NBER WORKING PAPER SERIES

NEW METHODS FOR ESTIMATING LABOR SUPPLY FUNCTIONS: A SURVEY

James J. Heckman

Thomas E. MaCurdy

Working Paper No. 858

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

February 1982

This research was supported by NSF Grant SOC-77-27136 and Grant 10-P-9074819-01 from the Social Security Administration to NBER. This paper was presented at a conference on Labor Markets, Magdalen College, Oxford, England, September, 1979. The research reported here is part of the NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #858
February 1982

New Methods for Estimating Labor Supply Functions: A Survey

Abstract

This paper surveys new methods for estimating labor supply functions. A unified framework of analysis is presented. All recent models of labor supply are special cases of a general index function model developed for the analysis of dummy endogenous variables.

Professor James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637

(312) 753-4533 or -1106

Professor Thomas E. MaCurdy
Department of Economics
Stanford University
Stanford, California 94305

(415) 497-9694

Introduction

Recent empirical studies of labor supply have their foundations in the statistical theory of "index functions." This theory offers a general methodology to correct for sample selection biases and provides a conceptually simple framework for modeling corner solutions to the consumer's utility maximization problem. In this survey we show how recent studies on labor supply dealing with the topics of labor force participation, fixed costs of work, and taxes can all be fit within a general "index function framework."

The statistical theory of index functions has its foundation in the literature on dummy endogenous variables in a system of simultaneous equations. This literature is based on the notion that discrete endogenous variables are generated by continuous latent variables crossing thresholds. Tobin's (1958) seminal paper on estimating the demand for consumer durables is the first application of index function theory in economics. In Tobin's model, the intensity of demand for the durable good is the index function. Due to population variation in reservation prices, many consumers are at corner solutions. Only if the intensity of demand exceeds some minimum level (the threshold) does the individual purchase the durable good.

Over the past decade, the "index function framework" has been used extensively to analyze many problems in the area of consumer choice, including the analysis of quantal (i.e., discrete) choice (McFadden, 1974, 1976 and Domencich and McFadden, 1976). This framework also forms the basis for recent work dealing with problems arising from the use of nonrandom samples and sample selection biases. A general discussion of index functions and their relationship to simultaneous equation models which incorporate both continuous and discrete variables appears in Heckman (1978).

This survey begins with a discussion of the basic statistical framework found in many recent labor supply studies. In Section II we interpret Heckman's (1974a) model of joint labor supply, participation, and wage rates within this framework. Section III considers extensions of this model to incorporate fixed costs associated with labor supply, and it interprets the studies of Cogan (1976, 1979), Hanoch (1976, 1979), and Hausman (1979). Section IV develops a basic model for analyzing progressive tax schemes and labor supply. In Section V we generalize this model to allow for regressive as well as progressive taxes along the lines proposed by Burtless and Hausman (1978), Hausman (1979), and Wales and Woodland (1979). Discussion of specific methods of estimation is deferred to an appendix.

I. The Basic Index Model

The prototype for all of the models considered in this paper is a simple binary choice model. Let $V_{(1)}(R_1, \nu)$ be the best attainable utility for a consumer who does not work. R_1 is unearned income and ν is an unobserved "taste" component. Let $V_{(2)}(R_2, W, \nu)$ be his best attainable utility given that he works. His net wage is W and R_2 is his unearned income measured net of money costs of work (and other money transactions costs). R_2 may differ from R_1 due to the presence of work-related fixed costs. The net wage is irrelevant in evaluating the utility of the no work state.

If $V_{(2)} > V_{(1)}$, the consumer works. Otherwise, he does not. Write $Y_1 = V_{(2)} - V_{(1)}$, then if

$$(1) \qquad Y_1 = V_{(2)} - V_{(1)} > 0$$

the consumer works. This condition, and closely related conditions, underlie much recent work in labor supply.

If the consumer works, one may define an hours of work equation for the consumer. If equilibrium hours of work are determined by equating marginal benefits of work with marginal costs, one may use Roy's identity to derive the equation governing hours of work as

$$H = \frac{\partial V_{(2)}}{\partial W} \Big/ \frac{\partial V_{(2)}}{\partial R_2}$$

The hours of work equation may be written as

$$Y_2 = H(R_2, W, \nu).$$

In the ensuing analysis, we use this notation for the hours of work equation even when the conditions required for the application of Roy's theorem no longer apply.

Condition (1) is a prototype of a class of sample selection conditions that has received considerable attention in the recent econometric literature. In order to focus on the essential statistical issues involved, consider two linear functions

$$(2) \qquad Y_1 = Z_1\beta_1 + \varepsilon_1$$

$$(3) \qquad Y_2 = Z_2\beta_2 + \varepsilon_2.$$

ε_1 and ε_2 are mean zero random variables with finite second moments which are distributed independently of the vectors Z_1 and Z_2. Suppose we seek to estimate β_2. However, we only have data on individuals in a sample for which $Y_1 > 0$.

The regression for Y_2 given Z_2 and $Y_1 > 0$ is

$$(4) \qquad E(Y_2|Z_2, Y_1 > 0) = Z_2\beta_2 + E(\varepsilon_2|Y_1 > 0)$$

$$= Z_2\beta_2 + E(\varepsilon_2|\varepsilon_1 > - Z_1\beta_1).$$

If ε_1 and ε_2 are independent, $E(\varepsilon_2|\varepsilon_1 > - Z_1\beta_1) = 0$. Otherwise, the conditional mean of ε_2 depends on Z_1 and, in particular, on the probability that an observation with characteristics Z_1 is observed.

Write the joint density of ε_1, ε_2 as $f(\varepsilon_1, \varepsilon_2|\theta)$ where θ is a vector of parameters that generate the density. The probability that $Y_1 > 0$ is simply

$$1 - F_1(-Z_1\beta_1|\theta) = \int_{-\infty}^{\infty}\int_{-Z_1\beta_1}^{\infty} f(\varepsilon_1, \varepsilon_2|\theta)d\varepsilon_1 d\varepsilon_2$$

where F_1 is the marginal distribution function of ε_1. The conditional density of ε_2 given $Y_1 > 0$ is

$$k(\varepsilon_2|\varepsilon_1 > -Z_1\beta_1, \theta) = \frac{\int_{-Z_1\beta_1}^{\infty} f(\varepsilon_1, \varepsilon_2|\theta)d\varepsilon_1}{1 - F_1(-Z_1\beta_1|\theta)} .$$

Thus

$$E(\varepsilon_2|\varepsilon_1 > -Z_1\beta_1) = \int_{-\infty}^{\infty} \varepsilon_2 \; k(\varepsilon_2|\varepsilon_1 > -Z_1\beta_1, \theta)d\varepsilon_2$$
$$\equiv g(-Z_1\beta_1, \theta);$$

and, so,

$$E(Y_2|Z_2, Y_1 > 0) = Z_2\beta_2 + g(-Z_1\beta_1, \theta). \tag{5}$$

A regression of Y_2 on Z_2 that ignores the sample selection rule omits the term $g(\cdot)$ from the regression and standard specification error arguments apply.

For example, consider a variable Z_{2j} that appears in both Z_1 and Z_2. Let Y_2 be hours of work. A regression of Y_2 on Z_2 that does not correct for the sample selection term g estimates to a first order of approximation, instead of β_{2j},

$$\hat{\beta}_{2j} = \beta_{2j} + \frac{\partial g}{\partial Z_{2j}} .$$

The estimated value of β_{2j} differs from the true value by the second term on the right-hand side.

The essential point is that as we change Z_{2j} we alter the effective composition of the sample of workers. Computed partial derivatives with respect to Z_{2j} combine the ceteris paribus effect of changing Z_{2j} holding tastes fixed with the effect of changes in Z_{2j} on altering the sample distribution of tastes for work. The sample distribution of tastes for work is the distribution of ε conditional on condition (1) being met. This final effect is a consequence of entry and exit of observations from the sample due to the fact that condition (1) must be satisfied.

Condition (1) encompasses two distinct ideas which are sometimes confused in the literature. The first idea is that of self-selection. An individual chooses either to work or not to work. From an initial random sample of consumers, a sample of workers is not random in view of condition (1). The second idea is a more general concept--that of sample selection--which includes the first idea as a special case. From an "ideal" random sample, some rule is used to generate an observed sample of individuals. These rules may or may not be the consequences of choices made by the consumers being studied. For example, in the negative income tax experiments, decisions were made to "experiment" on low income populations. The effect of this restriction is that a decision rule generates the observed samples of workers and nonworkers. Since earnings are generated, in part, by tastes for work, these restrictions on sample membership operate on labor supply estimates based on selected samples in much the same way as condition (1). Econometric solutions to the general sample selection bias problem and the self-selection bias problem are identical. Much of the modern work on female labor supply and the analysis of experimental data, to be discussed below, is designed to eliminate the effects of sample selection bias on estimated structural labor supply functions.

The index function model given by equations (2) — (5) underlies all of the recent work on truncation and sample selection. For example, Cain and Watts (1973, p. 343) and Hausman and Wise (1977) consider a censoring problem that arises in analyzing data from the Negative Income Tax experiments. Labor supply functions are fit for experimental participants. However, to be an experimental participant, earnings, E, are required to be below a certain cutoff level $\bar{E}$. We thus observe individuals in the experiment only if $E < \bar{E}$. In terms of the index function models, we may write

$$Y_1 = \bar{E} - E.$$

Write labor supply as $Y_2 = Z_2\beta_2 + \varepsilon_2$. We observe Y_2 only if $Y_1 \geq 0$. Since it is plausible that the disturbances of the earnings function are not distributed independently of the disturbances of the hours of work function, the analysis of equations (2) — (5) applies with full force to this case.

The index function model can readily be generalized to encompass a multiplicity of sample generation rules and a multiplicity of behavioral functions. Y_1 and Y_2 may be vectors, and the simultaneous satisfaction of a set of sample generation rules can be characterized by the requirement that Y_1 lies in some subset of the feasible range of Y_1.

A version of the index model that will occupy our attention below is one in which individuals may be in any one of m states of the world where the value of the random variable Y_1 determines the state. In particular, a consumer is in state i if Y_1 lies in a set Θ_i which is some prespecified subset of the sample space of Y_1. The labor supply of a consumer occupying state i is determined by a function $Y_2 = H_{(i)}(R_i, W_i, \nu)$. Thus, the function relevant for determining a consumer's hours of work is state dependent in the sense that its form and/or its arguments differ across the various states. In the simple binary index model given by equations (1) - (3), there are two states of the world (i.e., $m = 2$), and the sample space of Y_1 is divided into two sets: $\Theta_1 = \{Y_1 : Y_1 \leq 0\}$ and $\Theta_2 = \{Y_1 : Y_1 > 0\}$. When in state 1 (i.e., $Y_1 \in \Theta_1$), a consumer's hours of work is given by $Y_2 = H_1 = 0$; and when in state 2 (i.e., $Y_1 \in \Theta_2$), $Y_2 = H_2 = Z_2\beta_2 + \varepsilon_2$.

All of the statistical models of labor supply to be considered below are special cases of this simple index function model. In the Heckman (1974a) model, there is a "work" and a "no work" state. Y_1 is the difference between the market wage and the reservation wage at zero hours of work. If $Y_1 > 0$, the consumer works. Given that he works, labor supply is defined as $Y_2 = \frac{1}{\gamma} Y_1$ where γ is a substitution parameter. The Cogan (1979) model of fixed costs is also a two-state model. Y_1 here is the difference between hours of work if the consumer is constrained to work and incur fixed costs, on the one hand, and "reservation hours" of work, on the other hand. If $Y_1 > 0$, the consumer chooses to work, and Y_2 is the hours of work chosen by the consumer. In the analysis of labor supply under progressive taxes where the budget constraint is composed of linear segments (due to discrete jumps in the marginal tax rate), each segment and kink corresponds to a different state of the world. Thus, in contrast to the Heckman and Cogan models, there is more than one state associated with "work." Y_1 is the marginal rate of substitution function, and as it takes values in various intervals, the consumer's labor supply equilibrium occurs on different branches of the budget constraint. While the hours of work function has the same form for each of the work states, the arguments of the function are appropriately modified to reflect the different tax parameters facing consumers on the different branches and kinks of their budget constraints. The Burtless and Hausman (1978) model of labor supply and taxes is general enough to deal with both regressive and progressive taxes. Their basic framework is the same as the one for the progressive tax case previously mentioned, except Y_1 in their model is an unobserved random "taste" component of the preference function. As Y_1 falls in various regions, equilibrium occurs on different segments of the budget constraint. The index function models can accomodate a wide variety of errors in the variables, including errors that arise from the inability to observe directly the particular state of the world a consumer occupies.

II. The Elementary Model of Labor Supply without Fixed Costs and Taxes

Consider a simple model of labor supply that neglects fixed costs of work and taxes. A consumer faces parametric wage $\dot{W}$.[1] Let X be a Hicks composite commodity of goods and L a Hicks composite commodity of nonmarket time. The consumer's strictly quasiconcave preference function may be written as $U(X, L, \nu)$ where ν is a "taste shifter." For the population of consumers, the density of ν is written as $f(\nu)$. This function induces a distribution on U. The maximum amount of leisure is T. Income in the absence of work is R.

A consumer works only if his best work alternative is better than his best nonwork alternative (i.e., full leisure). In the elementary model, a global comparison between the best attainable utilities in the work and no work states can be reduced to a local comparison between the marginal value of leisure at the no work position (the slope of the consumer's highest attainable indifference curve at zero hours of work) and the wage rate.

The marginal rate of substitution is defined as

$$(6) \qquad M(R, H, \nu) = \frac{U_2(R+WH,\ T-H,\ \nu)}{U_1(R+WH,\ T-H,\ \nu)}$$

where H is hours of work and $X = R+WH$. The reservation wage is $M(R, 0, \nu)$. The consumer works if

$$(7) \qquad M(R, 0, \nu) < W;$$

otherwise, he does not. If condition (1) is satisfied, the labor supply function is determined by solving the equation $M(R, H, \nu) = W$ for H to obtain

(8) $H = H(R, W, \nu)$.

There are three distinct concepts of labor supply or expected hours of work that are often confused in the literature. Consider a population of consumers who all face offer wage W and receive unearned income R but who have different ν's. The density $f(\nu)$ determines the distribution of "tastes for work" over the population. One measure of labor supply is the fraction of the population that works. Letting Θ denote the set of ν such that $M(R, 0, \nu) < W$, this fraction is

(9) $$P(W, R) = \int_{\Theta} f(\nu)d\nu = \text{Prob}\big(M(R, 0, \nu) < W\big)$$

where $\int_{\Theta}$ denotes integrating over the set Θ. The mean hours worked for <u>those employed</u> is

(10) $$E\big(H \mid M(R, 0, \nu) < W\big) = \frac{\int_{\Theta} H(R, W, \nu)f(\nu)d\nu}{P(W, R)} .$$

Yet a third measure of labor supply is the mean hours worked in the <u>entire population</u> which is given by

(11) $$E(H) = \int_{\Theta} H(R, W, \nu)f(\nu)d\nu$$

(remember $H(R, W, \nu) = 0$ for $\nu \notin \Theta$). The three measures of labor supply depend on some of the same parameters, but they are clearly distinct.

There is also some confusion in the literature concerning the appropriate interpretation of the partial derivatives of these different measures of labor supply. The partial derivatives of the hours of work function given by (8), H_W and H_R, produce the textbook uncompensated wage and income effects. It is crucial to note that these derivatives of P(W, R) with respect to W and R do not correspond to H_W and H_R (Lewis, 1967; Ben Porath, 1973). P_W must be positive, and H_W need not be. The partial derivatives of (10) or (11) with respect to W and R do not correspond to the Hicks-Slutsky terms H_W or H_R unless condition (7) is satisfied for everyone in the population. These simple points have been ignored in much of the literature. For example, Hall (1973) and Boskin (1973) interpret the partial derivatives of estimates of equation (11) with respect to W and R as estimates of H_W and H_R respectively. Others interpret partial derivatives of (10) (estimated from labor supply functions fit on samples of working individuals) as estimates of the Hicks-Slutsky parameters. If nonparticipation is a significant phenomenon in the population being sampled, estimates of (10) or (11) do not generate meaningful structural labor supply parameters.[2]

The model of Heckman (1974a) can be written within the index function framework. Write the marginal rate of substitution function given by (6) in semilog form as

$$\ell n\, M(R, H, \nu) = \alpha_0 + \alpha_1 R + \alpha_2 Z_2 + \gamma H + \nu \qquad (12)$$

where ν is a mean zero, normally distributed error term.[3] Market wage rates are written as

$$\ell n\ W = \beta_0 + \beta_1 Z_1 + V \tag{13}$$

where V is a normally distributed error term with mean zero. Solving equations (12) and (13) for hours of work for those observations satisfying $\ell n\ W > \ell n\ M(R, 0, \nu)$, one obtains

$$\begin{aligned} H &= \frac{1}{\gamma}\left[\ell n\ W - \ell n\ M(R, 0, \nu)\right] \\ &= \frac{1}{\gamma}(\beta_0 + \beta_1 Z_1 + V - \alpha_0 - \alpha_1 R - \alpha_2 Z_2 - \nu) \\ &= \frac{1}{\gamma}(\beta_0 - \alpha_0 + \beta_1 Z_1 - \alpha_1 R - \alpha_2 Z_2) + \frac{1}{\gamma}(V - \nu). \end{aligned} \tag{14}$$

In terms of the two-equation index model,

$$\begin{aligned} Y_1 &= \ell n\ W - \ell n\ M(R, 0, \nu) = (\beta_0 - \alpha_0 + \beta_1 Z_1 - \alpha_1 R - \alpha_2 Z_2) + (V - \nu) \\ Y_2 &= H = \frac{1}{\gamma} Y_1 \end{aligned} \tag{15}$$

so that the parameters of the sample selection rule ($Y_1 > 0$) are intimately related to the labor supply function. Assuming that ν and V are joint normally distributed, equations (13) and (14) generalize the "Tobit" model proposed by Tobin (1958). Provided that one variable appears in (12) that does not appear in (13), γ can be identified by a maximum likelihood procedure or a two-stage procedure.

We note, parenthetically, that in most data sets on labor supply, the condition that $Y_1 > 0$ must also be satisfied in order to observe the wage rate. Estimates of wage functions fit on samples of workers are subject to the same sort of sample selection bias that contaminates labor supply functions fit on subsamples of workers (Gronau, 1973). Assuming normally distributed error components, Heckman (1974a) builds a model that incorporates an hours of work equation (14) and a wage equation (13) that explicitly corrects for the effect of the condition $\{\ln W > \ln M(R, 0, \nu)\}$ on generating observations on workers.

III. Labor Supply Models with Fixed Costs

Cogan (1976, 1979) and Hanoch (1976, 1979) extend the simple model of the previous section by allowing for a nonconvex choice set arising from fixed costs of work such as commuting costs, expenditures on clothes needed for work, etc. The motivation for introducing fixed costs is to account for the small number of observations near zero in the hours of work distributions computed for workers. Our exposition follows more closely the work of Cogan.

Consider Figure I. A consumer's no work indifference curve is given by RA. For the simple model described above, if the wage rate is given by the slope of RC, the consumer works, and a standard interior solution labor supply function is generated. If the wage rate is given by the slope of line RB, the consumer does not work. The introduction of fixed money costs of work means that the consumer must pay fixed money cost F in order to work. The breakeven wage, which causes the consumer to be indifferent between work (with the fixed cost) and no work, is given by the slope of the line connecting points R-F and D. If the consumer is given a wage with this slope, and works H_d hours, he is indifferent between work and nonwork. At any higher wage rate, he will work. As money costs are increased, so are reservation wage rates and the minimum number of hours that the consumer works if he works at all. The relevant reservation wage for labor supply is the slope of indifference curve RA at position D.

The labor supply curve thus has a discontinuity. The consumer either does not work at all or works at least H_d hours. "H_d" is called "reservation hours" in the literature. The slope of the no work indifference curve at H_d hours, W_d, is termed the reservation wage. The labor supply function for those who work is essentially a standard labor supply curve with unearned income R reduced by amount F, the fixed money costs of work.

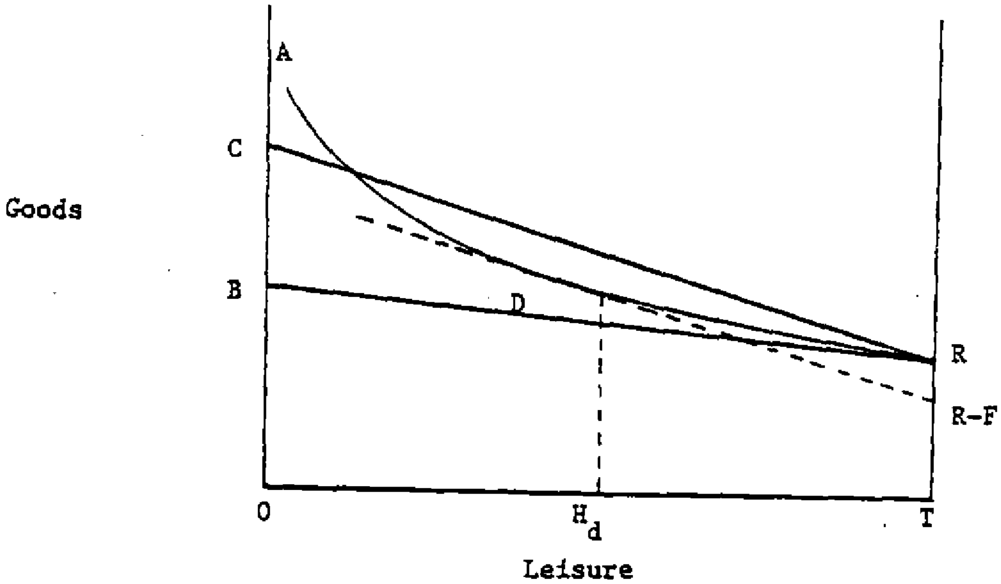

Figure 1

The analysis of fixed time costs of work parallels our discussion of fixed money costs of work except that time costs reduce hours worked by working individuals while money costs increase hours worked by working individuals (assuming leisure is a normal good). No new idea is involved.

The most direct way to solve for W_d and H_d is to use the indirect utility function. "W_d" is defined as that value of W such that

$$V(R-F,\ W_d,\ \nu) = U(R,\ T,\ \nu) \tag{16}$$

i.e., that value of W_d such that given the fixed costs of work, F, the consumer is indifferent between working at wage W_d and not working at all. This is the procedure used by Cogan (1979).

From equation (16) one can solve for W_d and by Roy's Identity one can solve for H_d,

$$H_d = \frac{V_W}{V_R} .$$

Thus one can write

$$H_d = H_d(R,\ F,\ \nu)$$

$$W_d = W_d(R,\ F,\ \nu)$$

$$H = H(R-F,\ W, \nu) .$$

Since the H_d, W_d, and H functions are derived from a common utility function, cross-equation restrictions connect these three functions. Without assuming explicit functional forms for the utility function it is difficult to impose these restrictions. Cogan does not impose a specific functional form and so does not exploit all the available information in the system. In practice, one does not have data on fixed costs. F is assumed to depend on a large set of observed variables, some of which do not enter the labor supply equation in their own right.

The consumer works provided that

$$H > H_d.$$

In terms of the index function model, we may write

$$Y_1 = H - H_d$$

$$Y_2 = H.$$

Thus we observe Y_2 only when $Y_1 > 0$.

Using standard sample selection bias correction procedures it is possible to obtain estimates of the parameters of the H function. Assuming a functional form for H and H_d, writing $H - H_d$ in reduced form, and assuming that one variable appears in H that does not appear in H_d (e.g., the wage rate),[1] it is possible to estimate the H_d function from the reduced form probability that the consumer works if the sample at hand contains both workers and nonworkers.

Thus, if

$$H = Z\beta + W\gamma + \varepsilon_1$$

$$H_d = Z\phi + \varepsilon_2$$

where Z is a set of exogenous variables, we have $H - H_d = Z(\beta - \phi) + W\gamma + \varepsilon_1 - \varepsilon_2$. Assuming that $\varepsilon_1 - \varepsilon_2$ is normally distributed, the probability that $H - H_d > 0$ is a probit probability. From probit analysis it is possible to estimate $\frac{(\beta - \phi)}{[\mathrm{Var}(\varepsilon_1 - \varepsilon_2)]^{\frac{1}{2}}}$ and $\frac{\gamma}{[\mathrm{Var}(\varepsilon_1 - \varepsilon_2)]^{\frac{1}{2}}}$. Combining these estimates with those of β and γ from the hours of work function, one can estimate ϕ.

All of the tests for the presence or absence of fixed costs that have been conducted within the Cogan framework have taken the Heckman (1974a) specification given by equations (13) and (14) as the baseline model of labor supply without fixed costs. This model assumes a strict proportionality relationship between the H and the $H - H_d$ equations (i.e., between the Y_1 and the Y_2 indices). The simple Heckman specification of the labor supply curve may be drawn as BB' in Figure 2. The intercept of the function is the log of the reservation wage. The key point of Cogan's analysis is that with fixed money costs of work, the labor supply function looks like CC'C". The log reservation wage is higher than B, a working consumer works at least H_d hours, and the labor supply function is above the Heckman function (because leisure is assumed to be a normal good). If CC'C" is the true labor supply function, and the linear Heckman function is fit to the data, labor supply elasticities will be overstated because the intercept in the linear Heckman labor supply curve is the reservation wage. Cogan's test for the presence of fixed costs amounts to determining whether the curve CC'C" explains the data significantly better than the curve BB'; if it does, then there are fixed costs associated with working.

Cogan's test crucially depends on an assumed functional form for the labor supply equation given by (14). If one permits nonlinearities in the log wage rate variable in this equation, then a nonlinear curve like BD in Figure 2 is also consistent with the proportionality assumption (see Heckman (1974b) for such an analysis). For all practical purposes this nonlinearity captures the essential features of Cogan's specification; namely, most consumers work a large number of hours if they work at all. Fixed costs may make a linear model of labor supply into a nonlinear model, but there are many reasons for nonlinearity. Tests for proportionality are more appropriately interpreted as tests for the presence or absence of nonlinearity in labor supply functions. Fixed costs are a source of nonlinearities, but evidence for or against nonlinearity is certainly not evidence for or against fixed costs.

Hausman (1979) extends Cogan's analysis of fixed costs by utilizing cross-equation restrictions on the H and H_d equations, and by utilizing another piece of information that Cogan neglects: that the position of the indifference

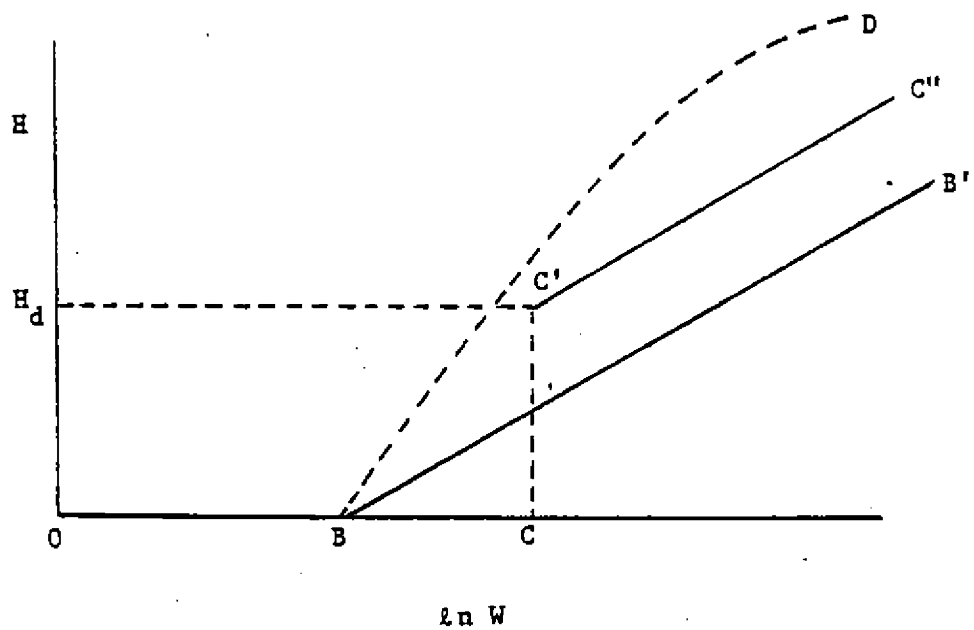

Figure 2

curve for nonworkers does not depend on fixed costs. Utilizing this information, he is able to estimate income and substitution effects of labor supply using only participation data. The price of these achievements is the imposition of stringent functional forms for preferences and assumptions about the way unobservables enter the model.

The utility of the consumer in the no work position is U(R, T, ν). Using the indirect utility function V(R-F, W, ν), Hausman is able to locate the best work alternative. It is possible that the indirect utility function is not defined for certain values of R, F, W and ν, but for such a case, the consumer does not work.

Hausman's specification of unobserved heterogeneity ν is strong but leads to econometrically useful results. Under his assumptions, U, V, and their difference, U - V, are monotonic functions of ν, a scalar random variable. Given W, R and an <u>exact</u> function for F (so that fixed costs are a function solely of observed variables),[3] Hausman can divide the domain of ν into two regions: a work region and a no work region. The boundary point for the region is given by ν^* which satisfies

$$V(R-F,\ W,\ \nu^*) = U(R,\ T,\ \nu^*).$$

Preferences are defined so that for $\nu > \nu^*$ the consumer works (i.e., $V(R-F, W, \nu) > U(R, T, \nu)$), and for $\nu \leq \nu^*$ the consumer does not work (i.e., $V(R-F, W, \nu) \leq U(R, T, \nu)$).

For each set of parameters (of V and U) and given R, W, and his assumed function for F, Hausman writes

$$Y_1 = V - U,$$

where Y_1 corresponds to the index function of equation (2). The probability of participation is simply the probability that $Y_1 \geq 0$. Given a distribution of "tastes," $f(\nu)$, the probability of working is

$$\int_{\nu^*}^{\infty} f(\nu)d\nu = \text{Prob}(Y_1 > 0).$$

A common set of parameters generate U and V. Hausman is able to estimate all of the parameters of the functions, and hence can generate all of the labor supply parameters, including income and substitution parameters, using only participation data. He is able to extract as much information as Cogan using less data because he assumes that the same linear labor supply function applies to the entire preference map, whereas Cogan uses a linear specification only as an approximation for the labor supply function for workers. However, given hours of work data, his procedure produces no more information than the Cogan procedure.

Two key assumptions in Hausman's model not needed in Cogan's, are: (1) fixed costs components are perfectly predicted by measured variables, and (2) data on wage rates are available for all individuals in the sample (including nonworkers). If either of these assumptions is violated, a more involved procedure is required which amounts to solving for ν^* given the unobserved components of F and of W and integrating over both these unobserved components. While it is conceivable that Hausman's first assumption concerning the perfect measurability of F is true, his second requirement concerning the availability of data on offer wages for nonworkers is surely violated for most data sets. We defer discussion of the consequences of not being able to observe wages for some individuals until Section V, where we develop a general framework for dealing with such unobservables.

IV. Labor Supply Models with Progressive Taxes

In this section we extend the simple model of hours of work without fixed costs outlined in Section II to incorporate progressive taxes. In contrast to the models of labor supply discussed above where there are two possible states of the world (i.e., work or no work), here we consider a multistate model. Although this extended model cannot be readily applied to the regressive tax case, it provides the essential ingredients required to analyze the general case.

Provided that the tax function facing the consumer generates a continuously differentiable strictly convex constraint set, the introduction of taxes into the model poses few analytical difficulties.[1] Define after-tax income as

$$E = E(WH + R;\ \psi), \quad E' > 0, \quad E'' < 0,$$

where ψ is a vector of parameters of the tax function (including exemption parameters and the like). The marginal wage rate at zero hours of work is

$$\left.\frac{\partial E}{\partial H}\right|_{H=0} = E'(R;\ \psi).$$

Replacing $E'(R;\ \psi)$ for W in condition (7), the analysis of labor force participation is the same as that given before. If the modified participation condition is met, one can linearize the budget constraint around the equilibrium hours of work position and solve for the structural labor supply equation in terms of the equilibrium marginal wage $E'(WH + R;\ \psi)$ and the intercept of the linearized budget constraint at the zero hours of work position, $E(WH + R; \psi) - E'(WH + R; \psi)H$. The marginal wage replaces the gross wage in equation (8), and the intercept term replaces R in the equation.

We may write the structural labor supply equation as

$$H = H(E', E - E'H, \nu).$$

If labor supply is measured with error, or there are disturbances in the labor supply equation, one must instrument the marginal wage and intercept terms of the linearized budget constraint to achieve consistent estimates.[2,3] This analysis carries over fully to estimation of labor supply functions in the presence of an equilibrium wage-hours locus (or hedonic line) of the sort considered by Lewis (1969) and Rosen (1974) provided that the constraint set facing the consumer is continuous and convex.[4]

The institutional features of the U.S. tax system are such that the assumption of smooth, continuously differentiable constraint sets for after-tax income is counterfactual. The U.S. tax system induces kinks and flats in the post-tax income function. A progressive tax system generates a convex budget set with linear segments and kinks such as the one depicted in Figure 3. To simplify the exposition, we consider only a two-flat function. The ensuing analysis may easily be extended to a multiple kink constraint. Given initial income (after tax) of R_2, and a gross wage rate W, after-tax income in the presence of a kinked tax schedule may be characterized by a marginal tax rate of t_A on the first segment, $(0, \overline{H})$, and a higher marginal tax rate t_B on the second segment, $(\overline{H}, T)$.

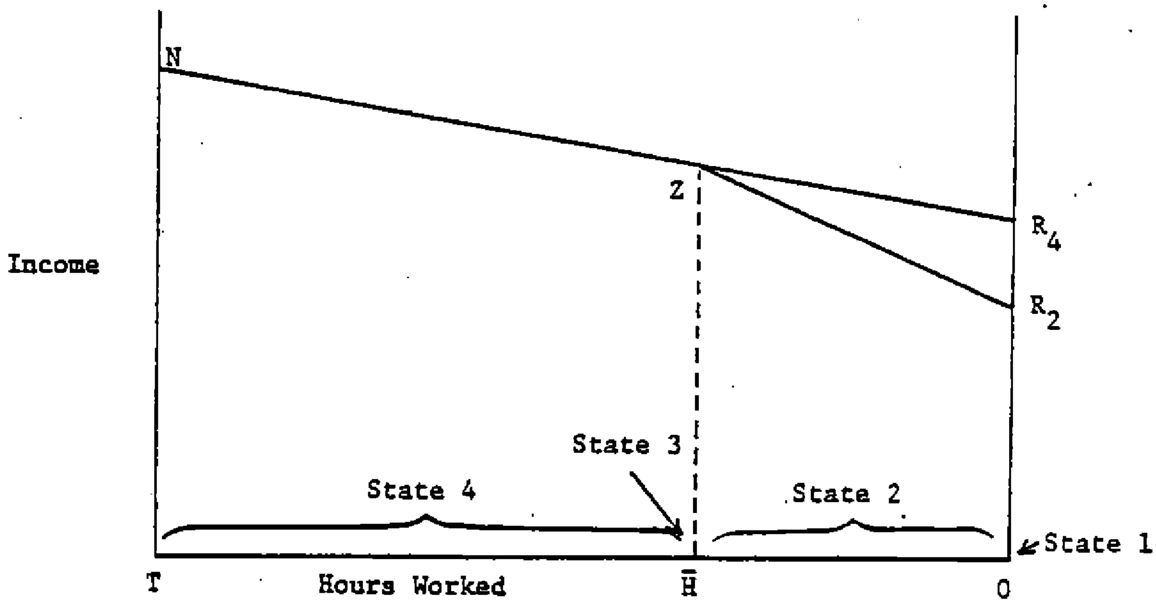

Figure 3

A consumer may occupy any one of four states of the world in this model of taxes. Each state corresponds to a different kink or segment of the budget constraint. A consumer who does not work is at kink point R_2 which constitutes state 1. A consumer who chooses to work, on the other hand, may be on segment R_2Z, or at kink point Z, or on segment ZN, which constitute states 2, 3, and 4, respectively. A consumer in state 2 faces a net marginal wage rate equal to $W_2 \equiv W(1 - t_A)$ and receives unearned income R_2. A consumer in state 4, on the other hand, faces an after-tax wage rate equal to $W_4 \equiv W(1 - t_B)$ and can be viewed as receiving the equivalent of $R_4 = R_2 + (W_2 - W_4)\bar{H} = R_2 + W(t_B - t_A)\bar{H}$ as unearned income.

As a consequence of convexity of preferences and the constraint set, a local comparison of the marginal rate of substitution function given by (6) and the after-tax wage rate at the kink points determines the location of an individual on the budget set. The consumer chooses not to work if

(17) $$M(R_2,\ 0,\ \nu) \geq W_2 \equiv W(1 - t_A).$$

The consumer works in the interval $(0, \overline{H})$ if

(18) $$M(R_2,\ 0,\ \nu) < W_2 \quad \text{and} \quad M(R_4,\ \overline{H},\ \nu) > W_2.$$

The consumer is in corner solution equilibrium at the kink Z provided that

(19) $$W_2 \geq M(R_4,\ \overline{H},\ \nu) \geq W_4 \equiv W(1 - t_B).$$

Finally, the consumer is at an interior solution in the interval $(\overline{H}, T)$ if

(20) $$M(R_4,\ \overline{H},\ \nu) < W_4.$$

It is straightforward to derive the implied labor supply function associated with each state of the world. To simplify the following exposition denote this function by $H_{(i)}$ for state i. In state 1, the no work state, obviously $H_{(1)} = 0$. At interior equilibrium on branch R_1Z, the labor supply function is determined by solving the equation $M(R_2,\ H_{(2)},\ \nu) = W_2 \equiv W(1 - t_A)$ for $H_{(2)}$; so, in state 2 hours of work is given by

(21) $$H_{(2)} = H(R_2,\ W_2,\ \nu).$$

In state 3, the corner equilibrium at Z, $H_{(3)} = \overline{H}$. Finally, in state 4 (an interior equilibrium along branch ZN), solving $M(R_4,\ H_{(4)},\ \nu) = W_4 \equiv W(1 - t_B)$ for $H_{(4)}$ implies a labor supply function of the form

$$(22) \qquad H_{(4)} = H(R_4, W_4, \nu).$$

While the functional form of the labor supply functions in states 2 and 4 is the same, the arguments of these functions differ as a consequence of the different tax rates facing consumers on the different branches of the budget constraint.

To set up this model as an index function model, the "taste shifter" component, ν, is a natural candidate for the index Y_1 to determine which state of the world the consumer occupies. For given values of R_1, W, $\bar{H}$, t_A, and t_B, the consumer chooses one of the four possible states depending on the region in which ν lies. The conditions relating $M(R_2, 0, \nu)$ and $M(R_4, \bar{H}, \nu)$ to after-tax wage rates listed above define these regions. Let Θ_1, Θ_2, Θ_3 and Θ_4 denote the subsets of the sample space of ν that satisfy conditions (17) — (20), respectively. If $\nu \in \Theta_i$ (i.e., ν is an element of the set Θ_i), the consumer chooses state i.

If we assume that the marginal rate of substitution function, $M(R, H, \nu)$, is monotonically increasing in ν, we obtain simple expressions for these sets. Each Θ_i is a single interval in the real line. Define ν_1^*, ν_2^*, and ν_3^* as those values of ν satisfying

$$M(R_2, 0, \nu_3^*) = W_2 \equiv W(1 - t_A)$$

$$(23) \qquad M(R_4, \bar{H}, \nu_2^*) = W_2$$

$$M(R_4, \bar{H}, \nu_1^*) = W_4 \equiv W(1 - t_B).$$

Convexity of preferences implies that $M(R_4, \bar{H}, \nu) > M(R_2, 0, \nu)$. Hence the monotonicity assumption implies that $\nu_1^* < \nu_2^* < \nu_3^*$. Conditions (17) - (20), then, define the regions as

$\Theta_1 = \{v : v_3^* \leq v\}$; $\Theta_2 = \{v : v_2^* < v < v_3^*\}$; $\Theta_2 = \{v : v_1^* \leq v \leq v_2^*\}$; and $\Theta_4 = \{v : v < v_1^*\}$.

Choosing a specification for the marginal rate of substitution function and a distribution for "tastes" in the population, $f(v)$, yields a complete statistical characterization of labor supply behavior. The probability that a consumer is in state i is

$$\text{(24)} \qquad \text{Prob}(v \in \Theta_i) = \int_{\Theta_i} f(v)dv.$$

The expected hours of work of a consumer who is known to be in state i is

$$\text{(25)} \qquad E(H \mid v \in \Theta_i) = E(H_{(i)} \mid v \in \Theta_i)$$

$$= \frac{\int_{\Theta_i} H_{(i)} f(v)dv}{\text{Prob}(v \in \Theta_i)}$$

The expected hours of work for a randomly chosen individual is

$$\text{(26)} \qquad E(H) = \sum_{i=1}^{4} E(H_{(i)} \mid v \in \Theta_i) \ \text{Prob}(v \in \Theta_i).$$

Estimation of structural labor supply parameters involves the same set of issues considered in the simple Heckman model except that instead of the single corner and single interior solution segments as in the Heckman model, there are two corners and two interior segments in the model with kinked progressive taxes considered here. In order to avoid sample selection bias in estimating structural labor supply functions, one must account for the conditioning that generates the observations (i.e., one must account for the particular branch or corner on which an observation is situated). It is obvious that, in this case, correcting for potential sample selection bias automatically corrects for the endogeneity in tax rates and unearned income levels. Indeed, in this model, endogeneity of taxes and sample selection bias come to the same thing.

To illustrate the procedure involved, consider the following empirical specification. Write $M(R, H, \nu)$ as

$$M(R, H, \nu) = m_0 + m_1 R + m_2 H + \nu. \tag{27}$$

Since M is monotonically increasing in ν, the formulas for ν_1^*, ν_2^*, and ν_3^* given by (23) provide a simple method for dividing the sample space of ν into the sets Θ_i. We obtain

$$\begin{aligned}
\nu_1^* &= W_4 - m_0 - m_1 R_4 - m_2 \overline{H} \\
\nu_2^* &= W_2 - m_0 - m_1 R_4 - m_2 \overline{H} = \nu_1^* + W(t_B - t_A) \\
\nu_3^* &= W_2 - m_0 - m_1 R_2 = \nu_2^* + m_1(R_4 - R_2) + m_2 \overline{H}
\end{aligned} \tag{28}$$

The probability that the consumer does not work is

$$\mathrm{Prob}(\nu \geq \nu_3^*) = \int_{\nu_3^*}^{\infty} f(\nu)d\nu = 1 - F(\nu_3^*) \tag{29}$$

where F is the cumulative distribution function of ν. The probability that the consumer is at interior equilibrium on the first segment, RZ, is

$$(30) \qquad \mathrm{Prob}(\nu_2^* < \nu < \nu_3^*) = \int_{\nu_2^*}^{\nu_3^*} f(\nu)d\nu = F(\nu_3^*) - F(\nu_2^*).$$

The probability of kink equilibrium at Z is

$$(31) \qquad \mathrm{Prob}(\nu_1^* \leq \nu \leq \nu_2^*) = \int_{\nu_1^*}^{\nu_2^*} f(\nu)d\nu = F(\nu_2^*) - F(\nu_1^*).$$

The probability of interior equilibrium along branch ZN is

$$(32) \qquad \mathrm{Prob}(\nu < \nu_1^*) = \int_{-\infty}^{\nu_1^*} f(\nu)d\nu = F(\nu_1^*).$$

An important assumption of this version of the model is that one can directly observe the state of the world each consumer occupies. Knowledge of a consumer's hours of work is all the information required. If $H = 0$, the consumer is in state 1; if $H \in (0, \bar{H})$, he is in state 2; if $H = \bar{H}$, he is in state 3; and if $H \in (\bar{H}, T)$, the consumer occupies state 4. By choosing a density function $f(\nu)$, it is possible to estimate directly the structural parameters determining the probabilities of each state of the world given by (29) - (32).

If we choose $f(\nu)$ as the normal density, this statistical model is an ordered probit scheme (see Johnson (1972) and Rossett-Nelson (1975)). Forming the sample likelihood, one can estimate all the parameters of the marginal rate of substitution function, and the variance of ν. (The variance of ν, σ^2, is estimated by normalizing by the standard deviation within each probability statement, and noting that the coefficient on $(1 - t)W$ is $1/\sigma$. The parameter m_2 is identified if the kink point in the after-tax income function comes at different hours of work for different consumers--which, plausibly, is the case.)

Since the parameters of the M function generate all of the parameters of the labor supply function, the ordered probit analysis suffices to determine the parameters of the labor supply function <u>without using any data on hours of work</u>. Given these probabilities, one can compute the conditional means of the interior solution of the hours of work function for each branch of the budget constraint and, by following a straightforward generalization of equation (5), achieve more efficient estimates of the labor supply parameters. However, no new parameter is estimated by this procedure. It is straightforward to write down the likelihood function for the full model and so achieve full efficiency in deriving the estimates.

The labor supply function implied by the linear specification for M associated with equilibrium on segment R_2Z is

$$H_{(2)} = H(R_2, W_2, \nu) = \alpha_0 + \alpha_1 R_2 + \alpha_2 W_2 - \frac{\nu}{m_2} \tag{33}$$

where $\alpha_0 = -\frac{m_0}{m_2}$, $\alpha_1 = -\frac{m_1}{m_2}$, and $\alpha_2 = \frac{1}{m_2}$. For branch ZN, it is

$$H_{(4)} = H(R_4, W_4, \nu) = \alpha_0 + \alpha_1 R_4 + \alpha_2 W_4 - \frac{\nu}{m_2} . \tag{34}$$

Thus, expected hours of work of a consumer whose equilibrium is on interval R_2Z is

(35) $$E(H_{(2)}|\nu_2^* < \nu < \nu_3^*) = \alpha_0 + \alpha_1 R_2 + \alpha_2 W_2 + E\left[-\frac{\nu}{m_2}\middle|\nu_2^* < \nu < \nu_3^*\right];$$

and expected hours of work of a consumer who is known to be on segment ZN is

(36) $$E(H_{(4)}|\nu < \nu_1^*) = \alpha_0 + \alpha_1 R_4 + \alpha_2 W_4 + E\left[-\frac{\nu}{m_2}\middle|\nu < \nu_1^*\right].$$

Writing $E\left[-\frac{\nu}{m_2}\middle|\nu_2^* < \nu < \nu_3^*\right]$ as $b_2\lambda_2$ and $E\left[-\frac{\nu}{m_2}\middle|\nu < \nu_1^*\right]$ as $b_4\lambda_4$, equation (26) implies that expected hours of work is

(37) $$\begin{aligned} E(H) = {} & (\alpha_0 + \alpha_1 R_2 + \alpha_2 W_2 + b_2\lambda_2)\big(F(\nu_3^*) - F(\nu_2^*)\big) + \bar{H}\big(F(\nu_2^*) - F(\nu_1^*)\big) \\ & + (\alpha_0 + \alpha_1 R_4 + \alpha_2 W_4 + b_4\lambda_4)F(\nu_1^*) \\ = {} & \bar{H}\big(F(\nu_2^*) - F(\nu_1^*)\big) + \alpha_0\big(F(\nu_3^*) - F(\nu_2^*) + F(\nu_1^*)\big) + \alpha_1\Big[R_2\big(F(\nu_3^*) - F(\nu_2^*)\big) + R_4 F(\nu_1^*)\Big] \\ & + \alpha_2\Big[W_2\big(F(\nu_3^*) - F(\nu_2^*)\big) + W_4 F(\nu_1^*)\Big] + b_2\lambda_2\big(F(\nu_3^*) - F(\nu_2^*)\big) + b_4\lambda_4 F(\nu_1^*). \end{aligned}$$

The empirical specification for hours of work implied by this analysis is

(38) $$H = E(H) + \varepsilon$$

where $E(H)$ is given by (37) and ε is a randomly distributed heteroscedastic error term with mean zero. This regression equation is estimated using a random sample of consumers; nonworkers are assigned $H = 0$.

One does not necessarily require complicated nonlinear methods to estimate the parameters of (37). Assuming that ν is normally distributed, it is possible to estimate each of the $F(\nu)$ functions by ordered probit analysis. Then one can form the product of the estimated $F(\nu)$ and the wage, tax, and income variables. Following a straightforward modification of the procedure of Heckman (1979), one can estimate λ_2 and λ_4 up to a factor of proportionality. Then regression of H on these variables will yield estimates of α_0, α_1, α_2, and the factors of proportionality on λ_2 and λ_4 (b_2 and b_4).[5]

There is a crucial implicit assumption in the preceding estimation procedures: hours of work are not measured with error, so measured hours reflect true or desired hours. If this is not so, data on hours of work do not suffice to allocate individuals to the correct branch of the budget constraint. The state of the world a consumer occupies can no longer be directly observed, and we confront a discrete data version of an errors in variables problem. If we use data on H to assign individuals to various states of the world, the ordered probit estimates the wrong $F(\nu)$ functions. Suppose that true hours of work, H, and measured hours, which we denote by H^*, are related by the equation $H^* = H + e$ where e is a disturbance representing measurement error which is distributed independently of H and all the explanatory variables. Then the probability that one observes H^* in a given interval is not the same as the probability that H falls in that interval.[6] For example, the probability that $H = \bar{H}$ is positive, but the probability that $H^* = \bar{H}$ is zero assuming e follows a continuous density function. Since we estimate the wrong $F(\nu)$ functions (by using ordered probit analysis) when assigning consumers to various states on the basis of their H^*, the variables formed by multiplying estimated $F(\nu)$ functions with W, t_A,

t_B, R_2, R_4, $\bar{H}$, and 1 in equation (37) are measured with error which produces inconsistent parameter estimates.

This measurement error problem only affects the proposed two-stage estimation technique. It does not lead to any serious complications in the above theory and empirical specifications. The expression for E(H) given by (37) is unchanged. Nowhere in its derivation do we require states of the world to be correctly observed. The equation for H* can always be written as

$$(39) \qquad H^* = E(H) + \varepsilon^*$$

where $\varepsilon^* = \varepsilon + e$.

This observation immediately suggests how to avoid the errors in variables problem discussed above in the two-stage procedure based on ordered probit analysis. Instead of using a two-stage procedure of the sort proposed above to estimate the hours of work function (by first predicting the $F(\nu)$ terms in equation (37) and forming variables using $F(\nu)$ for the second stage regression), one should estimate equation (37) directly by nonlinear least squares, exploiting all of the restrictions of the model, to generate parameter estimates. The ν_1^*, ν_2^*, ν_3^* terms contain parameters of the marginal rate of substitution function which are intimately linked with the parameters of the labor supply equation (see equation (28)).[7] As we will see in the next section, this principle is the technique used by Burtless-Hausman in their analysis of labor supply and taxes.

Allowing for measurement error in hours of work causes no essential change in the general formulas describing labor supply presented above. As previously noted, the formulas for the probabilities of occupying each state are exactly the same. The formula for the unconditional expected value of measured hours of work is the same as for true hours given by (37) except that H^* replaces H. Expected values in this instance are computed by integrating $H_{(i)}$ over the set Θ_i and the sample space of the measurement error component of hours using the joint density of ν and the measurement error component.

Two important assumptions maintained in this section are that data on potential wage rates are available for all individuals including nonworkers and that wage rates are exogenous variables. Relaxing these assumptions does not introduce major complications in the previous analysis.

Suppose that true market wage rates are generated by the function

$$(40) \qquad W = W(Q, \eta)$$

where Q includes a consumer's measured characteristics, and η is an error term representing the contribution of unmeasured characteristics. Conditions (17) — (20) still determine the state of the world a consumer occupies. Replacing W by $W(Q, \eta)$, we see that these conditions divide the sample space of (ν, η) into subsets associated with each state of the world; they define the sets Θ_i such that $(\nu, \eta) \in \Theta_i$ implies the consumer is in state i, and the probability of such an event is

$$(41) \qquad \text{Prob } ((\nu, \eta) \in \Theta_i) = \iint_{\Theta_i} k(\nu, \eta) d\nu d\eta$$

where $k(\nu, \eta)$ is the joint density of ν and η, and integration is carried out over the set Θ_i.

The labor supply functions associated with each state are unchanged except that $W(Q, \eta)$ replaces W in constructing the arguments of the functions for states 2 and 4 given by (21) and (22).[8] The expression for expected hours of work given by (26) becomes

$$(42) \qquad E(H) = \sum_{i=1}^{4} E\big(H_{(i)} \,|\, (\nu, \eta) \in \Theta_i\big)\ \mathrm{Prob}\big((\nu, \eta) \in \Theta_i\big)$$

where

$$(43) \qquad E\big(H_{(i)} \,|\, (\nu, \eta) \in \Theta_i\big) = \frac{\iint_{\Theta_i} H_{(i)} k(\nu, \eta)\, d\nu\, d\eta}{\mathrm{Prob}\big((\nu, \eta) \in \Theta_i\big)} \,.$$

To illustrate the problems that arise when one introduces a wage function, consider combining a wage function of the form

$$(44) \qquad W = B_0 + B_1 Q + \eta$$

with the linear specification of the marginal rate of substitution function used above to develop the empirical model of labor supply given by (37) and (38). Even if we assume that ν and η follow a joint normal distribution, an ordered probit analysis which allocates individuals to different states of the world no longer applies. While the conditions defining states given by (17) – (20) imply restrictions on linear combinations of ν and η, it is not possible to construct a single linear combination of ν and η whose value completely determines which state a consumer occupies.

In particular, define $\overline{W}$ as $E(W)$ (i.e., $\overline{W} = B_0 + B_1 Q$), and $\overline{R}_4 = E(R_4)$. Replacing W with $\overline{W}$ and R_4 with $\overline{R}_4$ in equations (28), we see that three randomly distributed error terms ω, ϕ_A, and ϕ_B defined by $\omega = \nu - \eta(1 - t_A)$, $\phi_A = \nu - \eta(1 - t_A) + m_1(R_4 - \overline{R}_4)$ and $\phi_B = \nu - \eta(1 - t_B) + m_1(R_4 - \overline{R}_4)$ determine a consumer's state of the world. According to conditions (17) - (20), a consumer is in state 1 if $\omega \geq \overline{W}(1 - t_A) - m_0 - m_1R_2 \equiv \omega^*$; in state 2 if $\omega < \omega^*$ and $\phi_A > \overline{W}(1 - t_A) - m_0 - m_1\overline{R}_4 - m_2\overline{H} \equiv \phi_A^*$; in state 3 if $\phi_A \leq \phi_A^*$ and $\phi_B \geq \overline{W}(1 - t_B) - m_0 - m_1\overline{R}_4 - m_2\overline{H} \equiv \phi_B^*$; and in state 4 if $\phi_B < \phi_B^*$. We see then, that the values of three different random variables (i.e., ω, ϕ_A and ϕ_B) determine occupancy of a state. Also, occupancy of either state 2 or 3 requires the simultaneous satisfaction of two conditions. Thus, while univariate probit analysis can be used to predict the probability of state 1 or 4, bivariate probit analysis is needed to specify the probability of state 2 or 3.

Modifying the expressions for expected hours of work to account for the linear wage function is straightforward using the notation of the previous paragraph. The labor supply functions for state 2 and for state 4 given by (33) and (34) become $H_{(2)} = \alpha_0 + \alpha_1R_2 + \alpha_2\overline{W}(1 - t_A) - \frac{\omega}{m_2}$ and $H_{(4)} = \alpha_0 + \alpha_1\overline{R}_4 + \alpha_2\overline{W}(1 - t_B) - \frac{\phi_B}{m_2}$. Thus, expected hours conditional on being in state 2 and in state 4 given by (35) and (36) become

$$E(H_{(2)}|\omega < \omega^* \text{ and } \phi_A > \phi_A^*) = \alpha_0 + \alpha_1R_2 + \alpha_2\overline{W}(1 - t_A) + E\left(-\frac{\omega}{m_2}\middle|\omega < \omega^* \text{ and } \phi_A > \phi_A^*\right)$$

and

$$E(H_{(4)}|\phi_B < \phi_B^*) = \alpha_0 + \alpha_1\overline{R}_4 + \alpha_2\overline{W}(1 - t_B) + E\left(-\frac{\phi_B}{m_2}\middle|\phi_B < \phi_B^*\right).$$

The labor supply functions associated with states 1 and 3 are unchanged. Given these new expressions for expected hours and the probabilities of occupying each state, one can easily modify expression (37) for the unconditional expected value of hours of work by making the appropriate substitutions.

Substituting this new expression for E(H) into the regression equation given by (38) creates an estimating equation for labor supply which allows wages to be endogenous and which does not require that wage offer data be available for all observations. This new specification can be estimated using the nonlinear least-squares procedure described above. To identify all the parameters, one must also estimate the wage equation using data on workers, adjusting for sample selection bias. This is accomplished by regressing W on $E(W|\text{consumer works}) = B_0 + B_1Q + E(\eta|\omega > \omega^*)$ where $E(\eta|\omega > \omega^*)$ can be computed using techniques proposed by Heckman (1979). It is also possible to estimate hours and wage equations jointly.

It is significant to note that η here represents the contribution of unobserved variables affecting true wages; it does not include a measurement error component for wages. Allowing for measurement errors in wages requires exactly the same type of adjustment in the formulas for probabilities and expected values as is required in treating measurement errors in hours. The formulas for probabilities and expected values given by (41), (42), and (43) remain valid when measurement error in wages is present, except that expected values are now computed by integrating $H_{(i)}$ times the joint density function of ν, $\dot{\eta}$, and the measurement error components over the set Θ_1 and the sample space of the measurement error components.

V. A General Treatment of Taxes and Labor Supply

This section formulates a model of hours of work that allows for regressive as well as progressive taxes. Here we develop a methodology to handle cases where local marginal comparisons do not fully characterize labor supply behavior following suggestions by Burtless and Hausman (1978), Hausman (1979), and Wales and Woodland (1979).

A regressive tax scheme leads to a budget set that is not convex. Figure 4 displays the case we consider here.[1] A marginal tax rate of t_A applies to the branch R_2S, and a lower marginal rate t_B applies to the branch SN.

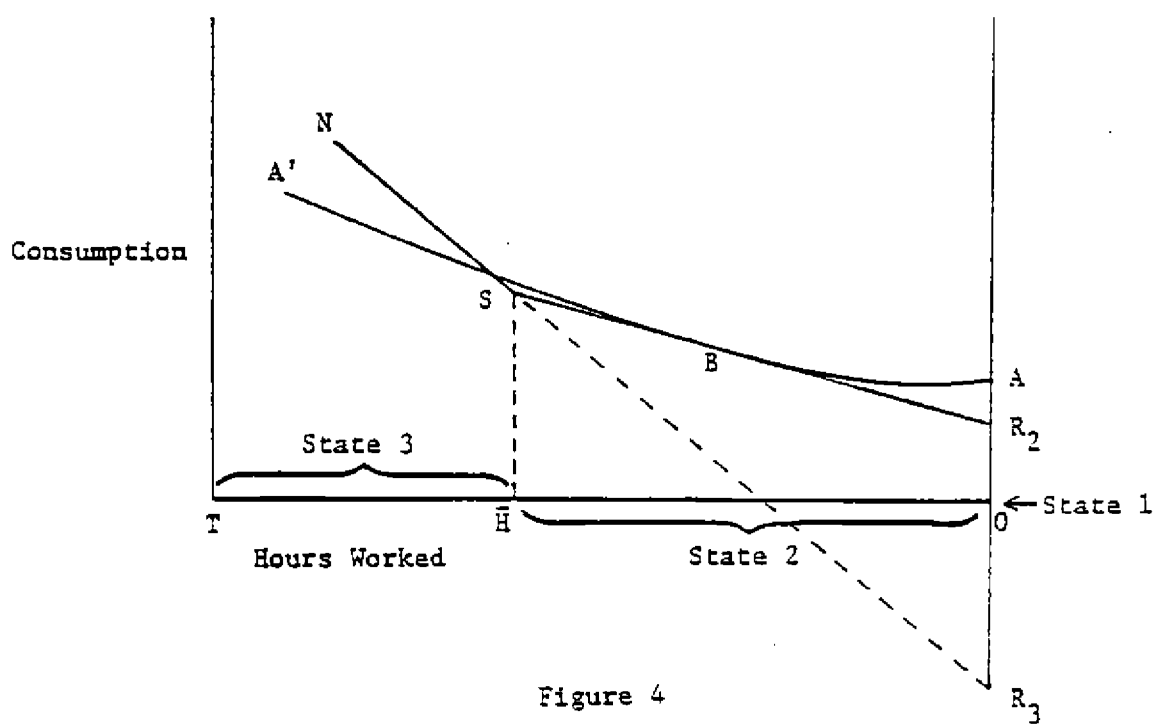

Figure 4

A consumer facing this budget set may choose any one of three possible states of the world: the no work position at kink point R_2 (i.e., state 1), or an interior equilibrium on either segment R_2S or segment SN (i.e., states 2 and 3).[2] A consumer in state 1 receives initial after-tax income R_2. In state 2, a consumer receives unearned income R_2 and works at an after-tax wage rate equal to $W_2 \equiv W(1 - t_A)$ where W is the gross wage. Finally, a consumer in state 3 earns after-tax wage rate $W_3 \equiv W(1 - t_B)$ and can be viewed as receiving the equivalent of R_3 as unearned income.

The analysis of kinked-nonconvex budget constraints involves an idea already considered in the analysis of fixed costs: a local comparison between the reservation wage and the market wage does not adequately characterize the work-no work decision. Due to the nonconvexity of constraints, existence of an interior solution on a branch does not imply that equilibrium will occur on the branch. Thus in Figure 4, point B assoicated with indifference map AA' is a possible interior equilibrium on branch R_2S that is clearly not the global optimum. Since local comparisons of the marginal rates of substitution function and after-tax wage rates cannot be used to determine the state of the world a consumer occupies, some features of the model developed in the previous section no longer apply.

An alternative strategy for determining the portion of the budget constraint on which a consumer decides to locate is the following. Write the direct preference function as $U(X, L, Z\alpha, \nu)$ where the exogenous variable Z represents the measured characteristics of a consumer, α is an unknown vector of parameters, and ν represents the unmeasured characteristics of a consumer that affect preferences. Using well-known methods, one may form the indirect preference function $V(R, W, Z\alpha, \nu)$. For interior solutions, labor supply functions may be written as

$$H = \frac{V_W}{V_R} = H(R, W, Z\alpha, \nu).$$

While the arguments of the functions $U(\cdot)$, $V(\cdot)$, and $H(\cdot)$ may differ across consumers, the functional forms are assumed to be the same for each consumer.

If a consumer is at an interior equilibrium on either segment R_2S or SN, then the equilibrium is defined by a tangency of the indifference curve and the budget constraint. Since this tangency indicates a point of maximum attainable utility, the indifference curve at this point represents a level of utility given by $V(R_i, W_i, Z\alpha, \nu)$ where R_i and W_i are respectively the after-tax unearned income and wage rate associated with segment i. Thus, hours of work at such a point must be given by V_W/V_R evaluated at R_i and W_i. For this point to be an admissible solution, the implied hours of work must lie between the two endpoints of the interval (i.e., equilibrium must occur on the budget segment). A consumer chooses not to work if utility at kink R_2, $U(R_2, T, Z\alpha, \nu)$, is greater than both $V(R_2, W_2, Z\alpha, \nu)$ and $V(R_3, W_3, Z\alpha, \nu)$, provided that these latter utility values represent solutions located on the budget constraint.

We have, then, a general technique for dividing the sample space of the "taste" component ν into the different regions representing the various states. Define the labor supply functions $H_{(1)}$, $H_{(2)}$ and $H_{(3)}$ as $H_{(1)} = 0$ and

$$\text{(45)} \qquad H_{(i)} = \frac{V_W(R_i, W_i, Z\alpha, \nu)}{V_R(R_i, W_i, Z\alpha, \nu)} = H(R_i, W_i, Z\alpha, \nu), \quad i = 2,3;$$

and define the admissible utility levels $V_{(1)}$, $V_{(2)}$ and $V_{(3)}$ as $V_{(1)} = U(R_2, T, Z\alpha, \nu)$, assumed to be greater than zero, and

$$(46) \qquad V_{(2)} = \begin{cases} V(R_2, W_2, Z\alpha, \nu) & \text{if} \quad 0 < H_{(2)} \leq \bar{H} \\ 0 & \text{otherwise} \end{cases}$$

and

$$(47) \qquad V_{(3)} = \begin{cases} V(R_3, W_3, Z\alpha, \nu) & \text{if} \quad \bar{H} < H_{(3)} \leq T \\ 0 & \text{otherwise.} \end{cases}$$

A consumer whose ν lies in the set

$$(48) \qquad \Theta_1 = \{\nu : V_{(1)} \geq V_{(2)} \quad \text{and} \quad V_{(1)} \geq V_{(3)}\}$$

will not work and occupies state 1. If ν lies in the set

$$(49) \qquad \Theta_2 = \{\nu : V_{(2)} > V_{(1)} \quad \text{and} \quad V_{(2)} > V_{(3)}\},$$

a consumer is at an interior solution on segment R_2S and occupies state 2. Finally, a consumer is at equilibrium in state 3 on segment SN if ν is an element of the set

$$(50) \qquad \Theta_3 = \{\nu : V_{(3)} > V_{(1)} \quad \text{and} \quad V_{(3)} > V_{(2)}\}.$$

The sets Θ_1, Θ_2, and Θ_3 do not intersect, and their union is the sample space of ν; thus, these sets are mutually exclusive and exhaustive. The functions $H_{(i)}$ determine the hours of work of an individual whose $\nu \in \Theta_i$.

Given these definitions of the sets Θ_i and the labor supply functions $H_{(i)}$, the analysis of the previous section applies with full force. Assuming

$f(\nu)$ represents the distribution of tastes in the population, the formula for the probability of a consumer occupying state i given by (24) and the formulas for the conditional and the unconditional expected values of labor supply given by (25) and (26) all remain valid. While the sets Θ_i may be difficult to specify in analytical form, the general theory and estimation procedures discussed in the previous section can be applied to analyze the models considered in this section as well. This includes the theory and the estimation procedures relating to the introduction of a wage function or arbitrary forms of measurement error in hours of work and wages.

If wage rates are determined by the function given by (44), where η reflects a randomly distributed error term affecting true market wages, then the definitions

$$\text{(51)} \qquad \Theta_i = \{(\nu, \eta) : V_{(i)} \geq V_{(j)} \quad \text{for all } j,\ j \neq i\}$$

replace the characterization of the sets Θ_i given by (48)–(50). A consumer whose $(\nu, \eta) \in \Theta_i$ chooses to occupy state i. Given these new definitions of Θ_i, the formulas for probabilities of occupying various states given by (41), and the formulas for expected values of labor supply given by (42) and (43) all apply without modification. If an errors in the variables problem exists for hours of work and wages, these formulas still apply except in computing the appropriate expected values it is also necessary to include integration over the measurement error components of hours and wages.

Burtless and Hausman (1978), Hausman (1979), and Wales and Woodland (1979) in their work on labor supply and taxes each use a variant of the general framework described above. All of these studies assume that hours of work are measured with error, and none of them treat wages as endogenous or measured with error.[3]

Burtless and Hausman assume a specific utility function that is monotonic in the unobserved component ν, and for which all consumers are guaranteed to work. In terms of the above model, a consumer occupies either state 2 or state 3 (i.e., the probability of occupying state 1 is zero). For their preference function, there exists a critical value ν^* such that at that value, given the wage rate, tax rate, and intercept income components, the consumer is in equilibrium on both branches. Working with the indirect utility function, V, ν^* is defined such that

$$(52) \qquad V_{(2)} = V(R_2, W_2, Z\alpha, \nu^*) = V(R_3, W_3, Z\alpha, \nu^*) = V_{(3)}$$

and for their special functional form a unique solution for ν^* is guaranteed to exist. For values of ν less than ν^* the consumer is in equilibrium on branch one. For values of ν greater than ν^*, the consumer is on branch two. Thus, ν^* defines the sets Θ_2 and Θ_3, and in this instance these sets are intervals of the real line.

Given their functional form for V, hours of work equations are defined by Roy's identity (see equation (45)). In the Burtless and Hausman world of two working states, the formula for expected labor supply given by (26) may be written as

$$E(H) = \int_{-\infty}^{\nu^*} H_{(2)} f(\nu) d\nu + \int_{\nu^*}^{\infty} H_{(3)} f(\nu) d\nu.$$

Burtless and Hausman estimate the parameters of their model using maximum likelihood procedures. Allowing for measurement error in hours of work, they assume that this measurement error and ν are independently distributed normal random variables. Their procedure, however, may be interpreted as employing the nonlinear least squares method described in the previous section. In particular, letting H^* denote measured hours of work, the criterion

$$\Sigma\left(H^* - E(H)\right)^2$$

is minimized with respect to the parameters of $f(\nu)$ and the parameters of the utility function (which, of course, are also parameters of the labor supply function). Notice that the value of ν^* changes with values of wages, taxes, unearned income, a consumer's measured characteristics, and the parameters of the preference function. Thus ν^* must be updated in any computational algorithm that determines the parameters of the model. Assuming normally distributed optimization error, this criterion is also a maximum likelihood criterion.

Hausman (1979) extends this procedure to allow for corner equilibrium at zero hours of work. Since we have already discussed the essential features of Hausman's model in Section III on fixed costs, we will not repeat that discussion here. As noted in that section, there is a serious problem with Hausman's analysis concerning his treatment of unobserved wages. To deal with the problem of missing wage rates for nonworkers, Hausman introduces a wage function of the sort considered above. He fails, however, to account properly for the presence of unobserved components of wages when constructing the sets Θ_i which define states of the world (see formulas (51)). Hausman estimates parameters of a wage equation corrected for censoring, and then treats the predicted values from this equation as if they were the true values of wages., This procedure is equivalent to assuming that the unobserved components of wages denoted above by η (see equation (44)) are identically equal to zero for all individuals in the sample. Thus, he defines the Θ_i sets only over the sample space of ν, the random component reflecting differences in consumer's "tastes." This leads to improper definitions of these sets if there are any unobserved random components determining wage rates, and this is true even if ν and η are independently distributed. The result is inconsistent estimates of the parameters of the preference and the labor supply functions.

The analyses of Burtless-Hausman and Hausman depend crucially on particular functional forms for preferences. For a general specification of preferences, and unobserved components, the indirect utility function, V, need not be monotonic in ν. More crucially, we are not guaranteed that a ν^* satisfying equation (52) exists, or if one does exist, that it is unique; in this case the sample space of ν cannot be simply partitioned into regions associated with equilibrium along different branches.

In contrast to the above studies, Wales and Woodland (1979) do not allow for unmeasured random differences in consumer's "tastes." In particular, the distribution of ν is assumed to be degenerate at a single point. This implies that the direct and the indirect preference functions and the labor supply function no longer depend on ν; they reduce to $U(X, L, Z\alpha)$, $V(R, W, Z\alpha)$, and $H(R, W, Z\alpha)$. Thus, preferences and labor supply behavior differ across consumers only to the extent that there are differences in measured characteristics.

In terms of the general framework described above, the Wales-Woodland model can be interpreted as one in which the probability of occupying one of the states is one and is zero for all other states. In fact, once a consumer's measured characteristics, wage rate, and unearned income are known, there is no uncertainty regarding his exact location on the budget constraint. Since the variables R, W, and $Z\alpha$ exactly determine a consumer's state of the world, the expression for the unconditional expectation for hours of work given by (37) becomes

$$
(53) \qquad E(H) = \begin{cases} H_{(1)} & \text{if} \quad (R,\ W,\ Z\alpha) \quad \text{imply state 1} \\ H_{(2)} & \text{if} \quad (R,\ W,\ Z\alpha) \quad \text{imply state 2} \\ H_{(3)} & \text{if} \quad (R,\ W,\ Z\alpha) \quad \text{imply state 3.} \end{cases}
$$

Each $H_{(i)}$, then, is a nonstochastic function of R, W, and $Z\alpha$ that is valid for only certain combinations of these variables. For other combinations, E(H) shifts to a different $H_{(j)}$ function associated with the state implied by the new combinations of observed variables.

The Wales-Woodland estimation procedure is to choose the parameters of the preference function to minimize

$$\Sigma\left(H^* - E(H)\right)^2$$

where the summation is over consumers, and H^* is measured hours of work (which need not be the same as true hours of work, H). By searching over potential values of the parameters of the preference function, one chooses the set of values that minimizes the above least squares criterion. In computing the value of this criterion (for a given set of parameter values), the following procedure is used to choose the labor supply function (i.e., E(H)) relevant for each consumer.

Given the functional form for preferences, the exogenous variables of the model (i.e., a consumer's measured characteristics), and values for the parameters of the preference function, one can compute the utility of the consumer at each point of the budget constraint--in particular, along each branch and at each kink point (including the no-work point R_2). By checking utility on each branch and verifying the existence of an interior solution, and by checking utility at each kink, one can literally solve the consumer's maximization problem by choosing that labor supply function associated with the highest utility. For each set of values for the parameters of the utility function, it is possible to calculate the unique equilibrium position, either interior or corner, for each consumer. As the parameter values change, the computed equilibrium hours of work changes.

When taxes are progressive (the only case Wales-Woodland consider), this procedure boils down to the following algorithm. For each value of α, one can compute the hours of work implied for each branch using the labor supply function associated with that branch implied by Roy's Identity. Thus, for the first branch (i.e., state 2) hours of work are given by $H_{(2)}$. If the predicted $H_{(2)}$ lies in the interval $(0, \bar{H}]$, the consumer's equilibrium position is assumed to lie along the first branch and E(H) is set equal to $H_{(2)}$.

Suppose that for the particular value of α under consideration, $H_{(2)}$ lies outside the interval $(0, \bar{H}]$, then the equilibrium position is not on the first branch; so, one checks the second branch. For the same α, one can compute the implied hours of work for this branch, $H_{(3)}$. If the $H_{(3)}$ lies in the interval $(\bar{H}, T]$, the consumer is assumed to be in equilibrium at an interior point on the second branch, and $E(H) = H_{(3)}$. Otherwise he is not. The same procedure can be used for a problem with more than two branches. From convexity of preferences and constraints, for each α, equilibrium can occur on only one branch at most.

Suppose that we never predict an H that lies in the correct interval. This can happen if there is a corner equilibrium. Each corner should then be checked including the no-work corner. Evaluate U at each corner. Pick the corner with the highest utility. This is the equilibrium value, and E(H) is set equal to that labor supply function associated with the kink. This procedure is guaranteed to locate the consumer's optimum for each value of α.

This procedure is immodest in that it assumes that given α, the econometrician can solve the consumer's problem as well as the consumer can. There is no information the consumer has that is relevant to selecting his desired hours of work that the econometrician does not know as well. There are no omitted variables in the model. Moreover, the model assumes that all consumers have the same α.

These are strong assumptions. If there are such omitted variables, or if α differs across consumers, the Wales-Woodland procedure generates inconsistent parameter estimates because their procedure can allocate consumers to the wrong branch of the budget constraint. Surprisingly, less restrictive assumptions that allow for random differences in unmeasured traits affecting "tastes" may lead to a simpler estimation scheme, such as the one proposed by Burtless and Hausman outlined above.

VI. Conclusion

The index function framework forms the basis for much of the recent work on labor supply. The use of this framework provides a convenient approach for the estimation of hours of work functions when budget constraints consist of several segments and kinks. The methodology presented in the last section is the most general. It not only admits both convex and non-convex constraint sets, it also allows for the endogeneity of wage rates, for the absence of wage data for some of the sample observations, and for arbitrary forms of measurement error in hours of work and wages.

APPENDIX

The Performance of Various Estimators in the Presence of Sample Selection Bias

Woodland and Wales (1978) report some sampling experiments with alternative estimators for labor supply and wage equations in the presence of sample selection bias. The model they consider is a variant of Heckman (1974a).

Their hours of work equation is

$$h = \beta_0 + \beta_1 w + \beta_2 x_2 + u \qquad \text{if } h > 0$$

$$h = 0 \qquad \text{otherwise.}$$

Their wage function is

$$w = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + e.$$

Wages are observed only if $h > 0$. In their sampling experiments u and e are assumed to be mean zero normally distributed variables with variances σ_u^2 and σ_e^2 respectively and interequation correlation of ρ.

The following methods are compared: (a) ordinary least squares (OLS); (b) maximum likelihood; (c) Amemiya's estimator (1973); (d) nonlinear least squares; (e) Heckman's maximum likelihood estimator (1974); (f) Heckman's two step estimator (1976, 1979). For a complete and thoroughly competent discussion of these estimators, see Wales and Woodland (1978). Table 1 reports on estimators based on samples of 5000 individuals. The number of workers is indicated by M. The true parameters are given in the first column. Least squares results are shown in the second column to be badly biased. The maximum likelihood estimators (method 1 and method 7) generate estimates quite close to the true values. Amemiya's procedure (method 2) and a one round iterate of Amemiya (method 2A which is one-Newton step toward the likelihood optimum from the initial consistent Amemiya estimator) is badly behaved. A nonlinear least squares procedure (method 3) produces estimates that are badly biased.

Essentially the same results are found for estimators that use the entire sample (see Table 2). OLS is badly biased. Maximum likelihood (method 4) generates estimates quite close to the true values. Two stage methods are less precise in generating parameter estimates.

Corrections for selection bias in the wage equation (results reported in Table 3) suggest that the two stage procedures and the maximum likelihood procedure do about equally well, but maximum likelihood is still marginally better.

It is important to note that these samples are generated under ideal conditions. It would be very interesting to compare the performance of these estimators, which are based on the normality assumption, in the situation in which the errors are nonnormal. It seems likely that the two stage estimators are more robust because the conditional means of the errors may still closely approximate the true conditional mean (i.e., the g term in equation (5)). Little is known about the performance of these estimators in the presence of other model misspecifications, but by analogy with the findings in the simultaneous equation literature it is likely that the two stage estimators are more robust to misspecification than are the maximum likelihood estimators.

TABLE 1: HOURS EQUATION ESTIMATES BASED ON THE SUB-SAMPLE OF WORKERS

	True	OLS	Method 1	Method 2	Method 2A	Method 3	Method 7
			$\rho = 0$	(M = 1634)			
β_0	-1.1854	.359	-1.027	-.185	-.508	-.591	-1.083
		(.044)	(.120)	(.358)	(.080)		(.115)
β_1	1.0	.596	.943	.790	.840	.847	.955
		(.017)	(.034)	(.066)	(.026)		(.038)
β_2	1.0	.625	.987	.841	.885	.889	1.011
		(.026)	(.047)	(.077)	(.037)		(.042)
σ_u	1.2472	.983	1.232	.925	1.057	1.035	1.234
			(.033)	(.160)	(.021)		(.034)
			$\rho = 0.5$	(M = 1581)			
β_0	-1.1854	-.073	-1.243	-1.691	-.725	-1.436	-1.103
		(.042)	(.095)	(.453)	(.063)		(.093)
β_1	1.0	.822	1.157	1.256	1.037	1.211	.994
		(.018)	(.031)	(.094)	(.023)		(.033)
β_2	1.0	.744	1.057	1.115	.969	1.096	1.016
		(.022)	(.037)	(.096)	(.028)		(.027)
σ_u	1.0393	.853	.997	1.139	.855	1.062	1.011
			(.024)	(.137)	(.016)		(.026)
			$\rho = -0.5$	(M = 1626)			
β_0	-1.1854	.860	-1.071	.237	-.146	-.347	-1.225
		(.044)	(.179)	(.659)	(.095)		(.160)
β_1	1.0	.339	.695	.529	.564	.576	.959
		(.015)	(.038)	(.035)	(.026)		(.049)
β_2	1.0	.477	1.010	.771	.819	.849	1.042
		(.029)	(.068)	(.129)	(.047)		(.063)
σ_u	1.7283	1.127	1.594	1.118	1.280	1.314	1.711
			(.053)	(.291)	(.028)		(.067)

Notes to Table 1

Method 1 estimates are achieved by maximizing the conditional likelihood function for hours of work.

Method 2 is based on Amemiya's estimator for truncated samples (Amemiya, 1973).

Method 2A is one-Newton-Raphson iterate of the likelihood function used in method 1 using the Amemiya initial consistent estimator.

Method 3 is a nonlinear least squares procedure that jointly estimates the parameters of the regression function and the conditional mean of the random sample disturbance (for details, see Wales and Woodland, 1978, who propose this procedure).

Method 7 is Heckman's (1974a) maximum likelihood estimator for hours and wages conditioned for samples of workers. This estimator is proposed by Wales and Woodland (1978).

TABLE 2: HOURS EQUATION ESTIMATES BASED ON THE ENTIRE SAMPLE

	True	OLS	Method 4	Method 5A	Method 5B	Method 5C	Method 5D
				$\rho = 0$			
β_0	-1.1854	.261	-1.168	-.798,-.799	-.945,-.946	-.782	-.921
		(.015)	(.053)			(.670)	
β_1	1.0	.398	.984	.900, .876	.941, .917	.883	.921
		(.010)	(.023)			(.159)	
β_2	1.0	.329	1.028	.931	.963	.927	.956
		(.011)	(.027)	(.178)		(.176)	
μ_1	1.6934		1.6847*	1.416	1.538	1.403	1.519
				(.625)		(.616)	
σ_u	1.2472	.906	1.246				
			(.027)				
				$\rho = 0.5$			
β_0	-1.1854	.243	-1.217	-1.481,-1.481	-1.693,-1.694	-1.491	-1.663
		(.015)	(.040)			(.913)	
β_1	1.0	.400	1.008	1.082, 1.074	1.129, 1.147	1.081	1.129
		(.010)	(.018)			(.219)	
β_2	1.0	.335	1.014	1.099	1.144	1.102	1.136
		(.011)	(.021)	(.222)		(.221)	
μ_1	1.6934		1.707*	1.822	1.987	1.830	1.964
				(.805)		(.802)	
σ_u	1.0393	.909	1.047				
			(.021)				
				$\rho = -0.5$			
β_0	-1.1854	.271	-1.207	-.569,-.569	-.825,-.825	-.568	-.823
		(.015)	(.074)			(.650)	
β_1	1.0	.370	.957	.803, .801	.875, .854	.802	.854
		(.009)	(.032)			(.151)	
β_2	1.0	.321	1.024	.860	.927	.860	.927
		(.011)	(.031)	(.167)		(.167)	
μ_1	1.6934		1.720*	1.269	1.467	1.268	1.466
				(.591)		(.590)	
σ_u	1.7283	.899	1.705				
			(.048)				

*This estimate is derived from the others.

Notes to Table 2

Method 4 is Heckman's (1974a) maximum likelihood estimator for hours and wages.

Method 5A is the Heckman (1976, 1979) indirect least squares two-step estimator. The multiplicity of estimates for β_1 arises because this parameter is overidentified in the current model.

Method 5B is a GLS version of 5A.

Method 5C is the Heckman procedure as modified by Wales and Woodland (1978) or Heckman (1978) to resolve the overidentification problem.

Method 5D is a GLS version of 5C.

"μ_1" is the estimated conditional mean of the disturbance term in the hours of work equation. (This corresponds to g in equation (5).)

TABLE 3: WAGE EQUATION ESTIMATES

	True	OLS	Method 4	Method 5A	Method 5B	Method 7
			$\rho = 0$			
γ_0	0	.752	-.030	-.043	-.049	.024
		(.034)	(.045)	(.106)		(.080)
γ_1	1.0	.780	.997	.993	.993	.964
		(.026)	(.026)	(.039)		(.037)
γ_2	1.0	.821	1.019	1.026	1.030	1.011
		(.025)	(.026)	(.039)		(.037)
μ_2	.7749		.781*	.802	.807	
				(.128)		
σ_e	1.1455	1.044	1.160	1.164		1.155
			(.024)			(.027)
			$\rho = 0.5$			
γ_0	0	.783	-.006	.065	.062	.091
		(.025)	(.031)	(.089)		(.068)
γ_1	1.0	.794	.979	.970	.967	.974
		(.018)	(.018)	(.031)		(.029)
γ_2	1.0	.781	.987	.966	.968	.968
		(.018)	(.019)	(.032)		(.029)
μ_2	.7749		.774*	.703	.707	
				(.107)		
σ_e	0.9148	.753	.913	.882		.878
			(.019)			(.023)
			$\rho = -0.5$			
γ_0	0	.760	.043	-.039	-.030	.048
		(.048)	(.068)	(.140)		(.095)
γ_1	1.0	.827	1.002	1.023	1.022	1.006
		(.035)	(.035)	(.051)		(.047)
γ_2	1.0	.814	.984	1.006	1.004	.988
		(.037)	(.035)	(.050)		(.047)
μ_2	.7749		.728*	.805	.796	
				(.152)		
σ_e	1.6564	1.525	1.598	1.614		1.593
			(.031)			(.032)

*This estimate is derived from the others.

Notes to Table 3

All of the procedures have been defined in Tables 1 and 2. "$\hat{\mu}_2$" is the estimated conditional mean of the disturbance term in the wage equation.

Footnotes

Section II

[1]For expositional simplicity, we abstract from endogenous wages in the statistical analysis presented below. This assumption is not essential to the analysis.

[2]Defining $H_W = 0$ for nonworkers, it is straightforward to verify that the partial derivative of expected hours with respect to W from equation (11) exceeds the mean value of H_W in the population. Defining $H_R = 0$ for non-workers, if leisure is a normal good the partial derivative of expected hours with respect to R is smaller than the mean value of H_R in the population.

[3]To simplify notation, we supress Z_2 (i.e., a consumer's measured characteristics) as an argument of the functions $U(\cdot)$, $M(\cdot)$ and $H(\cdot)$.

Section III

[1]Recall that we are assuming that W is exogenous and is available for each individual in the sample exposition. Cogan (1979) allows for endogenous wages and his approach does not require wage data for nonworkers.

[2]Because of the assumptions about functional forms, any set of exogenous variables, including those that enter the utility function in their own right, serve to identify the effect of fixed costs on hours of work. Thus, as a consequence of his functional form, Hausman does not require the exclusion restriction needed by Cogan to identify the effect of fixed costs on labor supply.

Section IV

[1]In this section, we abstract from the very important problem that true taxes are not measured by published tax schedules, even schedules that carefully set out exemptions and deductions. If consumers spend real resources to avoid paying taxes, such tax avoidance costs are properly considered as part of the effective tax. With the exception of preliminary work of Gould (1979), this problem has not received attention in the literature.

[2]Instruments will be available provided that variables appearing in the wage function do not appear in the structural labor supply function. Such variables are valid instruments.

[3]The procedure suggested by Hall (1973) and Rosen (1976) that evaluates the marginal tax rates and intercept terms at a standard number of hours of work for everyone in the sample, generates inconsistent parameter estimates because E' and $E - E'H$ are evaluated at the wrong point.

[4]However, we have no guarantee in this problem that the constraint set will be so characterized. See Rosen (1974).

[5]Pellechio (1979) proposes estimating a model with kinked convex constraints in essentially this fashion.

Section IV (continued)

[6]Letting g(e) denote the density function of e, these probabilities are related by the equation

$$\begin{aligned}\text{Prob }(C_1 \le H^* \le C_2) &= \text{Prob }(C_1 \le H+e \le C_2)\\ &= \int_{-\infty}^{\infty} \text{Prob}(C_1 \le H+e \le C_2 | e) g(e) de\\ &= \int_{-\infty}^{\infty} \text{Prob}(C_1 - e \le H \le C_2 - e | e) g(e) de\\ &= \int_{-\infty}^{\infty} \text{Prob}(C_1 - e \le H \le C_2 - e) g(e) de\end{aligned}$$

where the last line follows from the independence assumption of H and e.

[7]It is clear that one need not rigidly chain the parameters of the hours of work function to the F functions to secure identification. Thus, one could estimate the parameters of equation (37) by restricting the parameters of the F and λ functions to be the same, and not exploiting the theory to generate the relationship between these parameters and α_1, α_2, and α_3. However, even though the model is formally identified, it is likely that parameter estimates obtained by this procedure would be imprecisely estimated. Exploiting all the restrictions of the theory requires making strong assumptions about functional form. But if these assumptions are not made, parameter estimates are likely to be imprecise.

[8]Notice that the arguments W_2, W_4, and R_4 each depend on W.

Section V

[1]Generalization to more than two branches involves no new principle. Constraint sets like R_2SZ are common in negative income tax experiments and in certain social programs.

[2]The kink at S is not treated as a state of the world because there is not a positive probability that a consumer will be at this point if the "taste" component ν follows a continuous density. In fact, for most preference functions, point S can never be an equilibrium.

[3]Hausman (1979) introduces and estimates a wage equation, but, as we discuss later in this section, his estimation procedures do not properly treat unobserved components of wages.

REFERENCES

Amemiya, T. "Qualitative Response Models." Annals of Economic and Social Measurement 4 (Summer 1975): 363-372.

__________. "Regression Analysis When the Dependent Variable is Truncated Normal." Econometrica (November 1973).

Ben-Porath, Yoram. "Labor Force Participation Rates and the Supply of Labor." Journal of Political Economy 81 (May-June 1973): 697-704.

Borjas, G., and Heckman, J. "Labor Supply Estimates for Public Policy Evaluation." Proceedings of the Thirty-first Annual Meeting of the Industrial Relations Research Center (1978): 320-331.

Boskin, Michael J. "The Economics of Labor Supply." In Glen G. Cain and Harold W. Watts, eds., Income Maintenance and Labor Supply. Chicago: Markham, 1973.

Burtless, G., and Hausman, J. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experimant." Journal of Political Economy 86, no. 6 (December 1978): 1103-1131.

Cain, G., and Watts, H. Income Maintenance and Labor Supply. Chicago: Markham Publishers, 1973.

Cogan, J. "Fixed Costs and Labor Supply." The Rand Corporation, May 1979. Econometrica, forthcoming.

__________. "Labor Supply with Time and Money Costs of Participation." The Rand Corporation, R-2044-HEW, October 1977.

DaVanzo, J.; Detray, D.; and Greenberg, D. "Estimating Labor Supply Response: A Sensitivity Analysis." The Rand Corporation, R-1372-OEO, December 1973.

Domencich, M., and McFadden, D. Urban Travel Demand: A Behavioral Analysis. Amsterdam: North-Holland Publishing Co., 1975.

Gould, W. "Taxes and Female Labor Supply." Ph.D. dissertation, University of California, Los Angeles, 1979.

Gronau, Reuben. "Wage Comparisons--A Selectivity Bias." Journal of Political Economy 82, (November-December 1974): 1119-1144.

Hall, Robert E. "Wages, Income and Hours of Work in the U.S. Labor Force." In Glen G. Cain and Harold W. Watts, eds., Income Maintenance and Labor Supply. Chicago: Markham, 1973.

Hanoch, G. "A Multivariate Model of Labor Supply: Methodology for Estimation." Published as a Rand Report in 1976, and in J. P. Smith, ed., Female Labor Supply: Theory and Estimation. Princeton, New Jersey: Princeton University Press, 1979.

Hausman, J. "The Effect of Wages, Taxes and Fixed Costs on Women's Labor Force Participation." Massachusetts Institute of Technology, May 1979.

Hausman, J., and Wise, D. "The Evaluation of Results from Truncated Samples: The New Jersey Negative Income Tax Experiment." Annals of Economic and Social Measurement 5 (Fall 1976): 421-446.

__________. "Social Experimentation, Truncated Distributions, and Efficient Estimation." Econometrica (May 1977): 919-938.

Heckman, J. "Shadow Prices, Market Wages and Labor Supply." Econometrica (July 1974). (a)

__________. "The Effect of Child Care Programs on Women's Work Effort." Journal of Political Economy (March-April 1974). (b)

__________. "The Common Structure of Statistical Models of Truncation, Sample Selection Bias and Limited Dependent Variables and a Simple Estimator for Such Models." Annals of Economic and Social Measurement (December 1976).

__________. "Dummy Endogenous Variables in a Simultaneous Equation System." Econometrica (July 1978).

__________. "Sample Selection Bias as a Specification Error." Econometrica (January 1979).

Johnson, T. "Qualitative and Limited Dependent Variables in Economic Relationships." Econometrica (May 1972): 440-455.

Lewis, H. Gregg. "On Income and Substitution Effects on Labor Force Participation." University of Chicago, 1967.

__________. "Interes del empleador en las horas de Trabajo del empleado" [Employer Interests in Employee Hours of Work]. Cuadernos de Economia, Catholic University, Chile, 1969.

McFadden, Daniel. "Quantal Choice: A Survey." Annals of Economic and Social Measurement 5 (Fall 1976): 363-390.

__________. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zaremba, ed., Frontiers in Econometrics. Academic Press, 1974.

Pellechio, A. "The Estimation of Labor Supply over Kinked Budget Constraints: Some New Econometric Methodology." Working Paper #387, National Bureau of Economic Research, Cambridge, August 1979.

Rosen, H. S. "Taxes in a Labor Supply Model with Joint Wage-Hours Determination." Econometrica (May 1976): 485-508.

Rosen, S. "Hedonic Prices and Implicit Markets." Journal of Political Economy (January 1974).

Rosett, Richard N., and Nelson, Forrest D. "Estimation of the Two-Limit Probit Regression Model." Econometrica 43 (January 1975): 141-146.

Smith, James P., ed. Female Labor Supply: Theory and Estimation. Princeton, New Jersey: Princeton University Press, forthcoming, 1981.

Tobin, J. "Estimation of a Model with Limited Dependent Variables." Econometrica (1958).

Wales, T. J., and Woodland, A. D. "Labour Supply and Progressive Taxes." Review of Economic Studies 46 (January 1979): 83-95.

__________. "Sample Selectivity and the Estimation of Labour Supply Functions." University of British Columbia, November 1978.