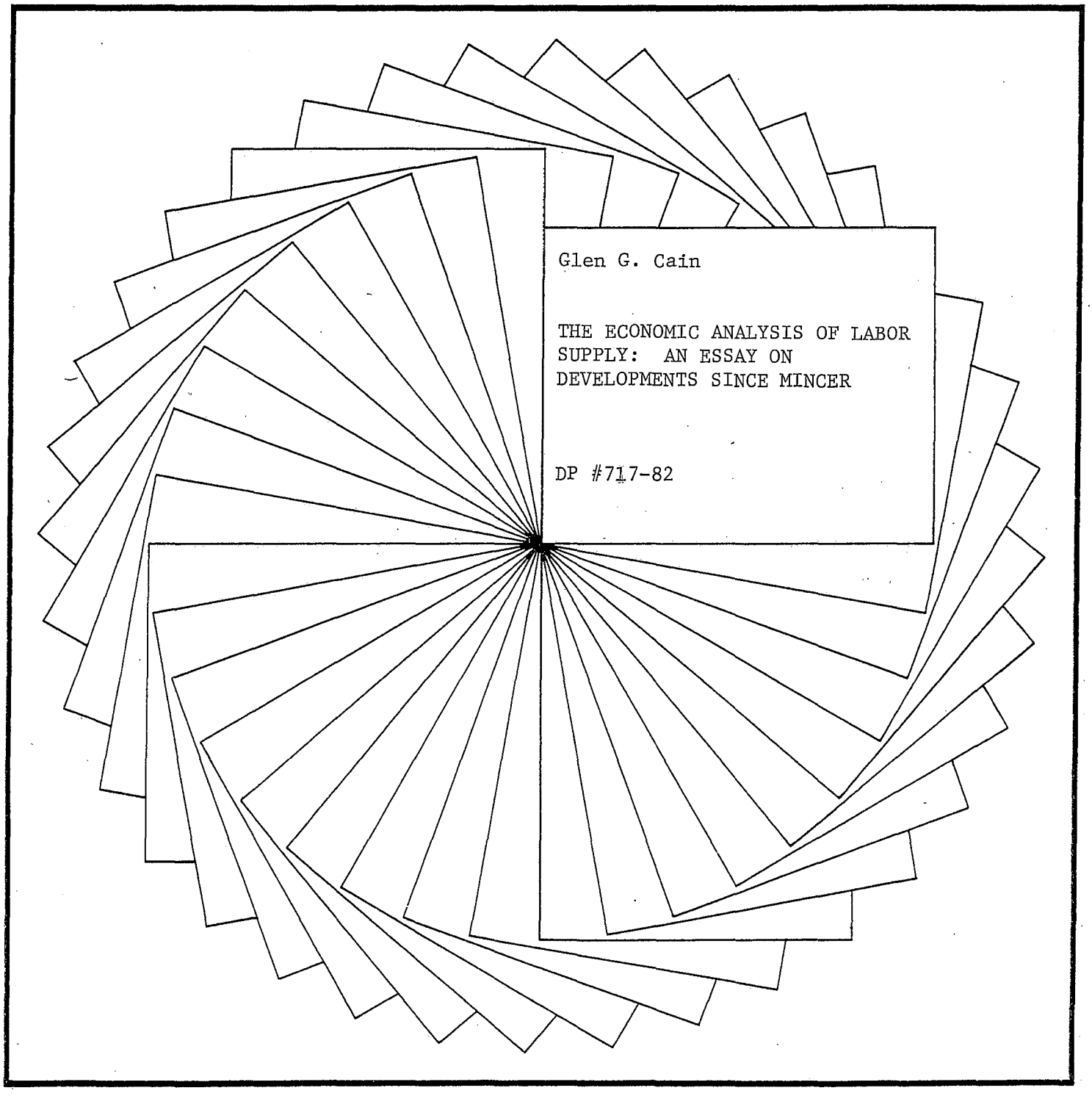
University of Wisconsin-Madison

IRP Discussion Papers



The Economic Analysis of Labor Supply:

An Essay on Developments since Mincer

Glen G. Cain
Institute for Research on Poverty and
Department of Economics
University of Wisconsin-Madison

January 1983

This paper was presented at a conference on labor supply, sponsored by the Social Science Research Council in December 1982. The research reported here was supported in part by the National Committee for Research on the 1980 Census and in part by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Department of Health and Human Services pursuant to the provisions of the Economic Opportunity Act of 1964.

ABSTRACT

Developments in the economic analysis of labor supply during the past 20 years are organized around five themes, and studies of the labor supply and wages of married women are used to illustrate these themes: (1) household decision-making theory; (2) empirical estimation; (3) the practical and policy uses of the research; (4) descriptive statistics; and (5) normative issues. The main methodological message is that a careful specification of the practical purposes of one's analysis is a necessary first step for making inferences about some population of interest. This step will motivate a closer examination of the sources of sample variation in the variables and of whether these correspond to the sources of interest in the population.

The Economic Analysis of Labor Supply: An Essay on Developments since Mincer

1. BACKGROUND

Three main themes may serve to organize a survey of labor supply research during the past twenty years--or, to mark a specific year, since 1961, when Jacob Mincer delivered his paper, "The Labor Force Participation of Married Women."[1] The three themes are the theory of labor supply, the empirical estimation of labor supply functions, and the practical, often policy-related, problems that sometimes motivate our research. Each theme appeared with clarity in that famous paper.

1. First, Mincer cast the theory of labor supply in its modern mold with three formulations. The household, or family, was specified as the appropriate unit of analysis, wherein income was pooled and a division of labor between home and market production occurred. This division of labor directed attention to two wage rates, home and market, for each family member. Second, an explicit formulation of income and substitution effects in terms of empirical variables was specified. Although it was not new to treat labor supply functions as the obverse of leisure-demand functions, and, therefore, in terms of consumer demand theory, Mincer's focus on married women enabled him to estimate separate income and substitution effects in labor supply models for the first time. Third, Mincer applied Milton Friedman's concept of permanent and transitory income effects (on consumption) to labor supply decisions for secondary earners in families, wives being the main example. Each of these theoretical formulations--household decisions, the specification of

income and substitution effects, and the issue of the timing of labor supply responses--has received much attention in the last twenty years. New theoretical ideas have also been added, some by Mincer in subsequent work. These will be discussed below.

2. Mincer's second contribution was his empirical estimates. These estimates were unusually interesting and useful because they dealt with a variety of data sets and appeared to be consistent across sets and consistent with the implications of the theory. The labor supply functions for married women showed negative income effects, which were consistent with leisure as a normal good, and positive substitution effects, as required by economic theories of rational behavior. Moreover, the substitution effects were larger than income effects in absolute value, which appeared consistent with the long-run rise in labor supply of wives during a period when incomes and wages rose at roughly parallel rates.

Mincer's rationalization for the larger substitution effect for wives compared to husbands was straightforward: Housework provides an alternative to market work for wives that is culturally sanctioned and, concerning fertility, biologically determined. Whereas men are assumed to have only leisure as an alternative to market work, housework allows wives more options, and this leads to a greater responsiveness (elasticity) by wives to the market wage rate. The movement for women's liberation has weakened but not yet overturned this argument.

We should not overlook Mincer's use of the unemployment rate and the wife's education and fertility as control variables along with the income and wage variables in the empirical model. Each control variable was included in the model with a theoretical, or maybe a commonsense, justification, and the estimated effect of each was interesting in its own right.

The last twenty years have witnessed a large number of empirical studies of labor supply functions, which have focused on income and substitution parameters, and enough surveys of these have been written to justify a survey of the surveys. However, this paper will not be that.

3. The third contribution of Mincer's 1962 paper was to address several practical, sometimes policy-related, problems that were important to the larger social science community and to the aware citizen, as well as to labor economists. First among these problems was the task of explaining and, therefore, predicting trends in labor supply. The immediate challenge was to explain the long-run rise in market work rates of women during a period when the labor supply of men had decreased sharply. A second important problem was the prediction and explanation of women's labor force participation over the business cycle. Mincer's theoretical model and empirical work were specified, either explicitly or implicitly, to deal with each of these obviously important issues. I will return to this point later.

It is worth pausing here to relate Mincer's quantitative answers to these two questions. First, approximately half of the rise in market work rates by women could be explained by a narrowly defined economic model that had income and wages as explanatory variables. This was a notable achievement in view of the facts that (a) the previously held economic empirical generalization about the labor supply function--namely, its negative slope--predicted a decline in work rates, and (b) the previous demographic explanations, which examined changes in the age, race, and marital composition, and in fertility rates, failed to explain the rise in women's labor force participation.[2]

Consider next Mincer's quantitative estimates of the response of married women's labor supply to the business cycle. He found that his empirical work predicted roughly offsetting gross flows of women (wives) into and out of the labor force over the business cycle. Thus, the predicted net response was zero, although research by Mincer and others subsequently showed a net procyclical response for the time series from the late 1940s to around 1970.[3] "Discouraged workers" outnumbered "added workers" in the terms used in this analysis. This issue has not received much attention during the last ten years, and the work that I am most familiar with shows that the dominance of the discouraged-worker effect among wives has disappeared in the last decade (Mitchell, 1980).

In summary, Mincer's contributions in his research up to the mid-1960s were theoretical, empirical, and practical. His answers to two practical problems were reasonable but called for further investigation.

My survey will take as its point of departure these three themes described above, with an emphasis on the third. What are the continuing and new practical problems that motivate labor economists to analyze labor supply? I will, however, add two themes: the role of descriptive statistics and the normative issues involved in research on labor supply.

2. A BRIEF SUMMARY OF THEORETICAL DEVELOPMENTS

Theoretical developments in the last twenty years can be grouped into the following three categories: the use of utility functions, the "endogenization" of a variety of household decisions that are related to labor supply, and attention to the timing of labor supply decisions.

1. First, the attempts to use utility theory to derive specific functional forms and to impose constraints on the labor supply function have improved the aesthetic presentation of the topic, but I do not believe it has contributed much to empirical estimation or to the practical issues being addressed. Soon after Mincer's work appeared, Marvin Kosters (1966) used a more formal theoretical analysis of demand functions to sharpen and correct the measurement of income and substitution effects according to the Slutsky formulas. Orley Ashenfelter and James Heckman (1973, 1974) made further refinements and added the constraint of equal cross-substitution elasticities between the husband's and wife's leisure demands. This constraint is based on the assumption of a single utility function for the household (read, "husband-and-wife"). I do not find this assumption appealing for labor supply decisions, about which there is often considerable disagreement and conflict between spouses. Furthermore, the constraint is usually rejected (or only weakly supported) when tested. The fact that about half of the marriages in the United States will dissolve should cast further doubt on the assumption of a single utility function, given the lifetime perspective for decisions about labor supply and human capital investments.

Finally, the message I derive from the attempts to use utility theory to impose functional forms is that the constraints are not dictated by persuasive and consistent theoretical arguments. Furthermore, the role of these constraints in saving degrees of freedom is unnecessary, because we have large samples to use for estimation. On these matters, I found useful Jonathan Dickinson's examination (1976, 1980) of the theoretical and empirical properties of the functional forms for labor supply estimation that were derived from utility functions.

2. Second, most of the variables that were assumed exogenous and used to explain women's labor supply in the "first generation" of Mincer-models became endogenous in one or more subsequent models. The income of the wife's husband and various forms of nonlabor income were endogenized by allowing his labor supply and family asset accumulation to depend on the wife's labor supply. The wife's wage rate and education were endogenized, as human capital models flourished to predict wages. Fertility, marital status, household composition, unemployment status of other family members, and geographic residence (and migration) are the other main endogenized variables. They are all "choice" variables in the household-decision model. Examples of research expressing such variables as endogenous are too numerous to list.

Among variables that are specific to the individual, only age, sex, and ethnicity are considered purely exogenous. Although a few market variables, such as unemployment rates and market prices, are also considered exogenous to the individual, these market variables tended to be neglected as aggregated (or grouped) data fell out of favor.

3. Third, several issues concerning the timing of labor supply decisions were treated theoretically and led to econometric innovations. A useful distinction was drawn between the decision of a worker to vary his or her time at work and the decision of a nonworker to enter the labor force.[4] In the short run the person may face constraints on desired time at work (see Ham, 1982). A related distinction is that between a static measure of the labor supply of the wife for a single period (whatever its length) and her period-by-period labor supply decisions. The latter includes the matter of intertemporal substitution of labor supply.[5]

The empirical estimation of a model that tries to incorporate all these developments would be intractable. The complications from these developments--or perhaps I should say, the complications in the real world that motivated these developments--have made it mandatory, in my view, to focus on the practical problems at hand: to determine which of these conceptually endogenous variables may be considered predetermined and what time period is relevant. I place a high priority on paying more attention to the third in my list of Mincer's original contributions.

3. PRACTICAL QUESTIONS AND DEVELOPMENTS IN RECENT LABOR SUPPLY RESEARCH TO DEAL WITH THESE QUESTIONS

The two questions about long-term trends and business-cycle responses that Mincer addressed remain issues for research, but they have been superseded by the following categories of practical problems.

1. Predicting the consequences of government policies that are intended to affect labor supply.

2. Explaining wage differences among men and women, with special attention to the role of sex differences in labor supply. The large number of studies about labor market discrimination against women reflects this attention.

3. Explaining demographic outcomes among women, such as marital status, fertility, and marital dissolution, with special attention to the role of labor supply and wages.

I will discuss only the first two in this paper.

A. Predicting Effects of Policy on Labor Supply

I find it useful to distinguish between the Mincer-type of estimation of "natural" market forces and the estimation of direct interventions in the market, usually by the government. In both cases we seek, generally, conditional expectations of labor supply variables; conditional on a set of putative exogenous variables. However, we may place less stringent demands on the estimation of natural market forces. For example, if we want only to predict, narrowly construed, we may relax the requirement of isolating the net causal impact of the right-hand-side variables depicting a market force (for example, wages). If we want only to explain the past, we may be satisfied with qualitative (hypothesis testing) rather than quantitative (estimation) answers. The demands we make on models used for policy analysis, in contrast, usually require estimating the net causal effect of the policy variable with reasonable precision.

A large variety of government policies affect labor supply, and their number sharply increased during the Great Society decade of the 1960s and into the 1970s as well. Macroeconomic policies to one side, the microeconomic policies may be categorized into three types: (1) those that attempt to change one's earning capacity, such as education and training programs; (2) those that affect the price of the provision of goods and services that are complementary (or substitutable) with labor, such as child-care or housing subsidies; (3) those that affect labor supply directly, in the sense of changing one's income and/or wage rate (without directly affecting one's earning capacity), such as taxes on earnings and income maintenance programs. Only income maintenance laws will be discussed in this paper, but they serve to illustrate the type of approaches and problems in labor supply research on policy effects.

Income maintenance is a term that covers a variety of particular programs, including Aid to Families with Dependent Children, Food Stamps, disability payments, social security retirement benefits, and unemployment insurance. It was, however, a proposed program, the negative income tax, that received the most analysis with respect to its potential effects on labor supply. Several well-known negative income tax experiments were part of this research, but I will not discuss these. The subject is too specialized, complicated, voluminous, and, I believe, has yet to be incorporated into the mainstream of empirical economic research.

A simple formulation of a policy model should help the discussion.[6] Let individuals be the units of observation.

y = the outcome of interest, such as some labor supply variable.

$\underline{x}$ = a general set of control variables that theory has suggested affect labor supply, where these will remain deliberately unitemized until a specific need arises.

t = the government program, which in its simplest formulation is defined as equal to 1 if the person is in the program and equal to 0 otherwise.

Using a linear and additive model for simplicity, we have

$$(1) \qquad y = \underline{x}B + At + u,$$

where u is an error term, and we seek an unbiased estimate of A.

The familiar problem is that t cannot generally be assumed to be independent of the other-than-$\underline{x}$ variables (represented in u) that have their own effect on y. In other words, we cannot assume that t has been randomly assigned, even controlling for $\underline{x}$; instead, t is endogenous to

the system of relationships and needs to be modeled. One such specification is

$$(2) \qquad t = g(\underline{x}', z) + e,$$

where g is allowed to be a nonlinear function in view of the 0, 1 limits on the values of t, $\underline{x}'$ is a subset (perhaps all) of $\underline{x}$, and z is one (or more) variable(s) (if available) that affects t but not y. The error term, e, is permitted to be correlated with u in recognition of the commonality of some omitted variables that affect t and y. The roles of the function g and of the variable z are probably well known to the reader. The nonzero correlation between e and u forces us, in the absence of certain other restrictions on the model, to rely on the variation in t induced by z and/or the nonlinear functional forms of $\underline{x}'$ to identify A in model (1).

Reliance on nonlinearities in $\underline{x}'$ to identify A is risky, in part because of the likely severe multicollinearity between $\underline{x}$ and the "instrumental" variation in t (for example, in $\hat{t}$ as predicted by (2)), and in part because the original linear specification in equation (1) was probably a simplification of an unknown true relationship. The exclusion restriction on z is often also hard to justify, although my own preference is to try to study explicitly the "selection process" associated with t in order to discover credible z variables.

With this as background, consider the intense attention given to the issue of the labor supply response to a change in income maintenance laws during the ten-year period, 1966-1976. The problem in estimation is readily revealed. Welfare status (a dummy variable) or the values of a variable describing the welfare system (0 for those not on welfare and

the positive number measuring the amount of welfare benefits for those on welfare) is the t variable. We immediately recognize that u and e contain common unobservable variables. As examples: unmeasured health variables (including mental health) that affect both t and y. Thus, estimation of A is difficult and subject to intense skepticism. Selection bias is the current term used to describe the alleged bias in the estimated A.

How does the Mincer legacy of estimating market forces, which I will refer to as the conventional model, fill the breach caused by this difficulty in estimating the program effect directly? Unsatisfactorily, is my verdict. The available strategies in using the conventional model may be ex post or a priori.

Ex post, labor supply is measured before and after the time when the policy has been implemented. The conventional model predicts what labor supply would have been in the absence of the policy. This prediction is compared to the observed change, and the difference is attributed to the policy. Clearly, this strategy depends on the assumption that no other variable(s), other than those in $\underline{x}$, has (have) changed during this time period. To be more precise, the assumption is that the change in the omitted variable(s) did not, on average, affect y during the period. Or, consider the related strategy in which several geographic units, rather than time alone, provide variation in t. In this case we have to be convinced that the geographic units--or the specific economic, political, etc., variables that distinguish those units--are not causal to both t and y. These assumptions about the ex post strategy are hard to swallow.

The a priori strategy is more common, and it has the virtue of applying to a contemplated or projected policy. First, translate the

changes in t into equivalent changes in the variables appearing in the conventional model. For example, a change in an income maintenance plan, t, is translated into changes in income, m, and the wage rate, w. Next, the parameter estimates from the conventional model are used to predict the change in y in response to the m-and-w changes induced by t.

Several problems arise. (1) Can we trust that our translation of t-variation into m-and-w variations is accurate and sufficiently complete in describing the relevant variation in t? The administration of income maintenance programs may make the actual values of m and w different from their ostensible values. Or, there may be aspects of the program that affect y in ways for which we have no parameter estimates.[7]

(2) What parameter values from among the variety available are to be used? The pessimistic answer is that the fact of the variety is reason not to trust any of them. The optimistic, and more constructive, answer is to use those conventional estimates where the sources of variation are the most similar to the variation induced by t. Note that I use the term "sources of variation," not "amount of variation." I want to ignore any problem of sampling error, so I will ignore considerations of sample size or of the amount of variation in the independent variable.

But consider the dilemma here. If we look for an application of the conventional model where the experience--that is, the variation in m and w--reflects conditions similar to t, we are back into the problem of selection bias. For example, unemployment compensation is one type of income maintenance program, and it may be similar to other types. But estimating the effect of unemployment compensation on labor supply merely illustrates the previously discussed problem of estimating t.

The other extreme is when the conventional model relies on natural market forces that have no direct connection to t-type variation. To illustrate, assume that all persons with work-conditioned income, such as welfare, unemployment compensation, disability payments, and pensions, are excluded from the sample used for estimation. Assume further that the only source of wage variation among the remaining observations is differences in human capital embodiments: innate ability and acquired abilities. And the only sources of variation in nonlabor income are dividends, rents, interest payments, and capital gains.

Without intending to be disrespectful, I submit that at this point the craft, rather than the science, of our empirical work takes over. First, we select as control variables those determinants of differences in human capital and nonlabor income, such as education and age, that we believe have their own effects on y. Second, we consider that the residual variation remaining in w and m potentially reflects the following four sources:

(a) unmeasured, long-run (or permanent) determinants, such as "tastes for leisure," that affect both y and w-and-m;

(b) unmeasured, short-run (or transitory) sources of variation in w-and-m that may or may not translate into the t-induced (permanent?) changes in w-and-m;

(c) errors in measurement in w-and-m, now introduced for the sake of realism;

(d) random, yet permanent, sources of variation in w-and-m.

Only source (d) unambiguously provides variation in w-and-m that corresponds to the desired t-type variation. It is a challenge to specify what these "random, yet permanent" sources are. One of my favoriate can-

didates, following Mincer, and before him, Paul Douglas,[8] is geographic location of residence, which lends itself to grouping, which, in turn helps suppress the nettlesome sources, (a), (b), and (c). What is required, speaking loosely, is that geographic residence be considered an "accident," a happenstance, and that the nationwide market is sufficiently sluggish that w-and-m differences across (among) geographic units are sustained. Clearly, this particular source of w-and-m variation is no panacea. Even the apparent empirical consistency of obtaining all the "correct" signs for parameters when using standard metropolitan statistical areas (SMSAs) as units of observations was broken by Martin Dooley's work with the 1970 census (Dooley, 1982).

It is easy to get bogged down in philosophical discussions about methodology in this context. Estimating the conventional model may be described as a search for "structural" or "pure" wage and income parameters, which, when obtained, are applied with or without apology to the policy context. The apology comes when we admit that maybe the people involved in t are "different" or that the administrative context may change the parameters. No apology is required if we believe that all those characteristics that make the participants or the administrative context "different" have been held constant in the conventional model.

Recently two strategies have been proposed to measure "pure" w-and-m parameters. At the outset I will say that I admire the work, but I am not persuaded that either provides a panacea. One strategy is the class of selection bias adjustments dealing with the two-equation models like (1) and (2) above.

It is important to note that the two-equation model, used above to estimate a policy effect, was initially given prominence by Heckman

(1974) in his innovations to correct a selection bias for a conventional model that involved natural market forces. Thus, the main variables of interest in the conventional model, w and m, may be considered t-type variables. As with t, we realize that w-and-m values have not been randomly assigned but are the result of some selection process. Equation (2) expresses this. I have already mentioned the reasons why I believe the procedures for estimating A in these classes of models are shaky--which in no sense implies that the models or the procedures are wrong. The problem is likely to be that the data do not provide us with z-type variables or do not provide us with sufficient residual variation in the nonlinear estimate of $\hat{t}$, given that $\underline{x}$ is held constant.

The second strategy is to rely on panel data, which has only recently become widely available for labor supply research. The appeal of panel data is well known. By measuring changes in w-and-m for the same person, all their covariation with unmeasured permanent characteristics of the person (like tastes) may be assumed to be eliminated. Nevertheless, the question remains: What are the sources of variation in Δw and Δm, and are these sources sufficiently similar to those induced by t? Or, in estimating conventional models, the question is, Are the sources of variation in Δw and Δm in the sample sufficiently similar to the sources of variation in w and m in the context for which we want to make population inferences? Looking back to the nettlesome list, items (b) and (c) loom much more ominous with change data.

Panel data may be used along with the assumption, implicit in the original Mincer model, that one's lifetime income is known. Given this assumption, the variation in wage rates for a given person during the time-duration of the panel may be considered to yield pure substitution

effects--or, more precisely, pure effects on the intertemporal substitution of labor supply. As a corollary, actual income variation may be considered transitory, permitting the measure of transitory income effects. Clearly, if the assumption of perfect foresight is rejected, then the wage and income variables contain permanent income effects. These issues are explored by Heckman and Macurdy (1980).

Mincer's original model was intended to explain the long-run increase in the labor supply of wives, and he employed several devices to suppress intra-cohort intertemporal substitution of labor supply in order to focus on inter-cohort changes.

The implicit scenario for Mincer's model, given his objectives, is both simple and complicated at the same time. The variation in wages (say) across SMSAs corresponds to the average lifetime variation that confronts successive cohorts. The observed labor supply response (in terms of labor force participation rates) is an approximation to the average lifetime amount of labor supply for a cohort.[9] Geographic-migration responses to wage differences, which are possible in cross sections, are assumed away. (Obviously, there can be no migration from one cohort to another in response to wage variation facing different cohorts.) Intertemporal (within-cohort) substitution is ruled out by the following assumptions: (a) Fertility differences, which lead to intertemporal substitution, are exogenous and controlled. (b) The age structure is constant across SMSAs, or, in other contexts, age is controlled. (c) Cyclical variation (another cause of intertemporal substitution) was controlled by the unemployment rate among SMSAs. Now consider Mincer's assumptions about the (residual) wage variation in his sample of SMSAs.

(d) The wage variation was "permanent"--corresponding to the long-run changes that the economy offered successive cohorts of women. (e) The wage variation was attributable to exogenous, technologically determined demand conditions, as expressed by the industrial structure of SMSAs, and the wage variation corresponded to the exogenous technologically induced wage growth that occurred over time. The marginal responses in the labor supply of wives in the SMSA were assumed to have no effect on this market wage they faced.

Now consider the assumptions about two other arguments in the conventional demand function. (f) Tastes for leisure and/or home production and (g) "other prices"--price variation in goods and services complementary or substitutable with leisure and/or home-production--were all assumed to have a zero effect, on average. Either the variation in these variables was "averaged out" across SMSAs and had zero variance, or the nonzero components had a net zero effect on labor supply. Actually, another assumption would also work; namely, that any nonzero variation across SMSAs in, say, tastes, corresponded in its covariance with wages to that which existed over time.

The assumption about zero variance in tastes for leisure or home-production was made less palatable when Reuben Gronau (1973) and Heckman (1974) pointed to the wage-selection bias when measuring wages of women who work. With differences in home-productivity among women, a higher market wage is necessary to induce women who are more home-productive to enter the labor market.

Indeed, wage endogeneity was already a probable source of bias with cross-section data because the existing theory of human capital suggests that more time at work raises one's wage. Moreover, common observation,

and some ad hoc theorizing, suggests that, generally, full-time work is paid more than part-time work. Finally, recall that Mincer assumed exogeneity of fertility, wife's education, and husband's income, and this has also been challenged subsequently.

To accept Mincer's original model, given its purposes, we have to argue that these sources of bias are inconsequential, or that similar "biases" characterize the long-term time trends--in which case, Mincer's cross-section relationships were unbiased as explanations and predictions of the time series.

The research since Mincer has modified or dropped the assumptions listed above, but most of this subsequent work has either been aimed at a different objective--predicting the effects of policy changes, most notably--or the objectives have not been made clear.

As noted above, research explicitly aimed at explaining the long-term increase in work by wives has not received much attention in recent years. This is unfortunate but, to be realistically humble, there are no pressing policy decisions that depend on precision in explaining these long-term trends. Qualitative consistency of the kind Mincer achieved--in particular, a substitution (wage) parameter that is positive and larger in absolute value than the income parameter--may be sufficient, pedagogically and aesthetically, when presenting labor economics to students or the public.

What about using these models to predict the future? Again, it is difficult to make a case for the policy importance of anything but short- and near-term forecasts of labor supply, and the Mincer-type estimates of long-run parameters are probably not appropriate, even if we

were confident about the assigned forecasted values of w, m, and other variables in the model. Indeed, as I have argued elsewhere, it is difficult to improve on simple time extrapolations for predicting women's labor force participation rates (LFPRs) in the short- and near-term, given a careful separation of age and marital-status groupings (Cain, 1979).

Regressions of the conventional models using time-series data on women's (and wives') LFPRs have proven to be a disappointing alternative to the simple time extrapolations. The parameter estimates for income and wage effects are not stable, and the income effect is sometimes positive. In any case, they should be interpreted as measuring the effects of short-run changes. The selection of control variables, other than income, wage rates, and unemployment rates, tends to be arbitrary, and the problem of multicollinearity, already severe with just income and wages, becomes worse. Because most of the official annual time series began in 1947, there are at most 36 annual observations, and less if separate variables for black and white women are desired. (An illustration of these regressions is given by Mitchell, 1978, and Cain, 1979.) Again, these disappointing results are not necessarily serious, given the alternative of relatively simple time-series extrapolations. See R. Smith (1979) for a reasonable-appearing set of forecasts, using this method. The theoretical justification is straightforward; namely, that the complex set of factors that have operated in the past to determine women's labor supply will operate similarly over the short-run and near-term.

B. Explaining Wage Differences between Men and Women: Positive (Analytic), Normative, and Descriptive Issues, and Their Relation to Labor Supply Research

Wage differences are typically related to demand conditions or, on the supply side, to human capital investments. I briefly include the topic in my survey of developments in labor supply research for four reasons. First, an important component in models of long-run wage determination is a variable measuring labor supply. The roles of these two variables in labor supply models, as dependent and independent variables, are reversed in wage models. Either this signals a problem of specification error in conventional estimation of the models, or we have to examine closely the justification for the assumed exogeneity of whatever variable is on the right-hand side of the single-equation model. The latter interpretation leads to my second reason; namely, that sex differences in wages is another example of research that has suffered from a lack of attention to the practical purpose of the analysis. The third and fourth reasons are that the topic of sex differences in wages permits me to raise the issues of normative and descriptive purposes in research on labor supply.

Beginning with the last reason, recent time-series data and panel data have given us better descriptive measures of the "facts" regarding the labor supply and wages of various demographic groups. In particular, we currently have improved measures of wage rates. It was not unusual for older empirical studies of labor supply to use earnings-per-time-period divided by hours-worked-per-time-period as the wage. This was an obviously treacherous independent variable in a regression of hours

worked.[10] Another improvement is that panel data offer richer measures of labor supply, both by defining hours worked more extensively and by measuring the timing decisions of labor supply. We need further attention to these measurements, however. Three problems stand out in my view.

1. First, measures of labor supply that better correspond to "lifetime" values need to be constructed. These would be used with the lifetime models that attempt to explain long-term trends, and the measures also have a normative aspect, to be explained below.

With data from the census and supplementary sources, for example, the LFPRs and hours worked for cohorts of women may be traced over a 90-year period, 1890 to 1980. A lifetime measure of labor supply for a given cohort may be crudely measured by using six or seven census dates and interpolating values for between-census years. My calculations with this type of measure of lifetime labor supply show the increase in women's market work to be less dramatic than that illustrated by LFPRs, but it is still substantial. By assuming 16 hours per day of discretionary time during the ages 14 to 70, I calculate the ratio of hours worked to hours available. I find that around 1900 the average woman could expect to spend about 8 percent of her adult life (discretionary time) in market work. By 1970 this had increased to 13 percent. The average man in 1900 could expect to spend about 43 percent of his adult life in market work, and by 1970 only 25 percent.[11]

2. Second, measuring nonmarket uses of time is still in a primitive stage, and yet there are many positive issues regarding child care and schooling and some normative issues that depend on these measurements.

Consider our contrasting judgments of the above statistics on lifetime labor supply. We view the decline in market work by men to reflect an important component and source of the rise in their standard of living and economic well-being. At the same time we view the increase in market work by women to reflect increased opportunities and a rise in their economic well-being. The key to reconciling these contrasting judgments is the change in time spent in housework. The long-run rise in leisure attributed to men will also apply to women only if the decrease in women's housework has been greater than the increase in their market work. I know of no persuasive study of this issue that spans the past 100 years.[12]

3. Third, we need to exploit panel data to measure changes over time in labor supply and wages that will permit us to distinguish real changes in, say, work experience and wage rates from the effects of changes in the composition of the sample. For example, when asking whether the wage gap between men's and women's wages is widening or narrowing, it is essential to know whether the observed snapshots we get at different points in time reflect real changes in wages for quality-constant units of male and female labor or whether they reflect the changing composition of the male and female samples. Wage changes are more troublesome than such measures of labor supply as LFPRs precisely because wages are measured for a truncated population of only those in market employment. LFPRs are based on the entire population.

When the fraction of women who work is increasing, it is easy to show that the average wage of working women can decrease even though the market wage of every women may have increased. The new entrants are

expected to lower the average. Let us assume further that the average experience of women workers decreases because the inexperience of the new entrants dominates the increased experience owing to a simultaneous reduction in the exit rate of women workers. Now, during the period when women's LFPRs increased, the LFPRs of men decreased. The men who left the labor force, in particular the elderly or those with disabilities, are likely to be the low-wage workers. Thus, under these conditions, the stock of working men and women is changing to decrease the ratio of average women's wages to average men's wages, even though there may have been a rise in the market wage offered to the population of women relative to the population of men. Correcting for these selection biases is difficult.

One useful measure of the change in relative wages for men and women could be obtained by longitudinal data for those who worked continuously during the period. However, if women exit from the labor force more than men, then restricting the comparison to continuous workers would create a probable selection bias that would overstate the true change in the female/male wage ratio. This is because the change (increase) in wages, or "wage offers," would tend to be smaller for those who drop out of the labor force.

The main controversy about the wage gap between men and women concerns the measurement of sex discrimination, which may be defined as a higher wage paid to men than to women who are equal in productivity. I believe the term "productivity" should refer to the ability of the worker to perform a task, given an equal opportunity and willingness to do so. At any point in time the occupational distribution will show men with

more training and higher skills on average. A central question is, To what extent does this occupational and skill distribution reflect the voluntary choices of men and women regarding the division of labor between market and home sectors? Conversely, to what extent does this reflect market discrimination imposed on women? For some policy purposes this latter question is focused on whether the employer, other male workers, or customers are the specific sources of market discrimination. An important analytic issue concerns the misallocation of resources when two prices for the same good prevail. The normative issue about the inequity is obvious.

I doubt whether the usual econometric procedures for measuring labor market discrimination can answer these questions. The usual measure is the residual in average wages between working men and women after accounting for their "endowments" (or productivity characteristics) with a vector, $\underline{x}$, of control variables. The major problem is the ambiguity of the x's as representations of "endowments." Frequently the x's reflect market discrimination directly, and such x's should not be held constant. A glaring example occurs when the x's measure occupations. A more subtle example is the use of an x that measures labor market experience, defined as some measure of the quantity of labor supplied. This is the "simultaneity" problem referred to previously.

The point holds for any x-variable that is endogenous--that is, affected by the labor market. Economists view education, for example, as an investment that is to some extent responsive to rewards in the market. The number of children women desire may be influenced by the labor market opportunities for women, and so on. Thus, I object to the practice of

using the number of children as an instrument to represent experience, where the implicit claim is that the number of children, on its own, is unrelated to market rewards (and, therefore, unrelated to market discrimination).

Fundamentally, I doubt whether the question about "labor market discrimination" is specific to a practical policy issue. I do not question the importance of the global issues of efficiency and equity; rather, I doubt that our conventional econometric procedures can answer these issues.

It is interesting to examine the use of conventional multiple regression models in the analysis of discrimination in individual firms in court cases or other litigation proceedings stemming from anti-discrimination laws. These analyses have three advantages over the usual nationwide labor-market studies. First, the objectives are explicit--a guilty or innocent verdict! Second, a variety of characteristics--x-variables--may well be exogenous to a given employer, even though they are not exogenous to the labor market as a whole. Third, explicit information about the employer's criteria for hiring, retention, promotion, or pay may be used to select the x-variables. Regression analyses with nationwide samples usually suffer from ambiguities and vagueness about all these points. Unfortunately, the analyses of data from a single firm have two serious faults that invalidate their use for assessing market discrimination. First, the selection rules for becoming part of the data base are seldom known; second, the sampling variability is unknown but probably very large.

IV. CONCLUSION

I have focused on women's labor supply to illustrate a variety of methodological points about labor supply research in the past 20 years. I have included a discussion of wage differences between men and women to reinforce my main themes about labor supply research. These may be summarized as follows.

The pervasive endogeneity of the variables we use prevents us from estimating "complete" or "completely general" models. Indeed, I would go further and deny that the term "structural model" has any useful meaning, other than self-flattery, unless the user has carefully explained the purposes for the estimates and has explained how the sample is appropriate for these purposes.

A variety of practical, sometimes policy-related issues concerning labor supply can be addressed, however, because the question can often be narrowed to permit accepting many variables as predetermined. In these practical contexts the estimation question that I believe needs more attention is the following: What is the sample source of variation in the independent variable of interest, and does this source of variation correspond to the variation in the population about which we wish to make estimation inferences?

Our research on many practical issues has been and could be further helped by attention to refining a variety of descriptive statistics concerning labor supply and wages. Finally, often there are normative issues that deserve to be made explicit. On the one hand, they serve to motivate the research and, I believe, sharpen our focus on the practical issues. On the other hand, they serve to place some boundaries on the questions we can and cannot answer.

NOTES

[1]The paper was presented at a conference of the National Bureau of Economic Research in 1961 and published in 1962. See Mincer (1962). Throughout this paper I will discuss the labor supply of wives. The general points could apply, with minor changes, to the labor supply of any family member.

[2]Durand (1948) and Long (1958) had demonstrated that these demographic changes could not explain the rise in labor force participation rates of married women.

[3]See Mincer (1966), Bowen and Finegan (1969), and Cain and Mincer (1969).

[4]A partial list of references includes Ben-Porath (1973), Gronau (1973), Heckman (1974), Cogan (1980), and Hanoch (1980).

[5]A partial list of references includes Weiss (1972), Ghez and Becker (1975), Heckman and Macurdy (1980), and J.P. Smith (1980).

[6]I am indebted to several papers by James Heckman and to many discussions with Arthur Goldberger for clarifying my ideas about this model.

[7]Economists may claim that these aspects of the program are outside their jurisdiction, but the line is often hard to draw. "Work tests" or special inducements to work may completely change the basis for the prediction about the effects of an income maintenance law on labor supply.

[8]Douglas and Erika Schoenberg (1937) did not, however, attempt to separate income and substitution effects, as did Mincer.

[9]Cain and Dooley (1976) used five separate age groups with SMSA data, and we may refer to "each cohort." Mincer used the total population of adult women as a "representative cohort." To illustrate the interpretation of the labor supply measure, assume that the labor force participation rates of wives in two SMSAs are 40 percent and 20 percent, respectively. Then wives in the first SMSA are assumed to supply twice the amount of lifetime market work as wives in the second SMSA.

[10]Any random errors in measurement of hours worked tends to produce a spurious negative correlation between hours and earnings/hour.

[11]Unpublished tabulations, available from the author.

[12]For a widely publicized study that denied that the reduction in housework offset the increase in market work, see Vanek (1974).

REFERENCES

Ashenfelter, O. and J. J. Heckman. 1973. "Estimating labor-supply functions," in G. G. Cain and H. W. Watts, eds., Income Maintenance and Labor Supply. Chicago: Rand-McNally Co.

_____. 1974. "The estimation of income and substitution effects in a model of family labor supply," Econometrica, 42, January, 73-85.

Ben-Porath, Y. 1973. "Labor force participation rates and the supply of labor," Journal of Political Economy, 81, May/June, 697-704.

Bowen, W. G., and T. A. Finegan. 1969. The Economics of Labor Force Participation. Princeton, N.J.: Princeton University Press.

Cain, G. G. 1979. "Comment," on papers by R. Smith and J. Heckman, in C. B. Lloyd, E. S. Andrews, and C. S. Gilroy, eds., Women in the Labor Force. New York: Columbia University Press.

_____, and M. D. Dooley. 1976. "Estimation of a model of labor supply, fertility, and wages of married women," Journal of Political Economy, 84, December, S179-S199.

_____, and J. Mincer. 1969. "Urban poverty and labor force participation: Comment," American Economic Review, 59, March, 185-194.

Cogan, J. F. 1980. "Labor supply with costs of labor market entry," in J. P. Smith, ed., Female Labor Supply: Theory and Estimation. Princeton, N.J.: Princeton University Press.

Dickinson, J. G. 1976. "The estimation of income-leisure preference structures for prime age married males," Ph.D. dissertation, University of Michigan, Ann Arbor.

_____. 1980. "Parallel preference structures in labor supply and commodity demand: An adaptation of the Gorman polar form," Econometrica, 48, November, 1711-1726.

Dooley, M. D. 1982. "Labor supply and fertility of married women: An analysis with grouped and individual data from the 1970 U.S. Census," Journal of Human Resources, 17, Fall, 499-532.

Durand, J. D. 1948. The Labor Force in the United States, 1890-1960. New York: Social Science Research Council.

Ghez, G., and G. S. Becker. 1975. The Allocation of Goods and Time Over the Life Cycle. NBER, New York: Columbia University Press.

Gronau, R. 1973. "The intrafamily allocation of time: The value of the housewives' time," American Economic Review, 63, September, 634-651.

Ham, J. C. 1982. "Estimation of a labor supply model with censoring due to unemployment and underemployment," Review of Economic Studies, 59, 335-354.

Hanoch, G. 1980. "Hours and weeks in the theory of labor supply," in J.P. Smith, ed., Female Labor Supply: Theory and Estimation. Princeton, N.J.: Princeton University Press.

Heckman, J. J. 1974. "Shadow prices, market wages, and labor supply," Econometrica, 42, July, 679-694.

_____, and T. Macurdy. 1980. "A life cycle model of female labour supply," Review of Economic Studies, 47, January, 47-74.

Kosters, M. H. 1966. "Income and substitution effects in a family labor supply model," Ph.D. dissertation, University of Chicago.

Long, C. D. 1958. The Labor Force under Changing Income and Employment. NBER, Princeton, N.J.: Princeton University Press.

Mincer, J. 1962. "Labor force participation of married women," in Aspects of Labor Economics. Universities-NBER, Princeton, N.J.: Princeton University Press.

_____. 1968. "Labor force participation and unemployment: A review of recent evidence," in R. A. Gordon and M. S. Gordon, eds., Prosperity and Unemployment. New York: John Wiley and Sons, 73-112.

Mitchell, O. S. 1978. "Time series and cross section analysis of work rates of married and unmarried women," Ph.D. dissertation, University of Wisconsin, Madison.

_____. 1980. "The cyclical responsiveness of married females' labor supply: Added and discouraged worker effects," in Proceedings of the Thirty-Second Annual Meeting of the Industrial Relations Research Association. Madison, Wisconsin: IRRA, University of Wisconsin.

Schoenberg, E. H., and P. H. Douglas. 1937. "Studies in the supply curve of labor," Journal of Political Economy, 45, February, 45-79.

Smith, J. P. 1980. "Assets and labor supply," in J.P. Smith, ed., Female Labor Supply: Theory and Estimation, Princeton, N.J.: Princeton University Press.

Smith, R. E. 1979. Women in the Labor Force in 1990. Washington, D.C.: The Urban Institute.

Vanek, J. 1974. "Time spent in housework," Scientific American, November, 116-120.

Weiss, Y. 1972. "On the optimal lifetime pattern of labor supply," Economic Journal, 82, December, 1293-1315.