Federal Reserve Bank of Minneapolis
Research Department

A Structural Model of Multiple Welfare Program Participation and Labor Supply

Michael Keane and Robert Moffitt*

Working Paper 557

October 1995

ABSTRACT

One of the long-standing issues in the literature on transfer programs for the U.S. low-income population concerns the high cumulative marginal tax rate on earnings induced by participation in the multiplicity of programs offered by the government. Empirical work on the issue has reached an impasse partly because the analytic solution to the choice problem is intractable and partly because the model requires the estimation of multiple sets of equations with limited dependent variables, an estimation problem which until recently has been computationally infeasible. In this paper we estimate a model of labor supply and multiple program participation using methods of simulation estimation that enable us to solve both problems. The results show asymmetric wage and tax rate effects, with fairly large wage elasticities of labor supply but very inelastic responses to moderate changes in cumulative marginal tax rates, implying that high welfare tax rates do not necessarily induce major reductions in work effort.

*Keane, Federal Reserve Bank of Minneapolis and University of Minnesota; Moffitt, Brown University. The authors would like to thank Zvi Griliches, two anonymous referees, the participants at the Econometrica/SSRI conference on "Empirical Applications of Structural Models," Madison, May 1990, and the participants at several other conferences and university workshops for comments. Support from the Federal Reserve Bank of Minneapolis is also appreciated. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

The empirical literature on the effects of transfer programs on labor supply that has evolved over the past twenty years is fairly large (see Danziger et al, 1981 and Moffitt, 1992, for surveys). Most of the literature has applied the static labor supply model to the problem, and much of the work has been concerned with the proper treatment of the nonlinear budget constraints generated by most transfer programs (Hausman, 1985; Moffitt, 1986). The characteristic of past work which the present paper addresses is instead its concentration on the analysis of individual transfer programs, one at a time and in isolation from others, and its consequent failure to address the issues which arise in the analysis of multiple transfer programs.

The possible work disincentives and other inefficiencies created by the existence of multiple transfer programs is one of the most important, and one of the oldest, issues in the economics of transfer programs in the U.S. It is now the rule, rather than the exception, that families who participate in one program simultaneously participate in one or more other programs. In 1984, for example, 89 percent of recipients of transfers from the Aid to Families with Dependent Children (AFDC) program also received both Food Stamps and Medicaid benefits, and another 42 percent of these also received a fourth benefit, most often subsidized housing (Moffitt, 1992, Table 2). The situation was considerably different in the 1960s, before many programs now in existence were introduced. Currently, with marginal tax rates of 100 percent in AFDC, 30 percent in Food Stamps, and up to 30 percent in subsidized housing, and with Medicaid benefits lost in their entirety when AFDC eligibility ends--creating a budget-constraint "notch" with a tax rate in excess of 100 percent--the potential for work disincentives is clearly very high.

Despite the importance of the problem, there have been virtually no studies of the labor supply effects of multiple program participation, even within the conventional static labor supply model (much less within dynamic models). There are two reasons for this lack of progress. First, the existence of self-selection into different program combinations on the basis of unobserved heterogeneity components such as welfare stigma and other factors implies that the labor supply equation must be estimated jointly with a set of program participation equations. Yet the joint estimation of large numbers of equations with limited dependent variables has until recently been computationally infeasible. Second, and more special to structural modeling ("structural" in a sense to be precisely defined below), imposing a utility structure on the problem results in a choice problem whose analytic solution is intractable because the regions of the error space within which different program combinations are optimal are too complex to derive. As a result, for example, one of the few studies to address the multiple program problem (Fraker and Moffitt, 1988) estimated reduced-form equations even though the model was sufficiently small in size that the first problem (computation of the necessary multiple integrals) was entirely feasible.

For these two reasons there have been no studies to date which estimate a structural model of labor supply and multiple transfer program participation. In this paper we apply simulation estimation methods to solve the problem. In simulation estimation, choice probabilities are simulated using Monte Carlo methods rather than evaluated by conventional numerical methods. Simulation not only solves the problem of evaluating the multiple integrals arising from sets of limited-dependent-variable equations, as is well-known, but it also solves the problem of the intractable analytic solution to the choice problem. The latter solution is possible because the estimation of the relevant choice probabilities through simulation does not require an analytic derivation of the boundaries of the error space within which different program combinations are optimal.

We go beyond existing work on simulation methods by investigating the special issues which arise in selection models, for in our case we include a wage equation which is observed only for workers. We examine the relative advantages of using simulated maximum likelihood, or SML (Albright et al., 1977; Lerman and Manski, 1981) vs. the method of simulated moments, or MSM

(McFadden, 1989; Pakes and Pollard, 1989) in this context. For the latter, we illustrate the inappropriateness of conventional MSM methods for selection models and instead employ a modified MSM method based on a suggestion of McFadden and Ruud (1994) for such models. We find that the modified MSM method, although consistent in fixed simulation size while SML is not, is nevertheless computationally much more burdensome than SML when applied to selection models. We therefore test a new two-step estimator in which first-stage wage-equation estimates are used in a second-stage MSM estimation. We find this new method to dominate the two other methods in computational efficiency.

We estimate a model with 1984 data on female heads of family who choose from among three transfer programs--AFDC, Food Stamps, and subsidized housing.[1] Together with the labor supply equation, a four-equation model results. The method we employ generalizes to any number of programs. Section I below lays out the problem, while Section II outlines the solution method using simulation techniques. Section III discusses the data set we employ and how we construct the budget constraints for the individuals in the sample and Section IV presents the results. Section V presents simulations of the effects of reducing cumulative marginal tax rates on labor supply and of other policy programs (such as wage subsidies), as well as an out-of-sample comparison of the actual effect of a historical change in the AFDC tax rate to the effects predicted from our model.

I. The Estimation Problem in a Model of Labor Supply and Multiple Program Participation

Consider the problem of analyzing the effect of the availability of M different welfare programs on the labor supply of a population of individuals eligible for those programs. Each program bases its payment to a recipient on a particular "guarantee" amount--the payment to a family with no earnings and no income from any other transfer program--and on a set of "tax rates" which denote the amount by which the payment is reduced for each extra dollar of earnings and each extra dollar of benefits from other transfer programs. The tax rates on earnings and on other-program benefits may differ within and across the benefit formulas for different programs.

The analysis of labor supply response in this environment is relatively simple in principle. If the utility function is of the conventional form U(H,Y), where H is hours of work (a "bad") and Y is disposable income (a "good"), a straightforward approach would require simply computing the budget constraint (the value of Y for each value of H) for each of the 2^M possible program combinations and then deriving the envelope of these constraints. Labor supply choice in the face of the envelope constraint could then be estimated by methods appropriate to choice subject to what would presumably be a nonlinear budget locus. Unfortunately, this approach fails because of the existence of significant numbers of non-participating eligibles--families who could apparently increase their net income without changing hours of work but do not do so because they do not locate on the envelope of the constraint. Such behavior has been attributed to welfare stigma (Moffitt, 1983), but may also arise from a more general disutility of dealing with welfare bureaucracies or from time and money costs of program participation (taking time off to visit the welfare office, etc.).[2] The econometric difficulty created by this phenomenon arises because the unobservables affecting program participation are likely to be correlated with the unobservables affecting labor supply--for example, those most likely to participate are likely those with the lowest tastes for work. Consequently, labor supply choice cannot be estimated conditional on an assumed exogenously-chosen combination of programs in which the individual participates. The program-participation combination choice must be treated jointly with labor supply choice, and hence the labor supply equation must be estimated jointly with a set of program participation equations.

That the choice problem remains relatively simple but the estimation problem does not can be illustrated with the following model. Suppose the utility function is $U(H,Y,P_1,P_2,\ldots,P_m)$, where P_m is a dummy variable equal to

1 if the individual participates in program m and 0 if not. The presence of participation indicators in the preference function can be interpreted as representing either stigma influences or, more generally, as costs of participation (neither money nor time costs are directly measured in most data sets). Assume $\partial U/\partial H < 0$, $\partial U/\partial Y > 0$, and $\partial U/\partial P_m < 0$. For purposes of illustration consider the separable case:

$$U(H,Y,P_1,P_2,\ldots,P_m) = \bar{U}(H,Y) - \sum_{m=1}^{M} \psi_m P_m \tag{1}$$

where each ψ_m denotes the marginal disutility of participating in program m. Thus if ψ_m is sufficiently large, a particular program may not be chosen even though participation increases $\bar{U}$.

It is convenient both analytically and empirically to consider the choice of discrete H points rather than continuous H. Therefore consider the choice of H=0, H=20, and H=40 per week, taken as the choice of nonwork, part-time work, and full-time work, respectively. This approach, taken before in the transfer program literature (Zabalza et al.,1980; Fraker and Moffitt, 1988), avoids the task of computing the locations of the numerous segments and kinks of each of the 2^M budget constraints that would be required if continuous H were modeled. In the present context, discretization of H is particularly convenient because it allows us to model agents as facing a multinomial choice problem with a set of discrete participation-hours alternatives.

With three H points and M transfer programs there are $3*2^M$ discrete choices available to the individual. The budget constraint gives disposable income for each:

$$Y(H,P_1,P_2,\ldots,P_m) = wH + N + \sum_{m=1}^{M} P_m B_m(H) - T(H) \tag{2}$$

where w is the hourly wage rate, N is nontransfer nonlabor income, $B_m(H)$ is the benefit function for program m, and T(H) is the positive tax function. For notational simplicity we have suppressed the dependence of the functions $B_m(H)$ and T(H) on N, wH, and the benefits of other transfer programs in which the individual might participate (as well as individual and family characteristics that may affect taxes or transfers).

Letting $j=1,\ldots,3*2^M$ index alternatives from the choice set, the choice problem is simply:

$$\text{Choose alternative } j \text{ iff } U_j \geq U_k \quad \forall \quad k=1,\ldots,3*2^M \tag{3}$$

where U_j denotes the evaluation of (1) for combination j obtained by inserting (2) evaluated at that combination into (1) and by setting H and the P_m at their appropriate values for combination j.

The estimation problem can be seen by assuming a stochastic specification in which a labor-supply preference parameter α varies across the population--hence utility is $\overline{U}(H,Y;\alpha)$--and the M parameters ψ_m vary across the population as well; the parameters vary for reasons unobserved to the investigator but known to the individuals. Estimation of the resulting M+1 equations of the model by conventional ML methods must confront, first, a well-known problem of evaluating integrals of order (M+1) for the computation of the probabilities; such evaluation is in general infeasible with typical quadrature methods if the covariance matrix of the errors is unrestricted. Second, however, the analytical problem is intractable in any case because calculation of the limits of integration of the necessary integrals requires that the regions of the (M+1) dimensional error space within which each of the program combinations is optimal be determined analytically. Yet this requires that the N(N-1)/2 (where $N=3*2^M$) unique hyperplanes in the error space defined by $U_j-U_k=0$ for $j,k=1,\ldots\ldots,3*2^M$ be calculated along with their intersections and areas of dominance (many program combinations will be completely dominated in some ranges of the error space, for example). Determining these regions is a practical impossibility.[3] We note immediately that this problem arises because of the imposition of a particular utility structure on the problem and from the factor structure created by the M+1 underlying errors. A model without such utility structure imposed could be estimated more straightforwardly, providing the first problem (of evaluation of multiple integrals) could be solved.[4]

Both of these problems are solved by simulation estimation. It is well-known that this method can solve the first problem but less recognized in the literature on welfare programs and similar problems that it can solve the second problem as well. To simulate choice probabilities, the regions of the space over which integration is taken need not be analytically derived; it need only be determined which of the participation possibilities has greatest utility for each draw from the distribution of the errors. The details of the computational algorithm are discussed below.

II. The Empirical Application and the Estimation Method

Empirical Application. Our application is to the labor supply and program participation decisions of female heads of family, the demographic group eligible for most U.S. welfare programs. We study their participation decisions with regard to three major programs: AFDC, Food Stamps, and subsidized housing. A fourth program, Medicaid, is not included in the choice set but is included as a benefit automatically conferred by the choice of AFDC (details given below). We therefore have M=3 and we set m=A for AFDC, m=F for Food Stamps, and m=R for subsidized (rental) housing. With three categories of H the number of alternatives is 24.

For the utility function in (1) above, we assume a conventional flexible-form, quadratic function representing a second-order Taylor series expansion in its arguments:

$$U(H,Y,P_1,P_2,\ldots,P_m) = \alpha H + Y - \beta_{HH}H^2 - \beta_{YY}Y^2 - \sum_{m=1}^{M} \psi_m P_m$$
$$+ \sum_{m=1}^{M} \sum_{n>m}^{M} \phi_{mn} P_m P_n + \beta_{HY} HY - \sum_{m=1}^{M} \delta_m H P_m - \sum_{m=1}^{M} \eta_m Y P_m \quad (4)$$

The marginal utility of Y at Y=0 is normalized to 1; the remaining parameters are therefore in dollar terms. We also permit interactions between the different participation programs (ϕ_{mn}); the parameters on these terms will be non-zero if the implicit cost of multiple program participation differs from the simple sum of the costs of participating in each singly (e.g., if the stigma or other costs of participation are less than proportionate to the number of programs); we will discuss this aspect of our specification more below. Our multinomial choice model therefore consists of (4) and (2), with solution (3).

Our stochastic structure permits α and the ψ_m to vary in the population conditional on a set of observable socioeconomic characteristics:

$$\alpha = \underset{\sim}{X}'\underset{\sim}{\overline{\alpha}} + \epsilon_\alpha \quad (5)$$

$$\psi_m = \underset{\sim}{X}'\underset{\sim}{\overline{\psi}}_m + \epsilon_m \quad , \quad m=A,F, \text{ or } R \quad (6)$$

where $\underset{\sim}{X}$ is a vector of characteristics and $\underset{\sim}{\overline{\alpha}}$ and $\underset{\sim}{\overline{\psi}}_m$ are vectors of coefficients. The parameter α represents the marginal disutility of work at H=0, and the parameters ψ_m represent the marginal disutilities, or costs, of program participation if there are no higher-order interactions in the preference function. We choose these parameters to be stochastic because they appear linearly in the pairwise utility differences among the 24 categories. That is, differencing (4) across alternatives with the same P_m but different H and Y gives choice equations for H that are linear in α, and differencing (4) across alternatives with the same H but different P_m and Y gives choice equations for P_m that are linear in the ψ_m. While linearity in errors is not necessary for computation whe simulation methods are used, it does increase the comparability of our specification with past work.

Our full model therefore can be derived by inserting (2), (5), and the three equations in (6) into (4). There are 24 possible combinations of the choice variables, and hence 24 "equations," and there are four error terms. Our model is "structural" in the sense that it has a particular factor structure of the errors that arises from the imposition of a particular utility function (albeit one with flexible form) and a presumption that the major source of variation in choices arises from heterogeneity in a selected set of preference parameters.

Since wage rates are unobserved for nonworkers, we specify a log wage equation as:

$$\ln(w) = \underline{X}'\underline{\nu} + \epsilon_w \tag{7}$$

We initially estimate (7) jointly with the labor-supply-participation choice model. However, we also estimate models which use predicted wages from first-stage estimates of (7) in a second-stage estimation of the choice model alone.

The five error terms in the model (ϵ_α, ϵ_A, ϵ_F, ϵ_R, and ϵ_w) are assumed to be distributed multivariate normal with an unrestricted covariance matrix with diagonal elements σ_j^2, $j=\alpha,A,F,R,w$, and with off-diagonal elements $\rho_{jk}\sigma_j\sigma_k$, $j,k=\alpha,A,F,R,w$. The elements of the covariance matrix are all identified by the normalizations in the model and the other parameters of the model are identified by the distributional assumptions and various nonlinearities.[5] However, to lessen their dependence on those functional form assumptions, we also impose exclusion restrictions in equations (5)-(7). We employ variables that affect program participation but not labor supply to identify the parameters of $\overline{U}$, by excluding some variables that are in (6) from (5). Also, as in more standard wage-labor-supply models, we exclude some variables in (5) from (7) and some variables in (7) from (5) in order to identify the parameters of each. We should also note that the effects of welfare benefits on behavior are identified in large part from their cross-sectional variation across different U.S. states (see below).

<u>Estimation Method</u>. Let $P(j|\theta, X_i)$ denote the probability of participation-hours combination j conditional on a vector of observed characteristics for individual i, X_i, and a vector of all parameters in the model, θ, where $j=1,\ldots,J$. In our application, $J=3*2^M=24$. Also let $P(j|\theta, X_i, w_i)$ denote the probability of the same event but conditioned on the observed wage w_i. Letting d_{ij} be an indicator equal to 1 if person i chooses participation-hours combination j and 0 otherwise, the log likelihood function for our model (assuming the wage equation is estimated jointly with the choice model) is

$$L(\theta) = \sum_{i\in E}\sum_{j=1}^{J} d_{ij} \ln[P(j|\theta, X_i, w_i)\phi(w_i|\theta, X_i)] + \sum_{i\in U}\sum_{j=1}^{J} d_{ij} \ln P(j|\theta, X_i) \tag{8}$$

where E is the set of individuals who are employed and U is the set who are not employed, and ϕ is the normal p.d.f. The score of the log likelihood is

$$\nabla_{\theta} L(\theta) = \sum_{i\in E}\sum_{j=1}^{J} d_{ij} \nabla_{\theta} \ln[P(j|\theta, X_i, w_i)\phi(w_i|\theta, X_i)] + \sum_{i\in U}\sum_{j=1}^{J} d_{ij} \nabla_{\theta} \ln P(j|\theta, X_i) \tag{9}$$

To estimate the parameters of the model we use two different simulation methods, one a method of simulated maximum likelihood (SML) (Albright et al., 1977; Lerman and Manski, 1981) and the second a method of simulated moments

(MSM) (McFadden, 1989; Pakes and Pollard, 1989). In the first, we directly the log likelihood function by simulating the probabilities $P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)$ and $P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)$ and inserting the simulated values into (8). We denote the simulated probabilities by $f(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)$ and $f(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)$, respectively. This procedure does not give an unbiased estimator of the log likelihood because, for a finite number of Monte Carlo draws,

$$E \ln f(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i) \neq \ln P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)$$

and

$$E \ln f(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i) \neq \ln P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)$$

The simulated log likelihood is only asymptotically unbiased as the number of draws used to simulated the choice probabilities grows large. Thus, an estimator of $\underset{\sim}{\theta}$ obtained by maximizing the simulated log-likelihood function obtains consistency only as simulation size goes to infinity.[6]

Part of the attractiveness of conventional MSM estimators for discrete choice models, on the other hand, is that they obtain consistency for finite simulation size (McFadden, 1989). This property results from rewriting the score in terms of (weighted) moments that have zero expectation in the population. Unfortunately, applying the conventional MSM procedure is inappropriate in our model because of the presence of the wage equation. In an attempt to apply the procedure discussed by McFadden (1989), one could use the identities

$$\sum_{j=1}^{J} P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i) \; \nabla_{\underset{\sim}{\theta}} \ln P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i) = 0$$

and

$$\sum_{j=1}^{J} P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i) \; \nabla_{\underset{\sim}{\theta}} \ln P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i) = 0$$

to rewrite the score (9) as

$$\nabla_{\theta} L(\theta) = \sum_{i \in E} \sum_{j=1}^{J} [d_{ij} - P(j|\theta, X_i, w_i)] \nabla_{\theta} \ln P(j|\theta, X_i, w_i)$$

$$+ \sum_{i \in E} \sum_{j=1}^{J} d_{ij} \nabla_{\theta} \ln \phi(w_i|\theta, X_i) \quad (10)$$

$$\sum_{i \in U} \sum_{j=1}^{J} [d_{ij} - P(j|\theta, X_i)] \nabla_{\theta} \ln P(j|\theta, X_i)$$

However, this score expression, although superficially having an MSM form, is inappropriate. The difficulty arises because the moments $[d_{ij} - P(j|\theta, X_i, w_i)]$ and $[d_{ij} - P(j|\theta, X_i)]$ are not mean zero because the labor supply, participation, and wage errors are not mean zero within either the worker or nonworker subpopulations. We instead put the model in an MSM framework by adapting methods suggested by McFadden and Ruud (1994).[7] The expectation of the score (9) conditional only on X_i (and not on w_i) is

$$E[\nabla_{\theta} L(\theta)|X] = \sum_{i=1}^{N} E[\sum_{j=9}^{24} d_{ij} \nabla_{\theta} \ln[P(j|\theta, X_i, w_i)\phi(w_i|\theta, X_i)] \mid X_i]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{8} P(j|\theta, X_i) \nabla_{\theta} \ln P(j|\theta, X_i) \quad (11)$$

where the expectation is taken over all random deviates in the model and where we order the J=24 alternatives so that the first eight are those for non-working alternatives and the rest are working alternatives. Since $E[\nabla_{\theta} L(\theta)|X]=0$ at the true θ (since we have conditioned only on X_i and not on w_i), the quantity $\nabla_{\theta} L(\theta) - E[\nabla_{\theta} L(\theta)|X]$ also has expectation zero at the true θ. Using (9) and (11), we obtain

$$
\begin{aligned}
\nabla_{\underset{\sim}{\theta}} L(\underset{\sim}{\theta}) &- E[\nabla_{\underset{\sim}{\theta}} L(\underset{\sim}{\theta}) | X] \\
&= \sum_{i=1}^{N} \Big\{ \sum_{j=9}^{24} d_{ij} \nabla_{\underset{\sim}{\theta}} \ln[P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)\phi(w_i|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \\
&\qquad - E[\sum_{j=9}^{24} d_{ij} \nabla_{\underset{\sim}{\theta}} \ln[P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)\phi(w_i|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \quad |X] \Big\} \qquad (12) \\
&+ \sum_{i=1}^{N} \sum_{j=1}^{8} [d_{ij} - P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \; \nabla_{\underset{\sim}{\theta}} \ln P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)
\end{aligned}
$$

It is not feasible to construct an unbiased simulator of this score expression because simulation error enters nonlinearly through the gradient terms. However, it is possible to use (12) to develop a method of moments estimator because the first two lines of (12) represent the sum over individuals of mean-zero moments and the third line represents a sum of individual weighted moments, each with mean zero. Replacing the gradients in the third line of (12) with weights $W_j(\underset{\sim}{\theta}, \underset{\sim}{X}_i)$ that are asymptotically correlated with the gradients but uncorrelated with the corresponding moment, and replacing the gradients of the log choice probabilities in the first two lines by approximants $Z_j(\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i)$, we obtain the following first-order conditions (FOC) for a method of moments estimator:

$$
\begin{aligned}
FOC(\underset{\sim}{\theta}) &= \sum_{i=1}^{N} \Big\{ \sum_{j=9}^{24} d_{ij} \; [Z_j(\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i) + \nabla_{\underset{\sim}{\theta}} \ln\phi(w_i|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \\
&\qquad - E[\sum_{j=9}^{24} d_{ij} \; [Z_j(\underset{\sim}{\theta}, \underset{\sim}{X}_i, w_i) + \nabla_{\underset{\sim}{\theta}} \ln\phi(w_i|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \; | \; \underset{\sim}{X}_i] \Big\} \qquad (13) \\
&+ \sum_{i=1}^{N} \sum_{j=1}^{8} [d_{ij} - P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)] \; W_j(\underset{\sim}{\theta}, \underset{\sim}{X}_i)
\end{aligned}
$$

Following McFadden (1989), an MSM estimator of $\underset{\sim}{\theta}$ can be constructed by substituting an unbiased simulator $f(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)$ for $P(j|\underset{\sim}{\theta}, \underset{\sim}{X}_i)$ and also substituting an unbiased simulator for the expectation in the second line of (13). The latter can be constructed by integrating out the wage distribution with the simulation construct

$$(1/N_w) \sum_{\tilde{w}_i} \sum_{j=9}^{24} f(j|\underset{\sim}{\theta},\underset{\sim}{X}_i,\tilde{w}_i) \; [Z_j(\underset{\sim}{\theta},\underset{\sim}{X}_i,\tilde{w}_i) + \nabla_{\underset{\sim}{\theta}} \ln\phi(\tilde{w}_i|\underset{\sim}{\theta},\underset{\sim}{X}_i)] \qquad (14)$$

where N_w is the number of random draws $\tilde{w}_i$ from the wage distribution. For $W_j(\underset{\sim}{\theta},\underset{\sim}{X}_i)$ we use $\nabla_{\underset{\sim}{\theta}} \ln f(j|\underset{\sim}{\theta},\underset{\sim}{X}_i)$ and for $Z_j(\underset{\sim}{\theta},\underset{\sim}{X}_i,\tilde{w}_i)$ we use $\nabla_{\underset{\sim}{\theta}} \ln f(j|\underset{\sim}{\theta},\underset{\sim}{X}_i,\tilde{w}_i)$.[8]

The main computational problem in this simulation method is that the number of simulations required is N_w times the number of draws used to form $f(j|\underset{\sim}{\theta},\underset{\sim}{X}_i,\tilde{w}_i)$ (i.e., the wage itself must be integrated out with simulation methods). As a consequence, although this MSM estimator is consistent in sample size for a fixed simulation size, its computational advantage over the SML method (which requires a large simulation size for consistency) is greatly reduced. Indeed, we find in our application (see below) that MSM is computationally more burdensome than SML.

The burden of MSM created by the requirement that the wage equation must be estimated jointly with the choice model is eliminated if the wage equation is instead estimated separately in a first-stage estimation. Then predicted wages may be formed for the nonworkers alone or for the entire sample. In this case the choice model can be estimated alone, conditional on the predicted wage, and the method of McFadden (1989) for the estimation of multinomial probit models can be used directly. The score for this model is a modified version of (10), with the second and third lines of that score deleted, with the first line summed over all workers and nonworkers, and with the fitted wage is inserted for w_i either for nonworkers alone or for both workers and nonworkers:[9]

$$\nabla_{\underset{\sim}{\theta}} L(\underset{\sim}{\theta}) = \sum_{i=1}^{N} \sum_{j=1}^{J} [d_{ij} - P(j|\underset{\sim}{\theta},\underset{\sim}{X}_i,\hat{w}_i)] \; \nabla_{\underset{\sim}{\theta}} \ln P(j|\underset{\sim}{\theta},\underset{\sim}{X}_i,\hat{w}_i) \qquad (15)$$

Implementing either simulation estimator is eased considerably by using simulators that are smooth functions of $\underset{\sim}{\theta}$, for this permits the use of standard gradient optimization procedures (see McFadden (1989)). The importance sampling smoothing methods proposed by McFadden are not particularly desirable for our problem because those methods require that the dimension of the vector of utility differences equal the number of error terms in the model. In our model a factor structure of errors of smaller dimension underlies the vector of utility differences. Therefore we instead adopt an alternative smoothing procedure discussed by McFadden which adds an i.i.d. extreme value error term to the utility of each alternative. The resulting choice probabilities are multinomial logit (conditional on the normal error terms contained in the U_k (see equation (3)) and are of the form

$$P(k|\underset{\sim}{\theta}) = \exp[U_k(\underset{\sim}{\theta})/\tau] \; / \; \{ \sum_{j=1}^{J} \exp[U_j(\underset{\sim}{\theta})/\tau] \} \tag{16}$$

where J=24 in our application and where τ is the standard deviation of the extreme value errors. This probability simulator is unbiased as τ approaches zero. In practice, we set τ at a value close to zero.

A critical feature of this smoothing method (which is not a feature of other methods) is that it allows us to simulate the choice probabilities without knowing the boundaries of the regions of integration that generate those probabilities. Equation (16) can be evaluated simply by calculating the $U_k(\underset{\sim}{\theta})$ at simulated values of the stochastic terms. This property is important because, as we noted previously,the relevant integration boundaries are intractable in our and similar applications.

III. Data and Variable Construction

We utilize data from the fourth wave of the first panel of the Survey of Income and Program Participation (SIPP), which was administered in the Fall of 1984. The SIPP is a nationally representative sample of the U.S. population of approximately twenty thousand households, and was especially designed to elicit accurate information on income and participation in various transfer programs. We select all female heads of family 18-64 with children under the age of 18 present. We also exclude families with high asset levels (over $4500) because they are far above the transfer-program asset limits and their behavior is likely to be structurally different from those with assets below that level. Our sample consists of 968 women.

The dependent variables are defined as of the month prior to interview. Participation in AFDC, Food Stamps, and subsidized housing are defined with regard to whether any participation took place in the month, and labor supply is defined as the average weekly hours of work in the month, with 1-to-35 defined as part-time and hours in excess of 35 defined as full-time (we will conduct a sensitivity test to these cutoff points). Table 1 shows the distribution of the sample across the categories. Over 50 percent of the female heads participate in no transfer program but 8 percent participate in all three and 21 percent participate in two.[10] Of those who participate in at least one program, two-thirds participate in more than one. The table also shows that the joint distribution of labor supply and participation is highly asymmetric. The largest cells in the table are for full-time workers not on any program, and for nonworkers participating in AFDC and Food Stamps only or in all three programs.[11] The data are not far from a bimodal distribution: almost no women on welfare work and almost all women off welfare work.

Hourly wage rates for those who work are computed from earnings and hours of work in the prior month, and are used to compute weekly earnings at part-time and full-time work. Nontransfer nonlabor income is computed as the sum of asset income and the income of others in the family. Variables are also constructed for a set of socioeconomic characteristics, including education, age, numbers of children, regional location, race, SMSA residence, and a number of state characteristics. The means of the variables are shown in Appendix Table A-1.

Benefits for each of the three programs and for positive taxes (federal and FICA) are computed at each of the three hours points. The benefit formulas are quite complex. They often interact--the Food Stamp and housing programs tax AFDC benefits, for example, and the AFDC programs in some states tax housing subsidies. Hence we calculate benefits separately for each program combination. The three programs also permit deductions for child-care expenses and for other work-related expenses. We include such deductions in our benefit calculations but, in order to maintain consistency with the budget constraint for women who work when off the programs, we assume that such expenses are incurred for all workers as well. The benefit formulas generally also have low maximum income limits that create notches in the budget constraint where benefits fall to zero discontinuously.

The details of the benefit, tax, and general work expense calculations are given in Appendix B, along with illustrative cumulative marginal tax rates that arise from participation in the three programs. As the Appendix shows, cumulative marginal rates are extremely high, usually exceeding 100 percent between H=0 and H=20 and often around .80 to .90 between H=20 and H=40. The net wage between H=0 and H=40 is generally less than $1 per hour at the mean of the population, and much lower for low-wage women. Given these results, it is not surprising that so few women on welfare work and that the labor-supply-participation distribution is close to bimodality.[12]

Finally, we introduce three extra parameters into the model to account for the in-kind nature of housing benefits, and to account for Medicaid and private health insurance. In general, in-kind transfers are not equivalent to cash from the viewpoint of the recipient. However, while the literature has shown Food Stamps to be essentially equivalent to cash, housing benefits appear to be valued at substantially less.[13] In addition to this issue, housing benefits differ in another important respect from AFDC and Food Stamps because such benefits are rationed. Public housing is available only to those who queue for several years and Section 8 subsidies are restricted in quantity. Housing benefits are thus not an entitlement. To capture both of these influences, we introduce a parameter γ_R which represents the extent to which housing benefits are discounted relative to cash and, as well, the extent to which participation rates respond to changes in housing benefits in the first place (in light of possible rationing). We therefore replace the budget constraint in (2) with

$$Y(H,Y,P_A,P_F,P_R) = wH + N + B_A(H)P_A + B_F(H)P_F + \gamma_R B_R(H)P_R - T(H) - E(H) \quad (17)$$

(general work expenses, E(H), are also included, as noted previously in this section). As well as capturing the extent of rationing, the parameter γ_R is observationally equivalent to a marginal utility parameter because (17) is inserted into the utility function for estimation.[14]

We also account for the relative availability of health insurance benefits by utilizing the Medicaid and private health insurance variables constructed for female heads in the SIPP by Moffitt and Wolfe (1992). Female heads on AFDC are automatically enrolled in the Medicaid program, which provides subsidized medical care, whereas female heads off AFDC are rarely eligible for Medicaid benefits at all.[15] Moffitt and Wolfe constructed a family-specific index of the expected value of Medicaid benefits for each female-headed family in the SIPP based on expected medical expenditures according to the health and other characteristics of the female head and her children, which we shall add to the budget constraint. They also calculated an expected value of private health insurance for women off AFDC, equal to the product of the predicted probability of private health insurance coverage and expected expenditures if covered, the latter once again a function of the family's health and other characteristics. We include the private health insurance variable in the budget constraint as well. However, while we introduce both Medicaid and private health insurance benefits, we allow their effects to differ from those of cash both because they are in-kind quantities, and because they are not measured in the same units as other benefits. With their addition, the final budget constraint becomes:

$$Y(H,P_A,P_F,P_R) = wH + N + B_A(H)P_A + B_F(H)P_F + \gamma_R B_R(H)P_R + \gamma_{MED}B_{MED}P_A + \gamma_{PHI}B_{PHI}(1-P_A) - T(H) - E(H) \quad (18)$$

where B_{MED} and B_{PHI} are the Moffitt-Wolfe values of Medicaid and private health insurance benefits, respectively, and γ_{MED} and γ_{PHI} are parameters to be estimated.[16]

IV. Results

Our initial estimates of the full model are obtained by estimating the labor supply and program participation equations jointly with the wage equation, are shown in Table 2. Our estimates use the SML method with 500 draws per individual, a value reached by successively increasing the number of draws until the parameters and standard errors stabilized. Computation proved extremely burdensome in this method, requiring 30 minutes of CPU time per iteration on an IBM 3090 mainframe computer. Estimates obtained when using 1000 draws as well as estimates using MSM are shown subsequently.[17]

The first portion of the table shows the coefficients on the elements of the $\underset{\sim}{X}$ vector for tastes for work, stigma costs, and the wage equation (see eqns.(5)-(7)). The specification of the vector shown in the table represents the final specification after considerable testing. The parameter estimates are mainly as expected from other work. Children significantly reduce labor supply; they also reduce welfare costs for AFDC and Food Stamps (i.e., they increase welfare participation), though the effect is not significant. Women who are older, who have higher levels of education, who are in good health, and who are white have higher levels of labor supply and usually have lower welfare participation propensities, though once again not always significantly.[18] The state unemployment rate has a very weak negative effect on labor supply, workers in SMSAs and in states with high fractions of employment in services (where low-income women are heavily concentrated) have higher wages, and individuals in states with high AFDC administrative expenses have lower AFDC participation rates but higher Food Stamp participation rates, possibly because these expenses are spent implementing more stringent administrative AFDC requirements.[19]

The second panel of the table shows the estimates of the utility function parameters and the covariance matrix of the errors. Initial estimates of the quadratic utility function (4) revealed a lack of significant interactions between H and Y, or between H or Y and program participation; the model shown in Table 2 therefore omits these terms. The remaining parameters are all significant at conventional levels. The utility function parameters β_{HH} and β_{YY} have no direct interpretation but they together determine wage and income elasticities.[20] At the means the uncompensated wage elasticity ϵ_w is 1.82 and the total income elasticity (ϵ_y) is -.21. The wage elasticity is in the high end of prior estimates for women, but most prior estimates have been obtained for married women rather than female heads. The income elasticity is on the low side of past estimates for married women, on the other hand. We test below whether measurement error in nonlabor income may contribute to this low estimate.[21]

Initial estimation revealed significant interactions between the three participation indicators (parameters ϕ_{mn} in eqn (4)). However, the estimates were difficult to interpret in their intended fashion, namely, as representing whether participation costs are strictly additive or strictly non-additive. After some experimentation, these interaction terms and the uninteracted participation terms were replaced in the utility function by the expression

$$\lambda\ (\psi_A P_A + \psi_F P_F + \psi_R P_R) + (1-\lambda)\ \mathrm{Max}(\psi_A P_A, \psi_F P_F, \psi_R P_R) \qquad (19)$$

where $0<\lambda<1$ is a parameter to be estimated. This specification allows participation costs to fall somewhere between perfect additivity ($\lambda=1$) and perfect non-additivity ($\lambda=0$), the latter corresponding to a situation where the stigma and other costs of participating in one program are not increased at all by participating in multiple programs. As seen in Table 2, the estimate of λ, although significantly different from zero, implies that the stigma and other costs of program are almost entirely non-additive.

The γ parameters show the importance of Medicaid and private health insurance, and of subsidized housing benefits, on choices. The results show significant effects of Medicaid benefits. Since those benefits are generally only available if the individual is on AFDC, this implies that Medicaid has a strong effect of drawing women onto the AFDC program and reducing their labor supply. Private health insurance benefits have a positive effect (and hence draw women off AFDC) but it is insignificant.

As for housing, the estimate of γ_R is small and has a high standard error, implying that subsidized housing benefits have no significant impact on participation in housing programs. Although this may reflect a low cash-equivalent value of this in-kind benefit, it more likely reflects extensive rationing in the allocation of subsidized housing units. As we noted previously, the stock of units is limited and there is excess demand for slots; the criteria by which slots are awarded appear to have little relation to the potential benefit the woman puts on such housing. Put differently, the public housing program is the one program we study which is not an entitlement program; participation is not guaranteed merely upon application and determination of eligibility. We therefore venture that modeling housing program participation as subject to administrative constraints rather than as an unrestricted utility-maximizing choice will be necessary to explain the determinants of participation. Given the lack of explanatory power of our model for housing benefits and the housing participation decision, we shall delete the housing portion of our model from the rest of our estimates in the paper.

Table 3 shows the estimates of our model without the housing equation. As expected, the deletion of the housing equation has little or no effect on any of the other parameters in the model, including the wage and income elasticities that are particularly important. The lower portion of the table showing the covariance matrix estimates indicates significant positive correlations between AFDC and Food Stamp errors, and between wage errors and Food Stamp "costs" (which are negatively related to participation rates) In addition, we should note that the Pearson χ^2 goodness-of-fit statistic (the ratio of the sum of squared deviations between actual and predicted outcome cell frequencies to the sum of squared predicted frequencies) falls from that in Table 2 because only 12 cells are fit instead of 24. However, the smaller model is much closer to the 5-percent χ^2 significance level than the larger model, indirectly reflecting the poor fit of the housing choices in the latter.[22]

Table 4 shows a comparison of the actual and fitted distributions of the data across the outcome categories of the model without housing, which gives a better sense of fit than the Pearson χ^2 alone. The fit of the model is reasonably good given the parsimony of the specification. There is a slight tendency for the model to overpredict the frequency of full-time workers off welfare and to underpredict the frequency of nonworkers on welfare, which together affect the marginals as well. Nevertheless, the fit is surprisingly good for many of the less frequent program combination categories.[23]

We subjected the model to a large number of alternative specifications and sensitivity tests, selected parameters from which are shown for a few of the specifications in Table 5. Sensitivity to our specification of the $\underset{\sim}{X}$ vector is shown in the first column, where estimates with only a small subset of those variables are obtained (only the children variables in the labor supply equation and only the administrative-expenditure variables in the cost equations). The most important parameters--β_{HH} and β_{YY}--are unaffected by this change, as are the wage and income elasticities. The health insurance variables increase in magnitude, however, no doubt because they are predicted values and are based in part on the excluded elements of the $\underset{\sim}{X}$ vector. While a likelihood ratio test rejects this specification in favor of that in Table 3, the results nevertheless support a conclusion that our key results do not appear to be especially sensitive to the other variables in the model.

The sensitivity of the results to the number of draws in the SML model (an important issue since the estimator is consistent only in the number of draws) is shown in the next column, which reports estimates computed with 1000 draws. The computational burden approximately doubles (hence one hour of CPU time per iteration) with this number of draws since computation time in the SML method is roughly linear in the number of draws, but the point estimates of the parameters are affected relatively little, save for a slight reduction in the income elasticity. The major impact of the larger number of draws is instead a slight increase in estimated standard errors.

Estimating the model with MSM instead of SML yields approximately the same parameter estimates, though with a somewhat larger value of β_{YY}, which translates into a somewhat smaller wage elasticity. In addition, the standard errors sometimes deteriorate substantially (e.g., for the health variables) although without much change in the point estimates of the parameters. However, MSM proved difficult to implement because of the large number of draws required. The estimates in the table are based upon 50 wage draws and 50 draws of the choice-equation errors, for a total of 2500 simulations per individual (for the evaluation of eqn (14)). This increases the computation time to two CPU hours per iteration, four times the time required for the SML method with 500 draws. However, 50 choice draws per individual is still only one-tenth the number used in the SML estimates, which necessarily reduces the accuracy of the estimation; MSM estimation with 500 choice draws would be prohibitively computationally intensive. In our application, therefore, we find a substantial computational disadvantage to MSM method as against SML, despite the requirement for a larger number of draws for consistency in the latter method.

Since the computational burden of MSM resides entirely in the requirement that the wage be integrated out, we provide in the next columns of the table estimates which use predicted wages obtained from first-stage OLS estimates of the wage equation.[24] Without the inclusion of the wage equation, estimation requires only that the 12 labor-supply-participation choice categories be fit by the model, which proved to be by far the least burdensome computational method of those we tried, including the SML method. Convergence was always rapidly achieved (5 CPU minutes per iteration) and iterations were well-behaved in their pattern. The estimates shown in the table use predicted wages for both workers and nonworkers, the former designed to correct for measurement error in the wage as well other types of endogeneities not represented elsewhere in the model. Estimates which used predicted wages only for the nonworkers were also obtained, however, and showed little change in the key parameter estimates.

The major conclusion from the wage IV method is that the parameter estimates are somewhat altered but not enough to change any of the substantive conclusions of the model. The β_{HH} parameter differs only slightly from its Table 3 value while the β_{YY} estimate falls somewhat more, with a corresponding fall in the income elasticity. The health insurance parameters also increase significantly in magnitude. Nevertheless, the Pearson χ^2 statistic for the choice component of the model drops dramatically from that in Table 3 and now fails to reject the specification at conventional levels (the 5-percent cutoff value is 19.7, as against the 15.6 value of the statistic in Table 5). A more detailed fit table corresponding to Table 4 for this model shows a much closer fit as well and a disappearance of the over- and underpredictions noted previously for the full model. This suggests that joint estimation of the wage equation, at least with the conventional types of restrictions we have imposed for it and its relation to the other equations, substantially reduces the goodness of the fit of the choice equations themselves. This result, together with the computational advantages of this model, make it an attractive alternative to the full model.

The last two columns test the importance of measurement error in nonlabor income, and of our definitions of the hours-worked categories. Using predicted instead of actual values for nonlabor income in the model increases all parameters, especially the β_{YY} parameter and the corresponding income elasticity, supporting an interpretation of measurement-error problems in nonlabor income as a cause of the rather low income elasticities obtained.[25]econd, we conduct an out-of-sample comparison by comparing the actual change in labor supply and program participation from a historical event (a major increase in the AFDC tax rate in 1981)--as measured in a different sample--to the change predicted by our model. The latter exercise serves both as a further policy exercise as well as an out-of-sample check of the validity of our model.

We also examine the sensitivity of our simulations to alternative specification by conducting them for two of our estimated models: (1) the full model estimated by SML, joint with the wage equation (as reported in Table 3), and (2) the choice model alone, estimated by MSM with an instrumental wage (as reported in Table 5). Although we have found that the latter estimation method is not only computationally much simpler but also appears to give approximately the same parameter estimates as those obtained with the former method, it is also desirable to determine whether both methods show approximately the same types of responses to changes in the budget constraint and policy parameters.

Table 6 shows simulations of the model obtained by computing mean probabilities for each of the twelve outcome categories over all individuals in the sample for different alterations in the budget constraint, using the SML estimates in Table 3.[26] The first row shows our baseline simulation while the second and third rows show the effects of reducing the marginal tax rates facing welfare recipients, a key issue. Interestingly, a reduction of the AFDC tax rate from its current level of 100% to 50% has scarcely any effect on labor supply, and it actually increases the participation rate in both AFDC and Food Stamps. This result is a reflection of a phenomenon noted in the welfare literature previously (see, e.g., Levy (1979) and Moffitt (1992)) sometimes referred to as the "breakeven problem." The problem arises because a reduction in a welfare-program tax rate raises the breakeven level (i.e., the point at which eligibility terminates), which draws some individuals onto the rolls with consequent reductions in labor supply. That this is occurring the present example is clear from Table 6, for the fraction of recipients working full-time actually falls when the tax rate is reduced; this arises because it is generally full-time workers who are made newly eligible, or nearly eligible, by the increase in the breakeven point, and it is they who reduce their labor supply when going onto the welfare rolls. The table also shows the effect on costs, defined as the increase in benefits of new entrants minus benefit reductions from existing recipients who work more or who leave the rolls, plus the change in the tax payments of both groups, as a percent of initial benefit payments minus tax payments. We find that there is little effect on costs from the tax rate reduction to 50%, for the benefit savings from increased numbers of working recipients is cancelled out by increased benefits for new entrants.

A more massive change of reducing both AFDC and Food Stamp tax rates to .10 succeeds in increasing labor supply by about 2 hours per week, but at the cost of increasing the AFDC caseload by about one-third and the Food Stamp caseload by about one-fourth. The tax-rate reduction in this case induces large numbers of nonworking recipients to go to work, either part-time or full-time, outweighing the labor supply reductions of the new recipients. However, costs rise by almost 80 percent for this reform because of the larger number of recipients.

The effects of these changes in marginal tax rates are in sharp contrast to the effects of increasing wage rates, for the two are not symmetrical for welfare programs. The remaining rows of Table 6 demonstrate this result. An increase in the gross hourly wage rate of $1 significantly reduces participation in both AFDC and Food Stamps and also increases expected weekly hours of work by about 3.5 hours. A wage change pivots the budget constraint around the origin and increases income if off welfare both above and below breakeven (the below-breakeven income increase pulls women off AFDC as well as the above-breakeven increase), in contrast to the tax-rate reduction. The magnitude of the hours increase is somewhat less than that implied by the wage elasticity (which was 1.94) primarily because about one-third of the sample does not work. The wage-increase increases costs by more than 160% (costs in this case are defined to include the increased wage costs).

The imposition of a minimum wage, which pulls up only the bottom portion of the wage distribution, has even greater downward effects on program participation and upward effects on labor supply. However, it costs less than the dollar-increase in the wage because the wages of high-skilled workers are not increased. A wage-rate subsidy--a policy often proposed to draw women off welfare--with a 50% subsidy rate up to a wage of $6 per hour is simulated to have about the same effects as the wage-increase, but once again at lower cost because the wage increases are concentrated at the lower end of the distribution. An increase in the EITC (the earned income tax credit) operates in a similar way, by increasing after-tax wages if off AFDC.[27] The effects on program participation and labor supply are in the same direction as those already discussed, but smaller in magnitude. The smaller magnitude is, in part, a result of the phaseout region of an EITC, which tends to reduce labor supply.

The final row of the table shows the effect of offering a small work subsidy to women off the welfare rolls as well as to women on the rolls (for the latter, to replace existing welfare deductions for work-related expenses). This policy both increases labor supply and reduces program participation, at reduced cost. The policy succeeds in drawing women off the rolls because the small work subsidy substantially reduces the work-discouraging effects of fixed costs of work (which we have in our model), and because many women with high levels of stigma are willing to leave the welfare rolls, even at reduced income, if such work subsidies are available. Costs are reduced because of the substantial numbers of women who exit the rolls.

Appendix Table A-2 shows the same simulations using the MSM estimates with IV wage. Because this method gives a much better fit to the choice distribution, the baseline means of the program participation and labor supply variables differ slightly from those in Table 6. In addition, the magnitudes of the changes in participation and labor supply are somewhat different than those obtained from the SML estimates. However, the directions and overall orders of magnitude of the changes in participation and labor supply are the same in every case as those in Table 6, and hence the same substantive conclusions would be drawn from both sets of simulations. We therefore find that, at least for the types of budget-constraint changes we have considered, the two estimation methods yield similar substantive economic implications.

Finally, we conduct an out-of-sample check on our model by simulating the effects of a major change in the AFDC tax rate in 1981. In the summer of that year, the U.S. Congress increased the AFDC tax rate from 67 percent to its current level of 100 percent. We obtain external data on the changes in labor supply and program participation among U.S. female heads from the year prior to the legislation, 1980, to 1984, the year of our data and analysis. We then obtain program rules for AFDC, Food Stamps, federal tax rates, and all other program rules delineated in Appendix B but for 1980 instead of 1984. We then use our 1984 SIPP data but apply the 1980 rules to simulate changes in program participation and labor supply from 1980 to 1984 (other minor parameters of the rules other than the AFDC tax rate changed between the years, but none was as important as that change). The external data is drawn from a combination of administrative sources and the March files of the Current Population Survey (CPS), as indicated in Table 7.[28]

Table 7 shows the results of our simulations and comparisons. The table shows a fairly remarkable similarity in the direction and general magnitude of the changes, despite the many differences in data sources. In both our simulations and the external data, there was a major reduction in the AFDC participation rate, of 10 percentage points simulated by our model and of 8.5 percentage points in the external data. Mean hours of work among female heads essentially did not change in either our simulations or the external data, consistent with the Table 6 simulations of a change in the AFDC tax rate of from 100 percent to 50 percent. As the changes in the distributions of hours of work in Table 7 indicate, the lack of change in mean hours was a result of a movement of women from part-time to full-time; that is, from women leaving the welfare rolls after the tax-rate increase to work while off welfare. The last row shows major increases in the fraction of AFDC recipients who do not work--which is a result, once again, of most working recipients having left the rolls after the change. We note as well that both our SML and MSM estimates provide very similar simulated changes.

V. Summary

In this paper we have applied simulation methods to address one of the long-standing issues in the analysis of the effects of transfer programs on labor supply, namely, the problem of multiple participation. We set up a relatively simple multinomial choice model for the choice of program combination and of labor supply conditional on that combination; our initial model has 24 choice cells. We find that simulated maximum likelihood (SML) estimation, conducted by jointly estimating a wage equation with the choice model, is feasible and well within computational capability. However, we find that the simulated method of moments (MSM) is extremely computationally burdensome if the wage equation is estimated jointly, despite the fact that it is consistent even with fixed numbers of draws, but that MSM estimation of the choice model alone but with predicted wages is the fastest of the alternative estimation methods and yet yields parameter estimates reasonably close to those of SML. These results should have implications for future work on applications with similar problems (e.g., with joint wage or other price equations).

Substantively, we find that participation in one of the major transfer programs for the poor in the U.S.--subsidized housing--is unrelated to housing benefits. We speculate that this results from rationing of housing units in the program. We also find that, while cumulative tax rates for recipients in multiple programs are very high, small-to-moderate reductions in those tax rates have very little effect on labor supply because of offsetting decreases in labor supply that arise from increased program entry; the tax-rate reductions also increase the welfare caseload. This finding is reinforced by an out-of-sample comparison to a 1981 tax-rate increase in the AFDC program. However, we find that many types of wage subsidies and related types of wage rate increases have both significant positive labor supply effects and decreased program participation effects.

Appendix B

Benefit, Tax, and Work Expense Algorithms

AFDC. The formula for the monthly AFDC benefit in 1984 was (all income amounts were converted to weekly for the model estimation):

$$B_A = \text{Min}\{ P, r[G_1 - \text{Max}(0, WH + N - C - E)] \} \quad \text{if } WH + N < (1.85)G_2$$
$$= 0 \text{ if not} \qquad (B1)$$

where P is the maximum payment permitted in a state, r is the "ratable reduction" (a number between 0 and 1 by which the benefit may be reduced), G_1 is the maximum amount paid, C is the child-care expense deduction (for workers only), E is other deductable work-related expense, and G_2 is the needs standard. The variables P, G_1, and G_2 vary by state and family size and are available from unpublished data from the Office of Family Assistance, Department of Health and Human Services. The ratable reduction, r, is available in U.S. Department of Health and Human Services (1985, p.335). Permissible work-related deductions in 1984 were $90 ($30 set-aside plus $60 maximum remaining expenses).

Work-related deductions (E) were set at $90 per week, the sum of a standard $30 deduction for all workers and a mean of $60 of extra deductions for AFDC recipients who work (U.S. House of Representatives, 1987, Table 25, p.435). Child-care deductions (C) were estimated for AFDC and for the other programs below as follows. Nationally, in 1984 average child care deductions for AFDC women were $93 per month for those who had positive deductions (U.S. House of Representatives, 1987, Table 25, p.435). Assuming these were generated by children under six and that there were two such children on average in these families, the deduction was approximately $46 per child per month. AFDC agencies generally assume child care expenses for part-time workers that are roughly half of those for full-time workers, so we assume the same in order to apportion the $46 average across part-time and full-time states. In our data there are 14 part-time working AFDC recipients for every 10 full-time AFDC working recipients, implying that mean deductions for the former were $33 per child per month and the latter, $66 per child per month for children under 6. The maximum allowable amount for part-time work, though not for full-time work, also varies by state. To capture cross-state variation, the $33 amount for part-time work is multiplied by the ratio of the state maximum for child-care expense for part-time work to the national average across states of all such maxima (state maxima for part-time work are taken from the individual state tables in U.S. Department of Health and Human Services (1985)). Finally, since only 20 percent of working AFDC recipients take a deduction (U.S. House of Representatives, 1987, Tables 23 and 25), the presence of the deduction was simulated along with the general simulation estimator with the mean probability set equal to .20.

The AFDC benefit was reduced in some cases for families in subsidized housing, as discussed below.

Food Stamps. We use the formula given in Fraker and Moffitt (1988, p.27). The formula for the monthly Food Stamp benefit in 1984 was:

$$\begin{aligned} B_F &= \text{Max } [M, G - .30 Y_{n1}] \quad \text{if} \quad WH + N < M_1 \text{ and } Y_{n1} < M_2 \\ &= 0 \quad \text{if not} \end{aligned} \tag{B2}$$

where

$$Y_{n1} = \text{Max } (0, .82WH + N + B_A - 95 - S) \tag{B3}$$

$$S = \text{Min } [134, \text{Max}(0, R - 0.5Y_{n2})] \tag{B4}$$

$$Y_{n2} = \text{Max } (0, .82WH + N + B_A - 95) \tag{B5}$$

where G is the Food Stamp guarantee, M is a minimum benefit, Y_{n1} is a first type of net income, M_1 is the gross income screen, M_2 is the net income screen, S is a shelter deduction, R is rent paid, and Y_{n2} is a second type of net income. G, M_1, and M_2 vary with family size and are obtained from unpublished data provided by the Food and Nutrition Service of the U.S. Department of Agriculture. No parameters vary by state since Food Stamps is a national program.

Subsidized Housing. Subsidized housing in the U.S. takes the form of either public housing or subsidized private rental housing (the Section 8 program). In both programs families with sufficiently low income and assets are eligible, and in both programs the tenant is obligated to pay rent according to a formula set by the government. In Section 8 housing the tenant pays the landlord the government-stipulated rent and the government pays the landlord the increment necessary to bring the total up to an amount known as the "fair market rent" for the unit (if the landlord charges a rent greater than this, the tenant must pay the landlord directly for the excess). In public housing, the government simply collects the rent and then provides the housing itself.

For present purposes the housing subsidy is taken as the difference between the tenant rental payment and the fair market rent. The latter is taken to be the same value for both public housing and private rental housing, since no information is available on the fair market value of public housing. Fair market rents by county and by bedroom size for 1984 were obtained from the July 5, 1984 issue of the Federal Register. The data are linked to families by assuming that required bedroom size is one less than family size (up to 3 rooms).

For participants not on AFDC or on AFDC in all but 10 states, the monthly rental payment (S) in 1984 was determined by the formula:

$$S = \text{Max}\ (.10Y_g,\ .30Y_n) \qquad \text{(B6)}$$

where

$$Y_g = WH + N + B_A \quad (B7)$$

$$Y_n = Y_g - 40K - C \quad (B8)$$

where Y_g is gross income, Y_n is net income, K is the number of children, and C is child care expense (calculated as previously described).

The rental formula for families on AFDC in the remaining 10 states was:

$$S = Max\ (.10Y_g,\ .30Y_n,\ rM) \quad (B9)$$

where r is the ratable reduction in the state AFDC program and M is the maximum shelter deduction permitted in the state AFDC rules. In these 10 states the AFDC departments explicitly set a maximum shelter allowance per family; HUD therefore assumes in these states that AFDC recipients will automatically receive r percent of this maximum (see equation (B1)). Values for M are taken from from U.S. Department of Health and Human Services (1985, pp.337-338).

In these 10 states there is the possibility that the AFDC benefit is reduced as well. If $S<M$ in these states, where the AFDC benefit used in the calculation of Y_g is that in equation (B1), the AFDC benefit is reduced by $r(M-S)$. This secondary benefit reduction arises because the AFDC rules in these states do not permit the payment of the maximum shelter allowance, M, if the actual shelter payment of subsidized housing participants is less than this amount (even though the housing agency assumes in its calculation that the maximum shelter allowance is provided).

In all states, families are ineligible for any type of subsidized housing if $Y_g>L$, where L is a "low income" limit set by the U.S. Department of Housing and Urban Development. L varies by area; 1984 values were obtained from unpublished data provided by HUD.

Positive Taxes. The female heads in the sample are assumed to have filed as heads of household in calendar 1984, to have taken the standard deduction, and to have taken one exemption per person in the family. AFDC benefits, Food Stamps, and housing subsidies are not included in income for tax purposes. Marginal tax rates and bracket endpoints are available from standard IRS sources. The earned income tax credit in 1984 was also assigned. The 1984 FICA tax rate was .067 up to $37800 of annual earnings.

General Work-Related Expenses. All workers are assumed to incur $90 per month of general work-related expenses (equal to the AFDC amount) and child-care expenses of C as described under the AFDC benefit formula section.

Illustrative Cumulative Tax Rates. Table B-1 shows illustrative benefits and cumulative tax rates for six states, two with high benefit levels (California, Minnesota), two with medium benefit levels (Ohio, Kansas), and two with low benefit levels (Alabama, Texas). Cumulative marginal tax rates are very high in all states, near to or exceeding 100 percent over the part-time range and very high over the part-time-to-full-time range as well. The implicit net wage between H=0 and H=40 is negative in two states and is never more than $.75 per hour in the other states. The high tax rates arise partly from the high rate in the AFDC program, which itself is as much a result of the limits on maximum income as well as the nominal 100 percent tax rate (see above description of benefit formula); that is, the notches created by the maximum income limits are often below 20 hours of work per week and always below 40 hours. Similar notches are present in the Food Stamp program. Housing benefits are phased out at much lower rates and, in fact, fairly high-income families are eligible for such benefits; however, as noted in the text, the housing program rations its stock and hence limits eligibility indirectly.

NOTES

1. Steinberg (1989) proposed that simulation methods be used for an analysis of these programs as well, but for a model quite different from ours.

2. Although the existence of such non-participating eligibles could conceivably be the result of ignorance or simple measurement error in the data, several studies have consistently shown that the probability of participation is strongly correlated with the potential benefit, the wage rate, and other economic measures of the relative benefit of participating (see Moffitt, 1992, for a review of these studies). These results provide strong evidence of purposeful behavior.

3. As noted earlier, Fraker and Moffitt (1988) therefore estimated reduced form equations even though they only considered two programs and hence needed only to evaluate trivariate integrals, a task well within the boundaries of current computational feasibility with quadrature methods.

4. For example, if each of the $J=3*2^M$ discrete choices were allowed its own separate error term, the problem would reduce to the textbook multinomial choice problem. But it is unlikely that a J-by-J covariance matrix (e.g., a 24-by-24 matrix in our application) is necessary when there are only M+1 dependent variables.

5. The variance of α is identified because we have three H categories rather than two. The variances of the three disutility parameters, ψ_m, are identified even though they implicitly enter into binary choice equations (i.e., differences in (4) across participation choices) because the normalization of the coefficient of Y in (4) to one implies that $1/\sigma_j$ is the coefficient on ΔY in the implicit program-participation equations. Thus these variances simply fix the other parameters of the model relative to the effect of Y, i.e., they scale the other parameters in dollar terms.

6. We obtain SML standard errors from the inverse of the outer product of the simulated scores.

7. We are grateful to Paul Ruud for pointing out to us how this adaptation could be achieved.

8. The draws used to construct $f(j|\underline{\theta},\underline{X}_i)$ and $f(j|\underline{\theta},\underline{X}_i,\tilde{w}_i)$, on the one hand, and $W_j(\underline{\theta},\underline{X}_i)$ and $Z_j(\underline{\theta},\underline{X}_i,\tilde{w}_i)$, on the other, are independent and held fixed throughout the optimization process. McFadden (1989) and Pakes and Pollard (1989) prove that such MSM estimators are consistent and asymptotically normal. McFadden (1989) gives formulas that may be used to construct a consistent estimate of the covariance matrix of $\hat{\underline{\theta}}$.

9. We condition on the distribution of w and continue to estimate the covariance of the wage error with the other errors in the model.

10. We should that, although AFDC participants are categorically eligible for Food Stamps, not all bother to collect, either because the benefit is too small or because they view Food Stamps as particularly stigmatizing.

11. In addition, a sizable fraction of nonworkers--20 percent--participate in no program, reflecting the presence of non-participating eligibles.

12. Indeed, marginal tax rates often exceed 100 percent between H=0 and H=40 in many states, which would be inconsistent with observing any workers in the data if welfare costs (stigma, etc.) were not in the model.

13. See Smeeding (1982). Food Stamp benefits are sufficiently low that most families would purchase that amount of food in the absence of the program.

14. A similar parameter for the AFDC benefit was also tested in the early analysis and found to be close to 1.0.

15. They can receive benefits only if they are in a state offering a "Medically Needy" program, have income less than 130 percent of the poverty line or have spent their income down to that level, and are not covered by private health insurance.

16. As equation (18) indicates, we ignore the Medically Needy program mentioned in the prior footnote. We tested a variable for women in this category, and its coefficient was close to zero and insignificant.

17. The CPU time is obviously machine-specific and would be lower if more powerful machines were used. However, it is the relative CPU times for the alternative estimation methods that are important for our investigation, as we noted below. We should also that the smoothing parameter τ is set to 2.0, which is a small number relative to values of U_k, which are typically around 300. Smaller values of τ were also tested, with no change in results.

18. Since the utility function is normalized by weekly income, Y (see equation (4)), all parameters in the model can be related to income units. For example, at H=20, an increase in the number of children less than 18 (with coefficient -0.26 in column (1)) is roughly equivalent to a reduction in weekly income of $5.20 (=.26*20) in utility terms (ignoring the quadratic income term). A one-year increase in age (coefficient=0.18 in the Food Stamp equation) lowers the utility of participating in the Food Stamp program by $1.80 per week (the coefficients are scaled by 10). The other coefficients can be interpreted analogously.

19. We include an SMSA variable and a variable for the percent of state employment in the service sector as identifiers to proxy local labor market conditions. We tested four state employment fractions (service, manufacturing, wholesale-retail trade, and other) but found only the service-sector variable to enter significantly.

20. The uncompensated wage elasticity and the total income elasticity for our utility function are, respectively, $(w/H)[1-2\delta N-4\delta wH]/(2z)$ and $-\delta W/z$, where $z=(\lambda+\delta w{**}2)$.

21. The determinants of the Hessians implied by the estimates satisfy the conditions for concavity of the utility function. Also, essentially all observations are in the region of positive marginal income utility of income (=consumption in this model), a condition not required by the quadratic utility function. Estimating the percentage of the observations falling into the region of positive marginal utility of leisure is more difficult because that marginal utility is stochastic in our model; hence, even though the mean

of α is negative (as it should be), some part of its distribution is positive with probability one. We ignore this problem.

22. The cutoff values for 23 and 11 d.o.f. are, respectively, 35.2 and 19.7.

23. As show below, the fit improves substantially when the wage equation is dropped from the model, implying that the fit of the wage equation--not represented in Table 4--distorts, to some extent, the fit of the choice data.

24. The same specification of the equation as that shown in Table 3 is employed.

25. The variables used to predict non-transfer nonlabor income are age, age squared, education, South, children less than 18, children less than 5, race, fair or poor health, household size, whether the woman is divorced or separated, and whether the woman is never-married (the residual marital status category is widowed). The last three of these variables are used for identification--the household size variable because earnings of other adults in the n this section we first show simulated effects of changes in the budget constraint and policy parameters on labor supply and program participation, in order to illustrate the implications of our estimates for the cumulative marginal tax rate problem in multiple programs; an

26. The simulations are performed separately on each individual in the sample, and then averaged over all individuals.

27. We assume the EITC is treated as "negative taxes" by the AFDC program, i.e., it is treated as income. Hence it has no effect on take-home income while on AFDC.

28. We inflate all 1980 dollar figures to 1984 for comparability with the date of the SIPP data. An alternative procedure would be to apply our estimated parameters to the 1980 CPS itself instead. However, we wish to minimize the use of the CPS because it gathers AFDC participation information only on an annual basis (i.e., whether on AFDC during an entire year) and because welfare participation is significantly underreported. Unfortunately, the SIPP only began in 1984 and hence was not collected in 1980.

REFERENCES

Albright, R.; S. Lerman; and C. Manski. "Report on the Development of an Estimation Program for the Multinomial Probit Model." Report Prepared by Cambridge Systematics for the Federal Highway Administration, 1977.

Danziger, S.; Haveman, R.; and R. Plotnick. "How Income Transfers Affect Work, Savings, and the Income Distribution." Journal of Economic Literature 19 (September 1981): 975-1028.

Fraker, T. and R. Moffitt. "The Effect of Food Stamps on Labor Supply: A Bivariate Selection Model." Journal of Public Economics 35 (February 1988): 1-24.

Hausman, J. "The Econometrics of Non-linear Budget Sets." Econometrica 53 (November 1985): 1255-1282.

Lerman, S. and C. Manski. "On the Use of Simulated Frequencies to Approximate Choice Probabilities." In Structural Analysis of Discrete Data with Econometric Applications, eds. C. Manski and D. McFadden and C. Manski. Cambridge: MIT Press, 1981.

Levy, F. "The Labor Supply of Female Heads, or AFDC Work Incentives Don't Work Too Well." Journal of Human Resources 14 (Winter 1979): 76-97.

McFadden, D. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." Econometrica 57 (September 1989): 995-1026.

McFadden, D. and P. Ruud. "Estimation by Simulation." Review of Economics and Statistics 76 (November 1994): 591-608.

Moffitt, R. "An Economic Model of Welfare Stigma." American Economic Review 73 (December 1983): 1023-1035.

__________. "The Econometrics of Piecewise-Linear Budget Constraints." Journal of Business and Economic Statistics 4 (July 1986): 317-328.

__________. "Incentive Effects of the U.S. Welfare System: A Review." Journal of Economic Literature 30 (March 1992): 1-61.

Moffitt, R. and B. Wolfe. "The Effect of the Medicaid Program on Welfare Participation and Labor Supply." Review of Economics and Statistics 74 (November 1992): 615-626.

Pakes, A. and D. Pollard. "Simulation and the Asymptotics of Optimization Estimators." Econometrica 57 (September 1989): 1027-1057.

Smeeding, T. "Alternative Methods for Valuing Selected In-Kind Benefits and Measuring Their Effect on Poverty." Technical Paper No. 50. Washington: Bureau of the Census, 1982.

Steinberg, B. "The Labor Supply Effects of Government Transfers in Cash and in Kind." Mimeographed, Vanderbilt University, 1989.

U.S. Department of Health and Human Services. Characteristics of State Plans for Aid to Families with Dependent Children: 1985 Edition. Washington:

Government Printing Office, 1985.

U.S. House of Representatives. "Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means." Washington: Government Printing Office, 1987.

Zabalza, A.; C. Pissarides; and M. Barton. "Social Security and the Choice Between Full-Time Work, Part-Time Work, and Retirement." Journal of Public Economics 14 (October 1980): 245-276.

Appendix Table A-1

Means and Standard Deviations of

Variables used in the Analysis

	Mean	Standard Deviation
<u>Control Variables</u>		
Years of education	11.48	2.50
Age	33.81	8.93
No. Children less than 18	2.06	1.24
No. Children less than 5	0.53	0.76
White dummy	0.61	0.49
South dummy	0.35	0.48
Poor or Fair Health dummy	0.14	0.34
State Unemployment Rate	7.71	1.83
SMSA dummy	0.59	0.49
Percent State Employment in Service Sector	0.21	0.03
State Monthly AFDC Administrative Expenses per recipient/100	4.79	1.73

Table A-1 (cont'd)

Budget Constraint Variables[a]		
Hourly wage (including mean wages of nonworkers)	$ 5.20	$ 2.39
Nonlabor income	4.36	15.46
Earnings at H=20	104.00	47.80
Earnings at H=40	208.00	95.60
AFDC Benefit at H=0	63.53	41.01
AFDC Benefit at H=20	13.74	25.65
AFDC Benefit at H=40	2.20	11.83
Food Stamp Benefit at H=0	41.49	17.67
Food Stamp Benefit at H=20	31.91	19.69
Food Stamp Benefit at H=40	15.45	19.10
Housing Benefit at H=0	94.67	24.08
Housing Benefit at H=20	81.39	27.53
Housing Benefit at H=40	56.58	34.39
Positive Taxes at H=20	8.01	5.98
Positive Taxes at H=40	23.94	20.98
Medicaid Benefit[b]	28.01	22.22
Private health insurance Benefit[c]	7.37	8.19

Notes:

a. All variables except wage are weekly. The AFDC, Food Stamp, and housing benefits shown are means over the sample if participating in all three programs; means for the other seven combinations are not shown for brevity. For nonworkers, earnings, benefit, and tax variables are evaluated at the individual's expected wage.
b. Expected medical expenditures if on Medicaid and if on AFDC.
c. Expected medical expenditures if off AFDC, equal to product of probability of private coverage and expected expenditures if covered.

Table A-2

Simulated Responses to Changes

in the Budget Constraint: MSM with 500 Draws, Wage IV

	Program Participation(%)		Work Hours Distribution(%)			Mean Hours Worked	Cost Change (%)
	P_A	P_F	NW	PT	FT		
Baseline	27.6	37.6	39.4	10.9	49.7	22.1	-
Decrease in AFDC tax rate from 1.00 to .50[a]	28.5	37.7	38.3	12.0	49.7	22.3	1
Decrease in AFDC and Food Stamp tax rates to .10	34.0	42.9	34.0	14.7	51.3	23.4	85
Wage Increase of $1	24.8	34.6	34.4	10.7	55.0	24.1	194
Minimum Wage of $5	24.4	34.5	33.6	11.5	54.9	24.3	86
Wage-Rate Subsidy[b]	24.8	34.9	34.6	11.1	54.3	24.0	74
Increase in EITC[c]	25.2	36.4	33.7	19.3	47.0	22.7	54
Universal Work Subsidy[d]	24.5	34.2	35.5	15.3	49.2	22.7	-13

Notes: See Table 6.

Table B-1

Illustrative Cumulative Tax Rates[a]

	Weekly Income			Tax Rate from H=0 to H=20	Tax Rate from H=20 to H=40
	H = 0	H = 20	H = 40		
<u>California</u>					
Earnings	0	104	208	-	-
AFDC Benefit	124	30	0	.90	.29
Food Stamp Benefit	16	16	0	0	.15
Housing Benefit	138	132	107	.06	.24
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	278	253	268	-	-
Cumulative Tax Rate	-	-	-	1.24	.86
<u>Minnesota</u>					
Earnings	0	104	208	-	-
AFDC Benefit	117	25	0	.88	.24
Food Stamp Benefit	19	19	0	0	.18
Housing Benefit	97	91	64	.06	.26
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	233	210	225	-	-
Cumulative Tax Rate	-	-	-	1.22	.85

Table B-1 (cont'd)

	Weekly Income			Tax Rate from H=0 to H=20	Tax Rate from H=20 to H=40
	H = 0	H = 20	H = 40		
<u>Ohio</u>					
Earnings	0	104	208	-	-
AFDC Benefit	60	0	0	.58	0
Food Stamp Benefit	44	30	4	.13	.29
Housing Benefit	87	71	37	.15	.33
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	191	176	202	-	-
Cumulative Tax Rate	-	-	-	1.14	.75
<u>Kansas</u>					
Earnings	0	104	208	-	-
AFDC Benefit	76	0	0	.73	0
Food Stamp Benefit	38	31	0	.07	.30
Housing Benefit	68	64	31	.04	.32
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	82	170	192	-	-
Cumulative Tax Rate	-	-	-	1.12	.79

Table B-1 (cont'd)

	Weekly Income			Tax Rate from H=0 to H=20	Tax Rate from H=20 to H=40
	H = 0	H = 20	H = 40		
<u>Alabama</u>					
Earnings	0	104	208	-	-
AFDC Benefit	23	0	0	.22	0
Food Stamp Benefit	48	31	0	.16	.30
Housing Benefit	94	67	34	.26	.32
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	165	173	195	-	-
Cumulative Tax Rate	-	-	-	.92	.79
<u>Texas</u>					
Earnings	0	104	208	-	-
AFDC Benefit	30	0	0	.29	0
Food Stamp Benefit	48	32	0	.16	.31
Housing Benefit	103	79	46	.23	.32
Taxes	0	-8	-26	.08	.17
Work Expenses	0	-21	-21	.20	0
Net Income	181	186	207	-	-
Cumulative Tax Rate	-	-	-	.95	.80

Notes:
a. Calculations assume wage of $5.20 and nonlabor income of $4. Child care expenses are set to zero. Participation in all three programs is assumed. The tax rate from H=0 to H=20 is calculated as column (1) minus column (2) divided by $104. The tax rate from H=20 to H=40 is calculated as column (2) minus column (3) divided by $104.

Table 1

Distribution of the Sample by Labor Supply

and Program Participation

	Labor Supply			Row Totals
	Nonworkers	Part-time	Full-time	
$P_A=0$, $P_F=0$, $P_R=0$	76 (15) (20)	57 (11) (56)	383 (74) (80)	516 (100) (53)
$P_A=1$, $P_F=0$, $P_R=0$	9 (53) (2)	1 (6) (1)	7 (41) (1)	17 (100) (2)
$P_A=0$, $P_F=1$, $P_R=0$	36 (41) (9)	20 (23) (20)	32 (36) (7)	88 (100) (9)
$P_A=1$, $P_F=1$, $P_R=0$	162 (93) (41)	11 (6) (11)	2 (1) (0)	175 (100) (18)
$P_A=0$, $P_F=0$, $P_R=1$	10 (16) (3)	6 (10) (6)	46 (74) (10)	62 (100) (6)
$P_A=1$, $P_F=0$, $P_R=1$	3 (100) (1)	0 (0) (0)	0 (0) (0)	3 (100) (0)
$P_A=0$, $P_F=1$, $P_R=1$	14 (52) (4)	4 (15) (4)	9 (33) (2)	27 (100) (3)
$P_A=1$, $P_F=1$, $P_R=1$	77 (96) (20)	2 (3) (2)	1 (1) (0)	80 (100) (8)
Column Totals	387 (40) (100)	101 (10) (100)	480 (50) (100)	968 (100) (100)

Notes: P_A, P_F, and P_R are dummies for participation in AFDC, Food Stamps, and subsidized housing, respectively. Entries in table show cell frequency, followed by row percent, followed by column percent.

Table 2

Simulated Maximum Likelihood Estimation Results
(500 Draws)

	Tastes for Work (α)	AFDC Costs (ψ_A)	Food Stamp Costs (ψ_F)	Housing Costs (ψ_R)	Wage Eqn. (w)
No. children less than 18	-0.26* (0.10)	-0.27 (0.22)	-0.27 (0.25)	0.05 (0.34)	--
No. children less than 5	-0.36* (0.14)	--	--	--	--
South	0.56* (0.19)	-0.46 (0.49)	0.68 (0.52)	-0.60 (0.88)	0.03 (0.05)
Education	0.02 (0.05)	0.44* (0.11)	0.55* (0.14)	0.12 (0.12)	0.08* (0.01)
Age	0.13 (0.08)	0.18* (0.03)	0.18* (0.04)	0.22* (0.07)	0.01* (0.02)
Age Squared/100	-0.15 (0.10)	--	--	--	-0.12* (0.02)
Fair or poor health	-0.50* (0.25)	-0.63 (0.67)	-1.25* (0.71)	-2.40* (1.27)	-0.16* (0.06)
Race (1=white)	0.26 (0.17)	1.75* (0.50)	2.20* (0.58)	6.42* (1.89)	0.03 (0.05)
State unemployment rate	-0.08* (0.04)	-	-	-	-
SMSA	-	-	-	-	0.03 (0.04)
State Service Employment Pct.	-	-	-	-	2.17* (0.77)
State AFDC Admin. expenses/100	-	0.57* (0.14)	-0.01 (0.14)	-0.61* (0.28)	-
Constant	1.62 (1.52)	-6.44* (2.03)	-7.51* (2.34)	2.55 (3.25)	-1.88* (0.32)
σ	1.76* (0.24)	4.37* (0.43)	5.51* (0.76)	9.96* (2.69)	0.51* (0.01)

Table 2 (cont'd)

Utility Function Parameters

β_{HH} [a]	3.86* (0.45)	ϵ_W	=	1.82
β_{YY} [b]	3.77* (1.38)	ϵ_Y	=	-0.21
λ	0.07* (0.03)			
γ_R	0.10 (0.09)			
γ_{MED}	0.48* (0.23)			
γ_{PHI}	0.62 (0.65)			

Correlation Matrix of Errors

	ϵ_A	ϵ_F	ϵ_R	ϵ_W
ϵ_α	-0.13* (0.06)	0.08 (0.06)	0.07 (0.06)	0.01 (0.09)
ϵ_A	-	0.69* (0.03)	0.19* (0.04)	0.04 (0.06)
ϵ_F	-	-	0.30* (0.05)	0.22* (0.05)
ϵ_R	-	-	-	0.08* (0.05)

Simulated Log Likelihood = -2249.1

Simulated Log Likelihood = -1822.3 (choices only)

Chi-Squared = 53.83

Notes:

Standard errors in parentheses
* = significant at the 10 percent level

Sample size = 968
Normalization: $\tau=2$
Parameters in three cost equations (including sigmas) are divided by 10.
[a] Multiplied by 100
[b] Multiplied by 10000

Table 3

Simulated Maximum Likelihood Estimation Results:
Without Public Housing
(500 Draws)

	Tastes for Work (α)	AFDC Costs (ψ_A)	Food Stamp Costs (ψ_F)	Wage Eqn. (w)
No. children less than 18	-0.16* (0.10)	-0.15 (0.22)	-0.20 (0.22)	--
No. children less than 5	-0.31* (0.14)	--	--	--
South	0.90* (0.22)	-1.60* (0.52)	-0.25 (0.50)	0.04 (0.05)
Education	0.02 (0.05)	0.34* (0.12)	0.44* (0.13)	0.08* (0.01)
Age	0.10 (0.09)	0.18* (0.04)	0.14* (0.04)	0.10* (0.02)
Age Squared/100	-0.13 (0.11)	--	--	-0.12* (0.02)
Fair or poor health	-0.59* (0.27)	0.26 (0.71)	-0.41 (0.69)	-0.18* (0.07)
Race (1=white)	0.33 (0.18)	1.13* (0.52)	1.45* (0.53)	0.03 (0.04)
State unemployment rate	-0.01 (0.04)	-	-	-
SMSA	-	-	-	0.03 (0.04)
State Service Employment Pct.	-	-	-	2.19* (0.78)
State AFDC Admin. expenses/100	-	0.44* (0.14)	-0.17 (0.13)	-
Constant	-2.36 (1.78)	-3.99* (2.07)	-4.54* (2.14)	-2.03* (0.33)
σ	1.65* (0.27)	4.10* (0.44)	4.67* (0.69)	0.51* (0.01)

Table 3 (cont'd)

Utility Function Parameters

β_{HH}[a]	3.92* (0.47)	ϵ_W	=	1.94
β_{YY}[b]	3.19* (1.49)	ϵ_Y	=	−0.18
λ	0.05* (0.02)			
γ_{MED}	0.50* (0.25)			
γ_{PHI}	0.73 (0.70)			

Correlation Matrix of Errors

	ϵ_A	ϵ_F	ϵ_W
ϵ_α	−0.00 (0.07)	0.06 (0.08)	−0.07 (0.12)
ϵ_A	–	0.58* (0.06)	−0.01 (0.07)
ϵ_F	–	–	0.24* (0.06)

Simulated Log Likelihood = −1826.5

Simulated Log Likelihood = −1391.4 (choices only)

Chi-Squared = 28.1

Notes:

Standard errors in parentheses
* = significant at the 10 percent level

Sample size = 968
Normalization: τ=2
Parameters in two cost equations (including sigmas) are divided by 10.
[a] Multiplied by 100
[b] Multiplied by 10000

Table 4

Actual and Fitted Distributions

of Labor Supply and Program Participation:

SML Estimates, 500 Draws

(Sample percentage distribution)

	Labor Supply			Row Totals
	Nonworkers	Part-time	Full-time	
		Actual		
P_A=0, P_F=0	8.9	6.5	45.4	59.7
P_A=1, P_F=0	1.2	0.1	0.7	2.1
P_A=0, P_F=1	5.2	2.5	4.2	11.9
P_A=1, P_F=1	24.7	1.3	0.3	26.4
Column Totals	40.0	10.2	49.6	100.0
		Fitted		
P_A=0, P_F=0	7.3	7.4	48.9	63.6
P_A=1, P_F=0	1.7	0.2	0.9	2.8
P_A=0, P_F=1	5.1	2.0	4.2	11.3
P_A=1, P_F=1	20.5	0.8	0.9	22.2
Column Totals	34.6	10.4	54.9	100.0

Table 5

Sensitivity Tests to the Specification

	SML		MSM			
	Small Model[a]	1000 Draws	Joint Wage Eqn.	Wage IV	Wage and Nonlabor Income IV	Narrow PT Def.[b]
β_{HH}[c]	3.56* (0.37)	4.30* (0.63)	4.05* (0.47)	4.41* (0.79)	4.60* (0.74)	3.43* (0.64)
β_{YY}[d]	4.05* (1.16)	2.73 (1.76)	4.32* (1.18)	2.25 (2.01)	4.38* (1.51)	4.66* (1.12)
ϵ_W	1.87	1.90	1.64	1.97	1.47	1.77
ϵ_Y	-0.24	-0.15	-0.22	-0.12	-0.20	-0.27
λ	0.06* (0.02)	0.08* (0.04)	0.08* (0.04)	0.05* (0.02)	0.05* (0.02)	0.09* (0.04)
γ_{MED}	0.70* (0.22)	0.66* (0.32)	0.78 (2.21)	0.82* (0.31)	1.91* (0.48)	2.33* (0.53)
γ_{PHI}	1.66* (0.63)	1.27 (0.88)	1.21 (1.18)	1.79* (0.81)	3.78* (1.03)	4.33* (1.04)
Simulated Log Likelihood	-1872.0	-1798.5	-1876.7	--	--	--
Simulated Log Likelihood (choices only)	-1428.1	-1365.2	-[e]	-1428.8	-1457.3	-1400.1
Chi-Squared	27.9	24.7	31.5	15.6	19.5	7.9

Notes:

Standard errors in parentheses
* = significant at the 10 percent level

Parameters in α, cost equations, and wage equations not shown.
Sample size = 968
Normalization: τ=2

[a] Regressor vector restricted to two children variables in α equation and administrative expense variable in cost equations.

[b] Part-time work defined as 11-to-29 hours per week; nonwork defined as 0-to-10 and full-time work defined as 30+. Wage and nonlabor income IV retained.

[c] Multiplied by 100

[d] Multiplied by 10000

[e] A choices-only log likelihood cannot be computed in this model.

Table 6

Simulated Responses to Changes

in the Budget Constraint: SML with 500 Draws

	Program Participation(%)		Work Hours Distribution(%)			Mean Hours Worked	Cost Change (%)
	P_A	P_F	NW	PT	FT		
Baseline	25.0	33.5	34.6	10.4	55.0	24.1	-
Decrease in AFDC tax rate from 1.00 to .50[a]	25.7	33.7	33.7	11.5	54.8	24.3	1
Decrease in AFDC and Food Stamp tax rates to .10	32.8	40.0	27.9	14.4	57.7	26.0	79
Wage Increase of $1	20.9	28.9	26.5	9.7	63.8	27.5	162
Minimum Wage of $5	19.1	26.8	22.4	10.8	66.8	28.9	128
Wage-Rate Subsidy[b]	20.3	28.3	24.8	10.6	64.6	28.0	89
Increase in EITC[c]	21.9	31.9	27.2	20.0	52.8	25.1	46
Universal Work Subsidy[d]	20.8	28.9	27.7	17.8	54.5	25.4	-3

Notes:

P_A = probability of participating in AFDC
P_F = probability of participating in Food Stamps
NW = probability of not working

PT = probability of working part-time
FT = probability of working full-time
Mean Hours Worked = 20*PT + 40*FT

[a] All income screens simultaneously eliminated (see Appendix B)

[b] Subsidy = .50*($6.00 - wage)

[c] A refundable tax credit equal to 30 percent of earnings up to a maximum $1500 annual credit, followed by a 20 percent phase-out rate.

[d] All AFDC and FSP deductions for work expenses are eliminated and replaced by a work subsidy defined by: Subsidy = $23 - .07*Income, and the subsidy is also offered to those off welfare.

Table 7

Out-of-Sample Comparisons of 1980-1984

Changes in AFDC Participation and Labor Supply

(Percentage-Point Changes)

	SIPP[a]		CPS and Administrative Data[b]
	SML	MSM	
Change In:			
AFDC Participation Rate	-10.9	-10.0	-8.5
Mean Hours of Work	0.5	0.6	0
Distribution of Hours of Work			
Nonwork	-0.1	-0.2	-0.1
Part-time	-2.5	-2.1	-0.1
Full-time	2.8	2.4	0.2
Percent of AFDC Recipients who do not work	14.6	15.8	10.0

Notes:

[a] Differences in simulations using 1984 SIPP but 1980 and 1984 program rules.

[b] Sources: AFDC participation rate: from CPS and administrative data, as reported in Moffitt (1992, Table 3; 1980 interpolated from 1979 and 1981). Hours of work mean and distribution: authors' tabulations from the March 1980 and March 1984 CPS. Percent of AFDC recipients who do not work: from AFDC administrative data, as reported in Moffitt (1992, Table 4; 1980 interpolated from 1979 and 1981).