hicks-Slutsky income and substitution effects.

The probability of finding a randomly selected member of the population with permanent wage W and asset position A in the labor force in any period, which is also the expected proportion of life that an individual will work, is

$$\text{(5)}\quad \sum_{t=0}^{T} \frac{t\, P_t(W, A)}{T}$$

The partial derivatives of this function are what is estimated in cross-section regressions of participation status on permanent wage rates and assets, and these partials are clearly *not* the usual income and substitution effects.

What type of "randomness" will lead to the simple model employed in the previous literature? One possibility is a random utility model (see Daniel McFadden) in which the consumer's work decision in each period is made by tossing a coin with probability $h (0 < h < 1)$ of working (assuming T is "large"). But transitory wage or cost variation will not lead to this model if the consumer is rational in the neoclassical sense of that term.

Estimates of equation (5) from the cross section cannot be used to estimate the distribution of participation sequences in the population or the distribution of the number of periods ever worked over the life cycle (given by the component parts of equation (5)) unless the random utility model is

[2] A similar analysis can be made if there are unobserved deviations in household production or tastes over the life cycle.

assumed, successive tosses of the coin are independent, and individuals are identical (the distribution of ϵ is degenerate). In the general case, one cannot use the cross-section mean to estimate the probability of any sequence of participation decisions over the life cycle if there is any unobserved heterogeneity in the population (see the author and Willis).

II. Dynamic Models of Labor Supply: Some Preliminary Results

In this section, I present the most elementary version of a class of dynamic models of labor supply that have recently been estimated by Thomas MaCurdy and myself. This model can be used to interpret the interrelationship among the various dimensions of labor supply analyzed in the literature and can also shed some revealing light on certain aberrant empirical findings in the "new" labor supply literature that rigorously analyzes one-period models of labor supply and applies them to the data at hand—typically a survey year or survey week.

The consumer's utility at age t is a strictly concave, twice differentiable nonsatiable function $U(t) = G(X(t)) + J(L(t))$, $t = 0, \ldots, T$, where $X(t)$ is the consumption of goods and $L(t)$ is the consumption of leisure at age t. The maximum value of leisure is 1. The real wage is $W(t)$ and the real price of goods is unity. In this simple model, human capital accumulation is ignored, although it is not ignored in the more sophisticated models that we have estimated. Assuming no uncertainty, a parametric interest rate r and rate of time preference ρ, the consumer maximizes

$$\sum_{t=0}^{T} U(t)(1+\rho)^{-t}$$

subject to the budget constraint

$$A = \sum_{t=0}^{T} (1+r)^{-t} \cdot [W(t)[1 - L(t)] - X(t)] \tag{6}$$

where A is initial wealth. This model is essentially a finite horizon Ramsey model.

For ages at which the consumer works, the demand functions for goods and leisure are

$$X(t) = X\left(\lambda\left(\frac{1+\rho}{1+r}\right)^t\right) = [G']^{-1}\left(\lambda\left(\frac{1+\rho}{1+r}\right)^t\right)$$

$$L(t) = L\left(\lambda W(t)\left(\frac{1+\rho}{1+r}\right)^t\right) = [J']^{-1}\left(\lambda W(t)\left(\frac{1+\rho}{1+r}\right)^t\right)$$

where $X' < 0$, $L' < 0$, and λ (> 0) is the marginal utility of wealth (the LaGrange multiplier associated with the budget constraint). Goods and leisure are assumed to be normal.

If the consumer does not work at age t, $L(t) = 1$, and

$$\frac{J'(1)}{\lambda}\left(\frac{1+\rho}{1+r}\right)^t > W(t) \tag{7}$$

The term on the left-hand side of the inequality is the shadow price of time at age *t evaluated at the equilibrium position*. The term on the right-hand side is the market wage. If the equilibrium reservation wage exceeds the market wage, the consumer does not work in period t.

The value of λ is determined jointly with the other decision variables by substituting the demand functions into resource constraint (6). Note that λ is a function of all the parameters of the model including the vector of lifetime wage rates.[3] Lifetime labor supply may be defined as the sum of the number of periods worked or the number of hours worked—two conceptually distinct measures which are related to the underlying set of parameters in different ways.

I briefly consider some of the implications of the model. They are spelled out in greater length in a companion paper (see

[3] For any monotonic transformation of the utility function $\Sigma\ U(t)\ (1 + \rho)^t$, λ defined in this way is invariant.

the author, 1977a). First consider the participation decision. Recent work on female labor supply utilizes inequality (7) to characterize the participation decision. The key assumption in that literature is that the market wage is independent of the reservation wage. In a one-period model of labor supply, this assumption is appropriate. It is also an appropriate assumption for determining the probability that a consumer will *ever* work in a life cycle model. However it is an inappropriate assumption for characterizing the participation decision in a given period if the consumer works in other periods, and if the wage in the given period is correlated with the wage in other periods, as G. Sedlacek has found is the case for women. Under the maintained assumptions, if consumers work in other periods, higher current wages are associated with lower values of λ and higher values of the reservation wage. It is thus possible that wage rates are inversely correlated with labor force participation at a point in time.

This point has bearing on some recent findings by R. Olsen and James Smith. They find that for certain groups of women, *lower* wage women are the ones more likely to participate. This finding can be explained in a one-period model by appealing to a large positive correlation between tastes for work and unobservable determinants of wages. The life cycle model can explain this fact by appealing to intertemporal correlation in wage rates and the plausible assumption that women who work today may also work some time in the future. It is significant that the "perverse" association between wage rates and participation status is found in demographic groups with the greatest volume of lifetime labor supply—such as married black women. It is in such groups that income effects are likely to be the largest.

This point also implies that the estimated value of the reservation wage or value of home time obtained through the procedures of Gronau and the author (1974) understates the true value of time at home for women with higher market wages, and overstates it for women with lower market wages, since a higher permanent wage leads to a higher value of time. The Gronau-Heckman procedures hold the market wage component of the reservation wage function fixed at the sample mean and thus understate the true *equilibrium* reservation wage for women with higher wages.

The crucial identifying restriction utilized in their work—that certain variables that determine market wages are excluded from the reservation wage equation—does not hold in a life cycle model in which some women work in periods other than the one under investigation. Only market wage variables that do not appear in other periods can serve as identifying variables and this is a limited set of variables.

Second, the probability that a randomly selected consumer will be in the labor force in a given period can be modeled as a discrete choice problem. The consumer can be in 2^{T+1} possible life cycle states corresponding to each possible participation sequence. For each sequence, there is a lifetime utility and the consumer is assumed to pick the sequence with the highest utility.[4] From the probability of different participation patterns one can derive the probability of participation at a point in time. The effect of permanent wages on participation is not a Hicks-Slutsky effect nor is it a corner solution participation effect. There are wealth effects of interpersonal wage differences that arise in evaluating the relative utility of different participation sequences.

Third, a frontal assault on the estimation of the 2^{T+1} choice discrete choice model leads to very difficult data requirements and intractable estimation problems. One can avoid these complications by following a suggestion of MaCurdy and noting that for a given consumer, λ stays fixed over the life cycle. Given λ, one can characterize consumption, labor supply, and participation decisions in terms of estimable func-

[4]This involves fixing hours of work at zero in certain periods while letting them be freely chosen in the remaining periods. Note that some sequences may be mathematically impossible to realize for certain values of the unobservable variables and for certain functional forms for preferences.

tions and current values of the variables. Thus, a simple estimation method consists of treating λ as a *fixed effect* for a consumer, and estimating λ and the parameters of the "λ constant" demand functions written above using longitudinal data. In the case of the life cycle participation decision, this leads to a multivariate probit model with fixed effects (see the author, 1977b). As I have noted elsewhere, given λ, the parameters of $[G']^{-1}$ and $[J']^{-1}$ and the budget constraint, one can determine all the parameters of the life cycle labor supply and consumption functions. Models of this sort have been estimated by MaCurdy and myself. By focusing on the parameters of the preference function, we generate the *entire* set of parameters of the life cycle labor supply functions.

An important feature of the λ constant demand system is that it permits one to avoid the inevitably *ad hoc* definitions of permanent and transitory wage and asset change, and it permits estimation of the effect of current wages as direct determinants of behavior (holding λ fixed) and as determinants of λ.

One can also parameterize λ as a function of observable variables. Then the proper specification of the labor supply equations requires current and future values of variables as determinants of current labor supply activity. In other work (see the author, 1977a), I demonstrate that future variables, such as children, unemployment of the husband, and the like, affect current behavior. In particular, estimates of the hours of work equation that omit future wages *understate* the effect of current change in wage rates on current labor supply.

REFERENCES

O. Ashenfelter and J. Heckman, "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply," *Econometrica*, Jan. 1974, *42*, 73–86.

Y. Ben-Porath, "Labor Force Participation Rates and the Supply of Labor," *J. Polit. Econ.*, May/June 1973, *81*, 697–704.

Glenn Cain, *The Labor Force Participation of Married Women*, Chicago 1966.

J. Cogan, "Labor Supply with Time and Money Costs of Participation," The Rand Corp., Aug. 1976.

R. Gronau, "The Effect of Children on the Housewife's Value of Time," *J. Polit. Econ.*, Mar./Apr. 1973, *80*, S30–S47.

G. Hanoch, "Hours and Weeks in the Theory of Labor Supply," The Rand Corp., R-1787, July 1976.

J. Heckman, "Shadow Prices, Market Wages and Labor Supply," *Econometrica*, July 1974, *42*, 679–94.

______, (1977a) "Dynamic Models of Female Labor Supply," unpublished paper, Univ. Chicago, Apr. 1977.

______, (1977b) "Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence Against the Hypothesis of Spurious State Dependence," unpublished paper, Univ. Chicago, July 1977.

______ **and R. Willis,** "A Beta Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," *J. Polit. Econ.*, Feb. 1977, *85*, 27–58.

M. Kosters, "Effects of an Income Tax on Labor Supply," in Arnold C. Harberger and Martin J. Bailey, eds., *The Taxation of Income from Capital*, Washington 1969.

H. G. Lewis, "On Income and Substitution Effects in Labor Force Participation," unpublished paper, Univ. Chicago 1967.

T. MaCurdy, "Labor Supply Decisions over The Life Cycle," unpublished paper, Univ. Chicago 1977.

D. McFadden, "Economic Applications of Psychological Choice Models," unpublished paper, Univ. California-Berkeley 1975.

J. Mincer, "Labor Force Participation of Married Women: A Study of Labor Supply," in Harold G. Lewis, ed., *Aspects of Labor Economics*, Princeton 1962.

______ **and H. Ofek,** "The Distribution of Lifetime Labor Force Participation of Married Women," unpublished paper, Columbia Univ., Aug. 1977.

R. Olsen, "An Econometric Model of Family Labor Supply," unpublished doctoral dissertation, Univ. Chicago, June 1977.

R. Rosett and F. Nelson, "Estimation of the Two-Limit Probit Regression Model," *Econometrica*, Jan. 1975, *43*, 141–46.

G. Sedlacek, "Dynamic Models of Female Wage Growth," unpublished paper, Univ. Chicago 1977.

J. Smith, "The Convergence to Racial Equality in Women's Wages," The Rand Corp., Sept. 1977.