FEMALE LABOR SUPPLY:

THEORY AND ESTIMATION

Introduction

INTRODUCTION

JAMES P. SMITH

THE RAND CORPORATION

In the last decade, research on female labor supply has expanded at a rate which even exceeds the remarkable rate of growth of the female labor force. The initial intellectual impulse was the work of Jacob Mincer (1962) who contributed a powerfully simple explanation of discrepancies between time series and cross-sectional market work patterns for white married women using the standard decomposition of income and price effects of traditional price theory. In his seminal paper on the allocation of time (1965), Gary Becker generalized the role of time in economic activity, so that time became a central element in decisions affecting fertility, health, location, and many other things. Becker's theory served also as the foundation for the household production model that became so popular in the early 1970's. Since the work of Mincer and Becker, several substantive methodological advancements have been made. The proposals for a National Income Maintenance Plan produced a number of small-scale negative income tax experiments that not only provided additional data sources for analysts, but raised questions about the usefulness of experimental versus nonexperimental data and the ability of economists to predict the potential impact of a national long-term income maintenance program. The simple model of labor supply choice was also extended in a number of dimensions. These included issues related to the family and the ability of husbands and wives to substitute their time in the market and nonmarket sectors; life-cycle patterns of time allocation; the interrelationship between life-cycle decisions regarding human capital accumulation and the consumption of market goods and leisure; uncertainty; and difficulties encountered with a corner solution when the majority of married women elect to work zero hours in the market. More recently, research has increasingly focused on the appropriate method of estimation. This trend is best characterized by the work of Gronau (1974) on selectivity bias, and of Heckman (1974b) on the joint distribution of hours and wages within censored samples.

Although this recent work has clearly extended the research frontier, the theory of labor supply and its related econometric methodology still contains many unresolved issues. The essays included in this volume address themselves to an important subset of these theoretical and econometric problems. The range of subjects covered is indeed impressive. In addition to the problems encountered with censored samples in estimating wage and hours functions, the topics include the choice between linear and nonlinear methods of estimation, the availability of alternative definitions of labor supply (participation, weeks worked, and annual hours), the role of time and money costs in the decisions to work and the extent of work, and the ability of the life-cycle approach to interpret otherwise anomalous empirical estimates. In each paper, theoretical models are tested empirically using a variety of existing micro-data sets.

This introduction has a threefold purpose. The first goal is to relate this volume to the existing literature so that the reader is better able to appreciate the motivation for these essays, the problems they are attempting to resolve, and their departure from conventional methods of analysis. Since much of the work concentrates on estimation techniques, I will summarize briefly the insights and problems associated with those articles that are the direct intellectual antecedents of the research contained in this book. The second purpose is to highlight the main theoretical and econometric innovations of the individual essays, so that the interested reader can see the forest from the midst of a fascinating but complex array of trees. Finally, I attempt to synthesize the principal empirical findings with an eye toward identifying the major similarities and differences emerging from the separate papers.

Until recently, the standard practice in empirical research was to estimate labor supply functions over samples of working women only. This was done either directly by confining the analysis exclusively to labor force participants or indirectly by imputing a wage to nonworking women from a wage equation estimated over a sample of workers. This restriction was forced by the absence of data on the value of time for women who were not working. Much of the econometric work on labor supply during the past five years has dealt with alternative methods of estimating wage and labor supply functions that are free of this censoring bias.

The first step toward resolving the censoring problem involved a clearer understanding on the nature of the decision to participate in the labor force. Economists frequently used alternative definitions of labor supply—annual hours, weekly hours, weeks worked, and labor force participation rates—interchangeably in empirical work. This practice was correctly criticized in papers by Ben-Porath (1973a) and Lewis (1972). They both argued that labor force participation involves a comparison between market wage offers ("W_m") and the value of home time at zero hours of work (the reservation wage "W_r"). A participation rate of 40 percent indicates simply that 40 percent of the women have market wages that exceed their reservation wages.

Ben-Porath considered two alternative interpretations of observing 40 percent of women in the labor force at any particular time. In the first, the 40 percent represents an estimate of the average desired fraction of time to spend in the labor force for every woman in that group. In any given week only some of these women will be labor force members (due to transitory and short-run factors). But over a sufficiently long time interval, each woman will spend 40 percent of her available time in the labor force. If this were an accurate characterization of the world, the standard Slusky interpretation of coefficients for wages and other variables in a participation equation as conveying information about income and substitution effects and as being qualitatively similar to hours of work equations would have some validity. There would, in fact, be little reason to distinguish among alternative definitions of labor supply. The second view suggested by Ben-Porath is that labor force participation rates reflect persistent differences among women. In this case, the participation rate indicates that 40 percent of the women desire in some long-run sense to be labor force members, while the other 60 percent do not. The difference in the interpretation we give to participation equations can be illustrated simply. Let there be no variation in market wages, but variation only in female home wages. Labor force participation rates would then measure the cumulative density of the women's reservation wage distribution up to the market wage. Since the slope of labor force participation rates with respect to the market wage is simply the reservation wage density function, the responsiveness of participation rates to wages depends directly on the tightness of the distribution of home wages.

Ben-Porath argued that the strong serial correlation observed over time in the probability that a given woman works is strong evidence that this second view is a much closer description of reality.

The insights of Ben-Porath and Lewis set the stage for the important work of Gronau. In a series of papers (1973a, 1973b, 1974), Gronau developed more formally the interpretation of the participation equation suggested by Ben-Porath and Lewis. The probability of participation was seen as depending upon whether offered market wages exceed reservation wages where

$$W_m = \alpha X + u_1$$

$$W_r = \beta Y + u_2$$

so that the probability of participation can be written as

$$\frac{\alpha X - \beta Y}{\sigma_p} > \frac{u_2 - u_1}{\sigma_p} = \frac{u_p}{\sigma_p}.$$

X and Y are the systematic determinants and u_1 and u_2 are normally distributed errors in the market and reservation wages equations, respectively. Since the participation error (u_p) is distributed as a standardized normal, probit estimation of this probability was appropriate. Although participation probabilities had been estimated before the work of Gronau, the interpretation placed on the coefficients was now quite different. For example, instead of using the wage coefficient to retrieve income and substitution parameters, it was now an estimate of $1/\sigma_p$ (the inverse of standard deviation of errors about the probability of participation). Gronau's initial interest was in deriving the unobserved potential market wage and home wage (the value of time) for nonworkers. However, without some additional strong a priori assumptions, he could not identify all the unknown parameters necessary to calculate the value of time with the participation equation alone. Since we observe only the wage of workers (the wage conditional on $W_m > W_r$), it was impossible to identify the full set of unobserved means and variances of the market and reservation wages for women. The assumptions Gronau invoked to estimate the value of time were indeed heroic. He proceeded by examining extreme cases in which the system was identified. First, he assumed that all women within a given education and age group expect the same wage, i.e., that there exists zero variance in residual market wage offers. Thus, differences in the participation of women within that group represent *only* differences in their reservation wages. One can then calculate the mean and standard deviation of the reservation wage distributions that would result in the observed labor force participation rate and mean wage of working women. The second alternative he considered was zero dispersion in home wages, so that labor force participation rates mirror only variation in market wages. By a similar procedure, he used the observed participation rate and wages of workers to deduce the implied mean and dispersion in market wages.

In closely related work, Gronau raised the issue of censoring or selectivity bias in the wage generating function for women. The difficulties encountered in estimating labor supply and wage equations are closely linked because identical sample censoring issues arise. The expected wage for workers may be written as

$$E(W_m) = \alpha X + E\left(u_1 \,\middle|\, \frac{\alpha X - \beta Y}{\sigma_p} > \frac{u_2 - u_1}{\sigma_p}\right).$$

If one estimates a wage equation using samples of working women, biases result because the same sets of variables that determine wages enter in as a criterion for sample eligibility. The estimated wage function confounds the true behavioral wage function with the rules for sample inclusion. The issue is illustrated in Figure I.1

For simplicity, let education be the sole variable determining market wages and assume further that education does not affect reservation wages. The solid line in Figure I.1 represents the true relation between education and market wages. However, by restricting the analysis to workers, our sample includes only women in the shaded area. In the

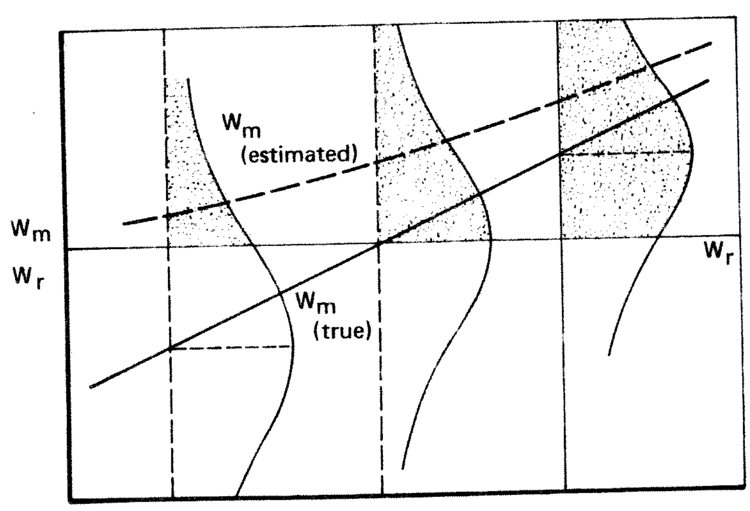

FIGURE I.1 Illustration of selectivity bias case illustrated above, the rate of return to education will be underestimated (the dotted line) because of the negative correlation introduced over the censored sample between education and the true residual in the wage equation. As Gronau pointed out, in making comparisons between groups of people who differ in their labor force participation rates (i.e., men and women, black females and white females), some of the observed differences may have little to do with behavioral dissimilarities but merely reflect sample censoring.

James Heckman's "Shadow Prices, Market Wages, and Labor Supply" (1974b) was a seminal methodological contribution in labor supply estimation. It extended the work of Gronau in a number of dimensions. First, it relaxed some of Gronau's restrictive assumptions by allowing for residual variation in both market and reservation wages and for correlation between their residuals. More importantly, it integrated into one consistent framework decisions regarding wages, hours, and participation. One limitation of Gronau's framework was that he focused on individual equations in the system in isolation—the probability of participation or the wage equation. Furthermore, he ignored completely the information available from the extent of hours worked among participants. Heckman's approach allows one to estimate a common set of parameters which underly the function determining the probability that a woman works, her hours of work, her observed wage rate, and her shadow wage. His conceptual framework rests on two behavioral relations—the market wage and the shadow wage functions. The shadow wage is defined as the marginal value placed on a woman's leisure evaluated at each unit of leisure (and hence hours of work). Heckman's model can be written as

$$W_s = b_1 H + BY + u_2$$
$$W_m = \alpha X + u_1.$$

A woman works if $W_m > W_{s|h=0}$ (her market wage exceeds her shadow wage at zero hours of work). For working women, shadow and market wages are equated at the margin so that one can solve for the hours equation.

$$H = \frac{1}{b_1}(W_m - BY) - \frac{u_2}{b_1} \quad \text{if} \quad \alpha X - BY > u_2 - u_1.$$
$$H = 0 \quad \text{if} \quad \alpha X - BY < u_2 - u_1.$$
$$W_m = \alpha X + u_1 \quad \text{if} \quad \alpha X - BY > u_2 - u_1.$$
$$W_m \text{ unobserved if } \alpha X - BY < u_2 - u_1.$$

Heckman writes out the likelihood function for the observed sample distribution of hours and wages for K workers and $N-K$ nonworkers as

$$\prod^{K} \frac{P(H, W_m)}{P(W_m > W_{s|h=0})} P(W_m > W_{s|h=0}) \prod^{N-K} P(W_m < W_{s|h=0}),$$

where $P(H, W_m)$ is the unconditional distribution of hours and wages and $P(W_m > W_{s|h=0})$ and $P(W_m < W_{s|h=0})$ are the probabilities of observing a woman working and not working. By differentiating this likelihood function with respect to all the unknowns $(\alpha, B, \sigma_1, \sigma_2, \sigma_{12})$, Heckman solves for the complete set of parameters, using nonlinear estimation. Although Heckman's work represents a significant methodological advance, it was computationally quite expensive and his technique was, therefore, never adopted in applied research.

These articles by Gronau and Heckman set the stage for that component of our research concerned primarily with econometric methodology. Because they were written prior to the other essays that treat statistical issues, the papers by Paul Schultz and John Cogan are placed first in this volume. This arrangement was also chosen because their papers contain a simple exposition and critique of the conventional methods of estimating labor supply equations.

Paul Schultz's "Estimating Labor Supply Functions for Married Women" summarizes informally many of the problems inherent in empirical work on labor supply. He includes a comprehensive discussion of the difficulties associated with the wage variable in particular. In addition to the censoring issue, biases may result from measurement error, simultaneity if hours and wages are determined jointly with other economic choices, the confounding and perhaps separate influences of transitory and permanent wage variation, and nonlinearities due to income tax rates. Schultz argues that the appropriate estimation method depends in part on the relative importance of these alternative sources of bias. For example, if measurement error in wages is the dominant factor, creating an instrumental variable for the wage rate using a sample of working women would not be a bad strategy to pursue. Since the purged residual variation about the wage equation would in this instance have no relation to the decision to participate in the market, sample censoring could be safely ignored. Because the other essays concentrate on the selectivity question, Shultz's cataloging of other perhaps equally important problems should serve as a useful caveat.

Using the 1967 Survey of Economic Opportunity, Schultz estimates labor supply functions for samples of married spouse-present women stratified into age groups and race. These functions are estimated by several alternative econometric techniques identifying, it is hoped, situations in which alternative methods yield basically identical estimates and those cases in which considerable care must be exercised in choosing among them. First, a labor supply function is estimated across all women using both Tobit and OLS and an instrument for the market wage calculated from the sample of working women. The second strategy consists in running a separate participation equation (using an OLS linear probability and a logistic model) and an OLS hours function over a sample of workers. Finally a reduced form Tobit for all women is estimated abandoning the possibility of retrieving any structural parameters.

In his paper, "Married Women's Labor Supply: A Comparison of Alternative Estimation Procedures," Cogan investigates more formally the biases inherent in some conventional labor supply functions. He deals first with the two traditional methods: (1) estimating over samples of working women, and (2) using the full sample of women but imputing a wage to nonworkers from a wage equation using the sample of workers. Cogan illustrates with great clarity the reasons why each approach is flawed. In so doing, he also clarifies the assumptions necessary to justify the use of either. He next critiques in detail the econometric methodology underlying Gronau's models, making explicit the statistical assumptions necessary to estimate the value of home time. He shows that if either of the two extremes considered by Gronau were, in fact, appropriate, then one of the traditional methods of estimating labor supply functions would also be correct. Since the traditional methods provide more information than Gronau's procedures, the usefulness of Gronau's techniques for applied empirical work is questionable.

Finally, Cogan compares estimates obtained with the two traditional methods with those reported by Heckman based on his maximum likelihood estimation. The 1967 National Longitudinal Survey is used in his empirical work. Systematic differences in the estimated parameters are substantial between Heckman's approach and the procedure of using only the subsample of working women. Systematic, but much smaller, differences are found between Heckman's method and the alternative of imputing wages to all women and estimating the hours equation over the complete sample of women. If Heckman's estimates are taken as the benchmark, the direction of the bias with the two traditional models is consistent with a priori expectations.

Even after the distinction is made between labor force participation and the extent of work-given participation, several measures of labor supply exist and have been used in empirical work—hours per week, weeks per year, annual hours. Labor supply can be specified over any time interval and within that interval many indices of the extent of work are possible. Theory provides little guidance about the "correct" supply concept to choose or about any relationships and constraints that may exist among them. From the perspective of preferences, the conventional model is couched in terms of aggregrate leisure time. An individual selects some total amount of leisure and market goods but the distribution of these leisure hours throughout the year is left open. But people do care for a variety of reasons about the timing of their leisure consumption. Leisure hours are clustered during particular days in a week and spread out during the weeks within a year. This problem of alternative definitions of labor supply is the subject of Giora Hanoch's "Hours and Weeks in a Theory of Labor Supply." Most of the work differentiating labor supply definitions has relied on fixed costs of work. Hanoch's model takes a very different perspective by assuming that within preference functions, individuals are not indifferent between the types of leisure they consume and when it is consumed.

Hanoch develops a theoretical model in which the nonperfect substitution between different types of leisure is explicit. He permits an individual to choose between leisure in a week that he does work and leisure in a week that he does not work. Based on this formulation, Hanoch derives a simultaneous multivariate model of wages, labor force participation, and annual hours and weeks worked. In particular, a consistent set of labor supply functions for both weeks worked and annual hours is derived and the relationship between them clarified. Hanoch shows that within his framework, using hours worked per week as a supply definition would be inappropriate, since it represents the ratio of two demand functions and permits no clear interpretation.

Perhaps Hanoch's real contribution is that his model possesses the flexibility to analyze a number of distinctions between units of leisure that the individual does not regard as perfect substitutes. The specific form adopted was clearly geared to justifying an empirical model that had both annual hours and weeks worked as potential dependent variables. But researchers interested in other sets of definitions of the extent of work can easily adapt the model to their purposes.

Until recently, economists viewed labor-leisure choice in a single-period framework. While this was clearly suited to explaining lifetime measures of average market participation in terms of long-run or permanent values of wealth and wage rates, individuals are faced with temporal variations in wage rates and other variables that could elicit timing responses about long-run desired levels. Moreover, the data indicated a clear age pattern to work time for both men and women. Male work age profiles have an inverted U shape with a peak in market hours occurring around age 45. The profiles for women are less easy to characterize, but they seem to be double-peaked with two periods of declining participation—one during childbearing and the other immediately before retirement. Ghez and Becker (1975) and Heckman (1971) formulated the basic theoretical structure for analyzing life-cycle timing patterns of market work. Their models explain timing decisions in terms of differences between rates of interest and time preference and life cycle movements in wage rates. Most recent research dealing with life-cycle labor supply has been predominantly theoretical, employing the techniques of optimal control to integrate a wider set of choice variables into life-cycle decision-making. Although the primary emphasis in this work concerns human capital investment patterns, the corresponding optimal age-consumption patterns of market goods and leisure have also been derived (Blinder and Weiss, 1975; Heckman, 1974c). Life-cycle models have also been used with some success to interpret empirical results of the income maintenance experiments (Metcalf, 1974; Smith, 1975) and retirement decisions, especially those connected with social security. Since the experiments were all limited in duration, they clearly induced timing behavior that contaminated their usefulness in predicting the potential impact of a permanent income maintenance plan.

The only essay in this volume that deals explicitly with life-cycle issues is James Smith's "Assets and Labor Supply." In empirical work, researchers have used variables measuring nonlabor income or assets to estimate pure wealth effects. The empirical difficulty was that estimated coefficients on assets were typically positive or, if negative, so small in absolute value that compensated labor supply elasticities were often negative. This problem also plagues many of the other essays in this volume (Cogan, Hanoch, Heckman). Using a life-cycle framework, Smith examines in detail the expected relationship between the observed asset levels of families and their labor supply. He argues that, to a large extent, assets at any age represent accumulated past savings and, hence, reflect desired consumption and income profiles. Smith demonstrates that economic theory suggests little of an unambiguous nature about the expected relationship between assets and labor supply. Rather, they are both simultaneously determined by similar economic variables over the life cycle, so that the empirical relationship should not be interpreted as reflecting a causal sequence from assets to market work.

In his empirical work, Smith uses the 1967 Survey of Economic Opportunity which contains information on assets, wages, and labor supply of family members. The implications of the life-cycle model for the savings and labor supply of each family member are tested. Although the quality of the asset data in the SEO is not ideal, his empirical results basically support the predictions of the theory.

In June 1975, James Heckman circulated to the Rand staff a five-page draft note on sample censoring that forms the nucleus of his paper, "Sample Selection Bias as a Specification Error." Building on the work of Amemiya (1973b) Heckman advanced the simple, neat view of censoring bias as another variant of specification error. Using well-known theorems involving truncated normal distributions he proves that the expected wage and hours for a sample of workers can be written as

$$E(W_m) = \alpha X + \frac{\sigma_{1p}}{\sigma_p}\lambda$$

$$E(h) = B_0(\alpha X - BY) + \sigma_p\lambda,$$

where λ is the ratio of height of the density to the right tail area, σ_{1p} is the correlation between the errors in the participation and wage equations, and σ_p is the standard deviation of the participation error. Therefore, the bias in estimating wage and hours over samples of working women results from the omission of the variable λ. If one knew or could estimate λ, then it would be possible to obtain consistent wage and hours equations with samples of workers. Heckman suggests estimating λ from a first-stage probit on the probability of participation. His formulation of the problem also clarifies the nature of the biases involved in using samples of working women. The bias depends on the sign of λ in the relevant equation as well as the sign of the coefficients in the auxilliary regression between λ and the variables in X and Y. Since λ is a negative function of the probability of participation, the issue reduces to whether variables in the X and Y vectors increase or decrease participation. The real value of this technique is its simplicity. The censoring correction can be estimated with inexpensive statistical programs easily available to applied researchers. Heckman's model clearly can be applied to a wider scope of censoring problems than just labor supply. It has already been used in the analysis of migration, unionism, educational choice, and the effects of welfare programs. For his empirical work, Heckman uses a sample from the 1967 National Longitudinal Survey for Mature Women. In interpreting his results, he proposed a modification of the labor supply function as an alternative to the fixed costs model of Hanoch and Cogan. Heckman argues that the issue is not so much fixed costs, but the limited availability of low-cost child care. These are conceptually distinct models and their implications are contrasted.

Giora Hanoch's "A Multivariate Model of Labor Supply: Methodology and Estimation," is a tour de force in dealing with a wide variety of estimation problems. In addition to incorporating Heckman's correction for participation selectivity bias, he considers a number of issues previously ignored or not included in a fully integrated system. First, based on his theoretical model in "Hours and Weeks in a Theory of Labor Supply," Hanoch's econometric method allows for the simultaneous joint determination of the supply of annual hours and weeks of work. The common practice in empirical research had been to treat these separately as two alternative, but unrelated, definitions of labor supply. The most fundamental contribution of this paper involves the impact of fixed costs of work. Since fixed time or money costs of work will impart a discontinuity in the supply function, the reservation wage at which an individual enters the market occurs at a positive quantity of working hours. If fixed costs are important, estimates of labor supply functions that ignore this discontinuity would seriously overestimate the measured wage elasticity of labor supply. In spite of the emphasis given to the zero truncation of working hours, there are probably corner limits that are as pervasive. One obvious restriction is that individuals cannot work more than 52 weeks per year. This 52 week limit imposes constraints on the responsiveness of the supply of weeks and annual hours as individuals approach the point of full-time work. Hanoch suggests methods that deal with this upper truncation in the estimation of labor supply. Finally, he considers the survey week selectivity bias that is ignored in all other research. In many data sets, information on hours worked during a week is only available for individuals who worked during the survey week. Because of this, analysts have restricted their empirical studies to those who have positive weeks and positive hours of work in the survey week. But survey week workers are not a random sample of all workers. If the survey week is a random week during the year, the people most likely to be found working are those with the largest weeks worked in the previous year. Hence, the probability of working during the survey week should be proportional to the fraction of weeks in a year that one works. Hanoch develops a statistical technique that controls for this bias. Instead of estimating his complex model with fully efficient, but expensive, methods, Hanoch outlines a step-by-step estimation procedure that is less costly to execute and quite feasible with existing computer programs. His model is estimated using a subsample of white married women from the 1967 Survey of Economic Opportunity.

The effect of time and money costs on labor supply is treated in greater depth in John Cogan's "Labor Supply with Costs of Labor Market Entry." The standard theoretical treatments of these costs analyze separately the impact of exogenous fixed time and money costs. Both time and money costs impart a discontinuity to the labor supply function, but they have different predictions about the extent of work among participants. Fixed money costs of work increase the hours worked of participants while fixed time costs reduce observed hours of workers. More realistically, Cogan contends that individuals have the potential to trade off time and money costs and that this choice depends in part on their earning capacity. The decision of where to live relative to place of work, the type and form of child-care arrangements, and the mode of transportation to work all indicate that such substitution possibilities exist. Cogan constructs a theoretical model in which people are confronted not with rigidly determined time or money costs, but rather with a locus of points representing all possible combinations of time and money costs of work. An individual chooses not only his desired bundle of leisure and market goods, but he must also minimize the cost of engaging in work. The comparative statics of this model are developed and the close link that exists between the labor supply function and the function describing the trade-off between time and money costs of work is illustrated. Cogan outlines a statistical method for estimating a labor supply function with endogenous time and money costs. The data used in Cogan's empirical analysis is taken from the 1976 Survey of the Michigan Panel of Income Dynamics.

The large quantity of empirical work presented in the volume precludes any attempt to summarize it completely. Instead, I will highlight those results relating to the two behavioral functions that form the core of our research—the wage and labor supply equations. The hope is to identify the major similarities and differences emerging from the separate papers.

Four of the papers offer evidence on the importance of selectivity bias in female wage functions. Cogan (Chapter 2) contrasts OLS wage regressions with those obtained with Heckman's maximum likelihood procedure, while Heckman (Chapter 5), Hanoch (Chapter 6), and Cogan (Chapter 7) employ the censoring correction proposed by Heckman (Chapter 5). Although each study adopts a different specification of the wage function using three separate data sets, a strikingly consistent story emerges. Any evaluation of the impact of sample censoring depends partly on the underlying reason for estimating a wage function. One may be interested primarily in predicting wages for women, in retrieving structural parameters in the wage functions, or simply in obtaining an instrument to be used in another equation (i.e., labor supply). The general consensus appears to be that sample censoring is a real but not overwhelming issue in the estimation of female wage functions. The estimated coefficients on λ are nontrivial (from .087 to .1282), but only in Hanoch's essay is the statistical significance impressive. All four studies report that OLS overpredicts wages of all women by a moderate amount. Evaluated at the mean characteristics in the respective samples, the OLS point estimates exceed the true average wage offer for women by 5.3 to 11.1 percent.[1] However, since OLS predicts the same wage for women independent of their current labor force status, this bias becomes considerably more severe when

[1] In the individual studies, the OLS wage prediction exceeds the mean wage offer by 5.3 percent (Cogan, Chapter 7), 5.5 percent (Heckman), 7.8 percent (Cogan, Chapter 2), and 11.1 percent (Hanoch).

imputing wages to nonworking women. For example, the studies by Heckman and Hanoch imply that OLS overstates wages for nonparticipants by over 20 percent. If wage equations are not corrected for censoring bias, the potential wages of nonworking women will apparently be substantially exaggerated.

Cogan and Heckman present OLS equations alongside the censored adjusted wage regressions. These comparisons, which are summarized in Table I.1, can be used to detect the magnitude of the sample-selection bias on individual parameters. The effect of sample censoring on the education coefficient appears to be very minor, producing only a slight underestimate. Ironically, it was partly the concern over a potential bias in the return to schooling that motivated the original research on censoring. In contrast, the coefficient on market experience does differ markedly. The OLS estimates of the return to market experience range from 15 percent (Cogan, Chapter 7) to 29 percent (Cogan, Chapter 2) below the true parameter, with Heckman's 22

TABLE I.1: Comparison of Female Wage Equations
(standard errors in parentheses below coefficients)

	Cogan, Chapter 2		*Heckman, Chapter 5*		*Cogan,[a] Chapter 7*	
	Corrected[b] for Censoring	OLS	Corrected for Censoring	OLS	Corrected for Censoring	OLS
Education	.0879 (.0074)	.0854 (.0073)	.0779 (.0057)	.0763 (.0056)	.096 (.007)	.092 (.007)
Experience	.0259 (.0046)	.0184 (.0030)	.0207 (.0035)	.0167 (.0021)	.251[c] (.042)	.213 (.026)
Constant	−.8761 (.206)	−.7176 (.191)	−.515	−.401	−.212 (.136)	−.113 (.105)
λ				.0878 (.059)	.088 (.076)	
Data Source	1967 National Longitudinal Survey		1967 National Longitudinal Survey		Michigan Income Dynamics	

[a] Cogan also included variables for southwest residence and age which were not altered by the selectivity correction.
[b] Based on Heckman's maximum likelihood procedure developed in Heckman (1974b).
[c] Cogan's variable is the log of market experience.

percent in an intermediate position. Apparently, selectivity bias produces a serious underestimate of the true impact of market experience on female wages. This sensitivity of the experience variable is not surprising in view of the high correlation of past market experience with the current participation decision. Finally, OLS systematically overestimates the intercept in all three studies. The difference in intercepts alone implies that OLS overstates the mean wage offer by 10 to 20 percent.

The only study that offers any evidence of the use of an OLS instrumental wage is Cogan's (Chapter 2). Cogan estimates a labor supply function for all women, using an instrumental wage from an OLS regression. The resulting measured wage elasticity exceeds the Heckman maximum likelihood estimate by 25 percent. This was expected, since OLS tends to overstate wages more for low-wage women thereby compressing the wage differences among women. However, none of the coefficients on the other variables in the labor supply function appear to be affected by the use of the wage instrument. Additional experimentation will be necessary before any final assessment on the use of the instrumental wage can be obtained.

While the similarity among these studies is encouraging, one should be cautious before extrapolating our results to other data sets and demographic groups. Because of the nonlinear nature of Heckman's censoring technique, wage equations estimated from other samples with different average economic and demographic characteristics may be quite different from those included in this volume. Moreover, the censoring correction is derived from the behavioral participation function. Schultz's evidence in Chapter 1 suggests that participation equations differ significantly across race and age groups. In view of this, it would not be surprising if future research provides a different view of the effect of censored samples on female wage functions.[2]

In contrast to the general uniformity of estimates for the wage equation, there is much more disagreement among the essays on the appropriate labor supply function. This partially reflects a dispute concerning the appropriate structural labor supply model, and also is a result of some sticky statistical problems in identification. The main theoretical contention concerns the possible existence of a discontinuity in the labor supply function due to fixed costs, as argued by

[2] For example, Smith (1979) reports a negative coefficient for λ in a sample of black women.

Cogan and Hanoch. I will summarize the results in the three main areas that deal with labor supply functions: the choice of linear and nonlinear methods of estimation, the importance of selectivity bias, and the existence of fixed costs.

Schultz reports experiments contrasting simple linear (OLS) and nonlinear maximum likelihood estimates. The Tobit expected value locus and the OLS supply function (estimated over all women) produce quite similar results at the sample means, indicating that OLS is a good linear approximation to Tobit. As one might suspect, they diverge significantly as we move away from the sample means. The conceptual advantage of Tobit is that one can also derive the index function. Although the expected value locus by incorporating the zero truncation of hours is the aggregrate labor supply function, the coefficients of the Tobit index reflect the parameters of the individual's utility function, in particular income and substituting effects. For some purposes, we are more interested in testing hypotheses about these utility parameters. Similarly, Schultz's OLS and logistic participation equations are almost identical when evaluated at the means. Apparently, these computationally less expensive linear methods provide useful guidance to the applied researcher in searching for an appropriate model specification, at least in situations where the data are not bunched at one of the extremes (mostly zeros or ones). The advantages of the nonlinear Tobit and logit (or probit) are mainly increased reliability of predictions for demographic groups that are not representative of the sample means. This general conclusion appears to be consistent with applied research reported in areas other than labor supply.

To fix ideas on selectivity bias and fixed costs, Figures I.2 and I.3 present labor supply functions with and without a discontinuity. In Figure I.2, the true supply function for the individual is the index line. However, because of the truncation of hours at zero (the decision of some women not to participate in the labor force), the expected number of working hours per female can be obtained from the expected value locus. This is the labor supply function in the sense that it measures at each wage the average amount of work effort per woman. Finally, an OLS labor supply function for workers will approximate the means of these distributions conditional on positive hours of work. The distance between the OLS function for workers and the index line measures the extent of selectivity bias at each wage. In

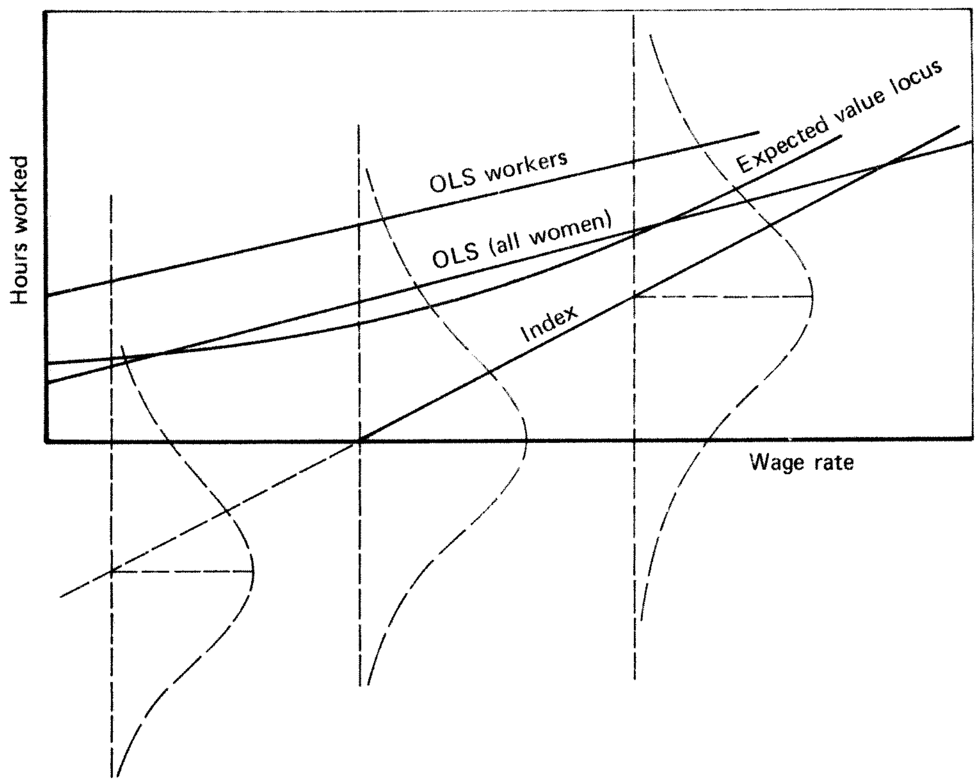

FIGURE I.2 Labor supply function without discontinuity

Figure I.3, the labor supply model with fixed costs is graphed. Because fixed costs eliminate the close link between the parameters of the participation and hours function (the parameters are proportional without fixed costs), the point at which the reservation wage equals the market wage can occur at a positive quantity of hours. If the discontinuity is large, the zero truncation will in fact be relatively unimportant for workers, and the slope of an OLS function over workers will closely approximate the index function. Thus the importance of selectivity may be severely diminished when the discontinuity is sufficiently large. For illustrative purposes, I have also placed in Figure I.3 the estimated index function that results from ignoring the discontinuity (and hence imposes an inappropriate constraint). Clearly this function can be considerably more elastic than the true index.

The papers by Schultz and Cogan (Chapter 2) were written prior to the work on fixed costs so they, of course, do not test the validity of the constraints imposed by ignoring a possible discontinuity in the labor supply function. Within the context of a no fixed costs model

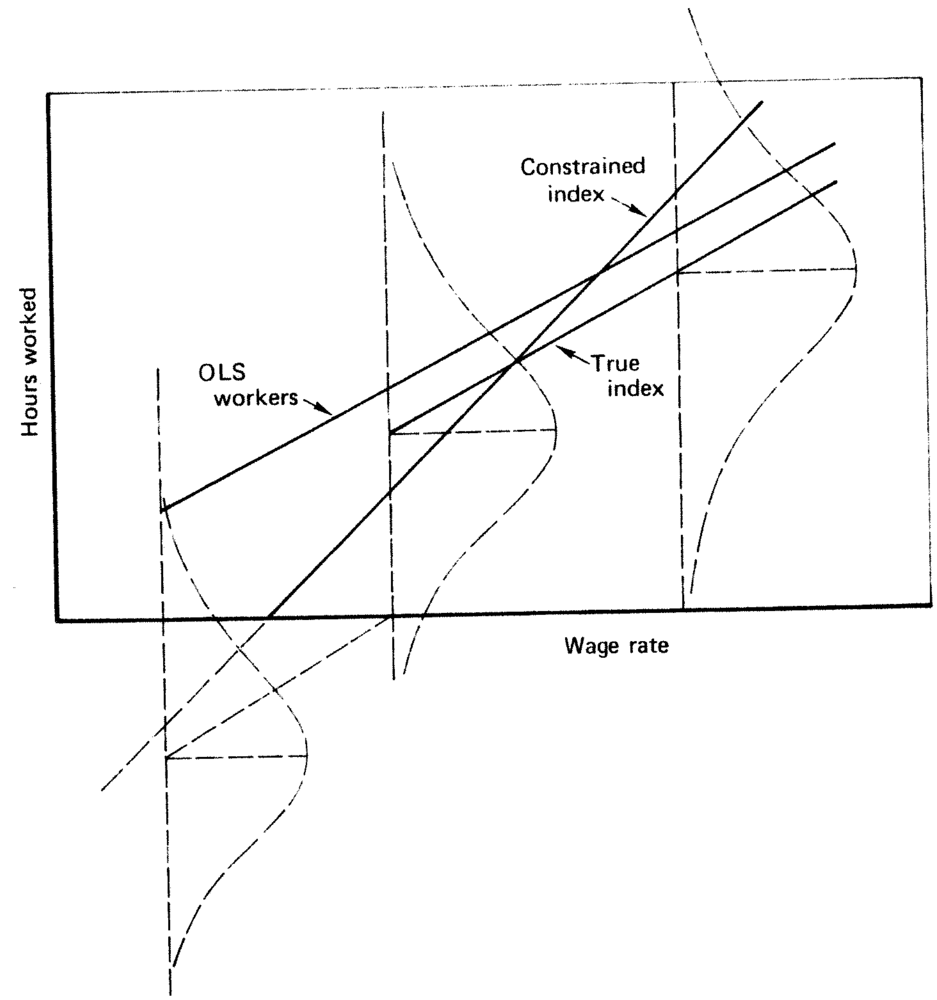

FIGURE I.3 Labor supply function with discontinuity

(Figure I.2), both studies provide evidence on the appropriateness of the two conventional methods of estimating labor supply functions—OLS over a sample of workers, and by the full sample of women with an instrumental OLS wage. Their OLS labor supply equations estimated over the sample of working women have substantially smaller wage elasticities than corresponding functions using the full sample of women. As is evident from Figure I.2 (where the OLS line is less steep), this is consistent with an important role for selectivity bias in the labor supply function. Cogan also reports that the OLS function on workers underpredicts reservation wages (which Cogan estimates as often being negative), so that nearly every woman is predicted to be a market participant. When an OLS instrumental wage is used in a Tobit function estimated over all women, Cogan's estimates indicate the true labor supply wage elasticity is overestimated. Intuitively we are compressing the horizontal (wage) axis by the use of this instrument and overstating the labor supply elasticity. Based on these two studies, it is clear that when we do not allow for a discontinuity in the labor supply function, the data strongly support the presence of a significant censoring bias in labor supply functions. Conventional methods of estimating labor supply relations will seriously bias the true response.

Unfortunately, the story becomes more muddled when the possibility of a discontinuity is considered. The three essays that allow both for selectivity bias and fixed costs are Heckman, Hanoch (Chapter 6), and Cogan (Chapter 7). The divergence between the estimates reflects the trade-off in the data between a discontinuous labor supply function and sample censoring. Heckman finds little support for the fixed costs model, but an important role for selectivity bias in labor supply functions. In contrast, the discontinuity in Cogan's labor supply equation is quite large (1,151 hours) so that for women who work, the zero truncation of hours is unimportant. Because of this, an OLS function estimated over workers may closely approximate the individual labor supply function. Hanoch reports evidence in favor of both fixed costs and selectivity. The minimum entry hours estimated by him are smaller (870 annual hours), allowing for an independent role for censoring. Both Hanoch and Cogan report relatively small wage elasticities in an individual's supply function, indicating that much of the adjustment in hours worked occurs through the participation response. Particularly intriguing in this regard is that, according to Hanoch, at high female wages when women work full-time work years, the labor supply function becomes backward bending resembling those estimated for males. In contrast, the wage response in the individual labor supply function estimated by Heckman is much larger. Heckman argues for an alternative definition of hours worked based on dividing earnings by the wage rate. This definition apparently fills out the hours density at low wages and partially accounts for the absence of a discontinuity in his paper.

Even if there were no disagreement about the correct structural labor supply model, the statistical problems in identifying labor supply functions alone would be severe. The major defect is that we have no strong a priori information for distinguishing between factors affecting participation decisions and those that affect hours, given participation.

As such, the estimated λ normally includes the same variables as those included directly in the individual labor supply function. Although λ is not ordinarily a linear function of these variables, the potential for collinearity is there. In practice, this has evidently presented estimation problems, and thus the inability of the data to enable us to discriminate clearly between alternative structural labor supply models may not be surprising.

With a problem as complex and far-reaching as labor supply and wages, the final chapter will never be written. Although we believe that these essays may have advanced the state of the art, many issues are unresolved and puzzles abound. Even within these essays, there is disagreement about the relative importance of sample censoring and fixed costs. Many fundamental methodological questions relating to experimental design and estimation of the data from the income maintenance experiments remain unanswered and are not even touched upon in this volume. The sensitivity of some of the estimation techniques advanced here to alternative specifications of the model is open to serious scholarly question. A limitation in all this work is that we do not take advantage of the longitudinal information available from many recent panel data sets. Panel data have considerable potential for addressing issues that are impossible to analyze with individual cross-sectional data sets. The question of persistent and transitory components of labor force participation over time and the estimation of life-cycle models using panel data require major statistical and theoretical innovations. But all that must remain the subject of future work.

ACKNOWLEDGMENTS

I am indebted to The Rand Corporation for supporting the rewriting and editing necessary to turn the original manuscripts into a book.