Preliminary and incomplete

Identification and Estimation of Structural Models of Labor Supply and Program Participation*

Bradley T. Heim
Office of Tax Analysis, U.S. Treasury, and
Department of Economics, Duke University

and

Bruce D. Meyer
Harris School of Public Policy Studies,
University of Chicago and NBER

First Draft: September 2005
Current Version: March 2006

ABSTRACT

In this paper, we estimate a structural model of employment, hours and program participation choices of single women over the 1984-1996 period. During the 1980s and 1990s, tax and welfare policy dramatically altered the labor supply and program participation incentives of single mothers. We use this setting to explore identification in structural labor supply models. Through the judicious use of special samples (specific states, years, women with certain numbers of children), control variables, and separate coefficients for different types of income, we isolate the different sources of variation in the after-tax reward to work. We explore the role of the intensive hours choice versus the extensive work/nonwork decision, examine the role of the point-in-time shape of the tax schedule for a given demographic group versus changes over time, the tax treatment of children, and the role of functional form. We also provide substantive results on effects of the Earned Income Tax Credit (EITC) and welfare programs. Studies analyzing the effects of the EITC using difference-in-difference methods have found that hours per year among those working increased in response to the EITC expansions. This change occurred even though both the income effect of the larger credits and the substitution effect of the higher phase-out rates implied a decline in hours. We address these surprising EITC results as well as the effects of welfare budget set changes using our joint model of labor supply and program participation.

*We thank Brian Jenn for his invaluable work on an earlier version of this project.

1. Introduction

During the 1980s and 1990s, tax and welfare policy dramatically changed the labor supply and program participation incentives of single mothers. Tax credits for the working were sharply expanded through the Earned Income Tax Credit (EITC). Welfare benefits for those not working were cut, while welfare benefits for the working were often increased. There were also large changes in Medicaid and other programs. Taken together, these policy changes were associated with large changes in the employment, hours and program participation of single mothers. Employment of single mothers increased by more than ten percentage points, and at the same time, hours among working mothers increased during the time period studied. Welfare program participation initially rose, but subsequently fell by more than forty percent. This paper addresses substantive unanswered questions about the effects of these recent policy changes.

Existing work has documented the dramatic changes in work incentives, particularly the EITC, and has shown how these policies affected employment rates (see Hotz and Scholz 2003 for a review). This line of work has often ignored welfare participation, however, and done a poor job of analyzing hours changes. These mostly difference-in-differences analyses found that hours increased in response to the EITC expansions even though both the income effect of the larger credits and substitution effects of the higher phase-out rates implied a decline. Given that potential reductions in hours among the working is one of the major policy concerns about the EITC, these surprising results are of great interest.

A second line of research has examined changes in welfare rules and the effects of such changes on welfare caseloads (see Grogger and Karoly (2005) for a review). Most of these

papers[1] do not estimate the effect of tax changes on welfare participation, despite indications in the first line of work that the EITC played a large role (often the dominant one in the early years) in employment changes. Thus, while employment, hours and participation are clearly linked, they have not been related in the literature.

More generally, the effects of taxes, social insurance and welfare programs on labor supply is one of the most important effects of government policies. The U.S. raises the vast majority of its federal revenue through taxes on labor and the primary distortion of transfer programs is arguably altered work incentives.

In this paper, we estimate a structural model of employment, hours and program participation choices of single women over the 1984-1996 period. There are several reasons why a structural model is well suited to analyze joint labor supply and welfare participation decisions. Joint modeling of program participation and labor supply in this context is not simply solved by estimating two equations. To appropriately capture the incentives to participate or the incentives to work, one needs to account for the jointness of and interactions between the two decisions. Welfare participation alters the budget constraint and may lead to stigma or transaction costs, all of which affect the labor supply decision. Similarly, the welfare participation decision depends on welfare benefits received which in turn depend on labor supply and earnings.

In addition, some of the policy changes of the 1990s, particularly the EITC, are expected to increase employment but reduce hours among the working because they raise virtual incomes and implicit tax rates at the most common hours choices. The EITC creates substantial nonconvexities in the budget constraint and appears to have had most of its effect on nonlocal

[1] Grogger (2003, 2004) are exceptions.

work/nonwork decisions. To better analyze such policies, we need to use an approach that captures the shape of the entire budget set. Furthermore, to simulate possible future policy changes, it is preferable to have a model that accounts for labor force participation, labor supply, as well as program participation choices in a coherent way.

Further, we will use this estimation method and the policy setting of the 1980's and 1990's to explore the role of the features of structural labor supply models that allow identification. The drawback to structural models as typically implemented is that it is difficult to see the sources of identification in the models and examine the sensitivity of the results to alternative assumptions. Typically, all sources of variation in the data are lumped together with functional form assumptions making it unclear what drives the results.[2] In our case, however, given the large changes in policy during this period that differentially affected groups of individuals, through the judicious use of special samples (specific states, years, women with certain numbers of children), control variables, and separate coefficients for different types of income, we can isolate the different sources of variation in the after-tax reward to work.

To that end, we will explore the role of the intensive hours choice versus the extensive work/nonwork decision. We will examine the role of the point-in-time shape of the tax schedule for a given demographic group versus changes over time and the treatment of children. We will also examine samples without demographic differences that lead to tax differences and where there are no tax schedule changes, to examine the role of tax changes in identifying estimated

[2] See, for example, Hausman (1981), Hoynes (1996), and van Soest (1995), among many others, in which functional form, wage variation, income variation, and variation in tax and social policies across states and/or income levels combine in some manner to identify parameters.

coefficients, as well as the role of functional form. This latter issue will also be addressed by trying more than one functional form for the utility function.

In an important sense, then, the approach in the paper will differ sharply from that in many structural estimation papers where the sources of identification are unclear. By focusing on different sources of identifying variation, trying alternative functional forms, and relaxing distributional assumptions, we hope to make clear what drives our results. The integration of natural experiment ideas with the sensible functional forms and economic restrictions of structural models is a key goal of this paper.

The paper proceeds as follows. We summarize the institutional background in Section 2, and in section 3 we discuss previous research. We provide an overview of our approach in Section 4. In Section 5 we explain our choice of data and in Section 6 describe the econometric methods in detail. Section 7 provides our initial results, while Section 8 concludes.

2. Summary of Institutional Background

The period we analyze, 1984-1996, was one of sharp changes in policy. The maximum EITC credit increased from $500 to $3556 leading to an an almost twelve-fold increase in real credits paid through the EITC. Medicaid eligibility for adults and children was expanded leading to a 73 percent increase in the number of children who were Medicaid recipients. Welfare benefits for those not working were reduced, but the implicit tax rate for those working was also cut, often raising benefits for working recipients. Many states implemented some type of waiver that altered their welfare program. Training programs were expanded and their focus changed,

and four significant federal childcare programs were added between 1988 and 1990.

The combined effect of these program changes greatly increased the incentive for single mothers to enter the workforce and, in most but not all cases, made welfare receipt less attractive. The effect on hours worked among those working of the combined cut in welfare guarantees but reduced implicit tax rates is theoretically ambiguous. However, the predicted effect of the EITC expansions on hours is predicted to be negative given the income effect of the higher credit and the higher implicit tax rates as the phase-out tax rates were increased.

Coincident with the policy changes were large increases in the employment of single mothers between 1984 and 1996. The pattern of welfare receipt had a more complicated pattern. Caseloads increased greatly through 1994 when a precipitous drop in the number of welfare cases began. The combination of work and welfare became more typical, so that by 1996 most of those who received welfare also worked.

3. Previous Research

The research on the effect of EITC on single mothers has focused on the credit's effects on the work/nonwork decision (see Hotz and Scholz 2003 for a nice summary). This work tends to find large effects of the EITC on employment. Several papers have also examined hours of those who work using difference-in-difference methods even though such methods are not especially well suited for the task (Eissa and Liebman 1996, Meyer and Rosenbaum 2001, Meyer 2002). These papers find the surprising result that hours among the working increased, despite the negative income and substitution effects for most eligibles due to the expansion of the credit.

There is a much smaller literature on the effects of the EITC on welfare participation (Grogger 2003, 2004) which finds that the EITC expansions reduced welfare participation. There is also related work on the British tax and transfer system by Blundell et al. (2000, 2005) and Brewer et al. (2003, 2005).

There is a much more voluminous literature on the effects of welfare reform and related policies on employment and welfare participation that is well summarized in Grogger and Karoly (2005). Much of this research goes beyond the time period on which we focus to the period where welfare reform was the dominant policy. We cannot adequately summarize this literature here, but it finds that work requirements, time limits, and the improvements in the economy all decrease welfare use, but lower implicit tax rates on work for welfare recipients often increase welfare use. All of these factors also tend to increase employment, but there is much less evidence on hours effects of welfare reform policies.

More broadly, this work is related to the large literature on structural estimation of labor supply and program participation, particularly studies by Keane and Moffitt (1998) and Hoynes (1996). In much of this literature, wage effects have generally been identified through cross-sectional wage variation and functional form assumptions, and rarely were different functional forms and sources of identifying variation compared. This may be particularly important in the current setting since, if wages are believed to be exogenous, the substantial correlation between part-time work and low wages for women may lead to the conclusion that there is a substantial wage elasticity for single women. It may be, however, that we are observing characteristics of jobs per se rather than people's decisions, because part-time jobs may just pay poorly. Furthermore, wages may be correlated with unobserved tastes for leisure. Thus, in this paper, we

examine the extent to which exogeneity and functional form assumptions drive resulting estimates in a structural model, by focusing on separate components of the variation in after-tax wages to identify wage effects.

4. Overview of Our Approach

We examine the employment, hours and program participation choices of single women over the 1984-1996 period. These years include dramatic change in tax and welfare policy that sharply altered labor supply and program participation incentives of single mothers. We incorporate state and federal income taxes including state and federal EITCs, payroll taxes, Medicaid, measures of changes in private health insurance offers, Aid to Families with Dependent Children (AFDC) welfare benefit levels, Food Stamps and the implicit tax rates on earnings due to Food Stamps and AFDC. We will use our model to estimate tax and welfare program effects on work decisions, work hours and program participation. These estimates will provide labor supply elasticities that we can apply to other situations. We will also estimate the extent of transaction costs or stigma costs of program participation as well as the costs of work.

We use 13 years of Current Population Survey (CPS) and Survey of Income and Program Participation (SIPP) data. We examine the years 1984-1996 because this period was one of dramatic changes in policy, but one when a complicated, but known budget constraint reasonably characterizes the choices facing these women. As welfare reform progressed, it became increasingly difficult to model the policy changes in terms of a budget constraint because of

diversion and other policies designed to reduce caseloads that are not easily summarized.[3]

The base model we estimate includes a number of methodological improvements relative to previous work. We build on structural models like Keane and Moffitt (1998) and Hoynes (1996) that start from a direct utility function and assume individuals choose from a discrete set of alternatives. We then non-parametrically estimate the distribution of wages and use our model of the employment decision to determine the selection of women into work. We allow preferences and wages to be related in a very flexible way. We also account for the pronounced under-reporting of welfare receipt in survey datasets.

Our model relies on using a discrete set of hours and wage points which simplifies the model in a way that allows us to explore many key issues. Integration over a continuous wage distribution (for nonworkers) with several complicated tax and benefit calculations at each value is otherwise required. This integration is much more computational intensive than our approach of selecting the appropriate tax and benefit values from a stored matrix and weighting them appropriately.

Because we have a very large sample and long time period, we can incorporate features in our analyses that past researchers could not. For example, we are able to account for characteristics of state welfare programs and labor markets that we could not summarize with control variables. In addition this approach enables us to analyze subpopulations and time-periods of special interest that allow us to explore the sources of identification in our model and more easily check the robustness of our results to alternative functional form assumptions.

---

[3] As a check on our estimates, we will also estimate models excluding the last couple of years of this period.

To enable us to do so, we have selected a sample (single women) and a time period (1984-1996) when there are very sharp changes in policy that affect certain groups, but not others. During this period, the EITC raised the after-tax reward to working for single mothers relative to single childless women by almost $2000 (compared to a mean after tax and benefit net return of under $7,500). The changes differed sharply in their size and timing by state and family size because of changes in state taxes, changes in tax treatment of different size families, and differences in state living costs. Furthermore, states differ with respect to their level of benefits, the amount of earnings that is disregarded, the rate at which benefits are reduced after the disregard, and the additional benefit for increments to family size, and these program parameters change substantially over time. In a sense, then, this aspect of the paper has the flavor of a natural experiment.

While the choice between income and leisure is one of the most basic economic decisions, we expect that our results will have some broader insights. In particular, we hope that our results may spur others to further explore identification in other types of structural models.

5. Data

We use March Current Population Survey (CPS) data for the years 1984-1996. We limit the sample to single women (widowed, divorced, and never married) who are between 19 and 44 years old and not in school. Women who were ill or disabled during the previous year or who had positive earned income but zero hours of work are also excluded. The strengths and weaknesses of the CPS are well known, but its large sample size (well over 100,000 observations

on single women over our period) is a clear advantage.[4] The key variables we use are employment, hours, wages, and welfare participation during the year.

As is the case with all household surveys with which we are familiar, receipt of many social insurance and welfare programs is under-reported. In recent years the fraction of months of AFDC or TANF received that are not reported has risen to over thirty percent in the Current Population Survey (CPS) and to almost this high a rate in the Survey of Income and Program Participation (SIPP) (see Primus et al. 1999, Bavier 1999).[5] Most of the under-reporting takes the form of not reporting any months at all (Moore, Marquis and Bogen 1996). It appears that about half of eligibles that do not report receipt are in fact non-reporting recipients. Under-reporting can bias upwards estimates of the transaction costs and stigma of welfare participation because these costs provide an explanation for what appears to be low takeup among eligibles.

Estimating under-reporting rates for each year for an indicator of welfare receipt anytime during the year is not straightforward. We have good estimates for under-reporting of months of receipt and dollars obtained by comparing weighted sample counts to administrative totals. Unfortunately, administrative data on the number of families receiving welfare in a given year are not available. Our current approach is to use the estimate of the fraction of under-reporting that is due to not reporting at all from Moore et al.(1996) and apply it to the estimates of the fraction of months not reported. We are also far along in arranging matching to administrative

---

[4] We have also done some preliminary analyses with the Survey of Income and Program Participation (SIPP) and may consider using it as well. The SIPP has a significantly smaller sample, but has better reporting of welfare receipt.

[5] See Moore, Stinson and Welniak (1997), Bound, Brown and Mathiowetz (2001), and Meyer and Sullivan (2003) for discussions of evidence on the under-reporting of social insurance and welfare program receipt in survey datasets.

microdata for three states. These data would allow us to improve our correction for misreporting by allowing us to infer how misreporting varies with in individual characteristics.

6. The Model and Likelihood Function for Employment, Hours and Program Participation

In building our model of labor supply and program participation, we start from a utility function that incorporates the stigma or transaction costs of program participation. We allow individual characteristics such as demographic variables to enter the preference parameters, and incorporate in the budget constraint the complexity of welfare programs and state and federal income taxes. However, to simplify the model we approximate individuals' choices by assuming that they decide between J different hours choices, and we approximate the wage distribution with K values. This last simplification allows us to estimate the wage distribution nonparametrically and allows us to include a very flexible relationship between wages and preferences for work. Furthermore, the approximation greatly simplifies computation, allowing us to explore the sources of identification in these models.

Utility

Specifically, we assume utility is a function of income Y, hours H worked in the year, an indicator for welfare participation P,[6] individual characteristics Z, a vector of parameters θ, and a term denoting unobserved tastes for leisure ϵ. Thus, we can write

[6] We assume that a woman participates in Food Stamps if and only if she participates in AFDC. We assume that a woman always participates in Medicaid if eligible (we could relax this later and add a second transaction cost term).

$$U= U(Y, H, P, Z(W), \theta, \epsilon).$$

The only nonstandard aspect to this utility function is that we write individual characteristics as Z(W) to make clear that Z may include the wage. Letting U take the Stone-Geary form for now (we may generalize this later), the utility function is

$$U=(\beta/(1+\beta))\ln(\mu - H - \rho P - \phi E) + (1/(1+\beta))\ln(Y + \eta) ,$$

where ρ captures transaction costs or stigma measured in hours, and ϕ captures the fixed costs of working. E is an indicator for employment; thus, if hours exceed zero, E=1, otherwise E=0. μ and η are respectively, the maximum hours available to a person and the minimum income consumption of a person. We allow tastes for leisure to vary in the population, i.e.

$\beta=\exp\{Z'\gamma + \beta_0 + \beta_1\epsilon\}$ so that the utility function becomes

$$U=(\exp\{Z'\gamma+\beta_0+\beta_1\epsilon\}/(1+\exp\{Z'\gamma + \beta_0 + \beta_1\epsilon\}))\ln((\mu-H-\rho P-\phi E)$$
$$+ (1/(1+ \exp\{Z'\gamma + \beta_0 + \beta_1\epsilon\})) \ln(Y + \eta),$$

where ϵ denotes unobserved heterogeneity in the taste for leisure.

An important modeling choice here is where individual characteristics, state or year dummies are incorporated. Where they are included should, of course, depend on what they are

intended to capture. While demographic variables and state and year dummies enter the taste for leisure, β, in our base case, we will also allow demographics such as the number of young children to enter work costs, ϕ. We will also consider allowing costs to vary with hours. In addition, it may make sense to include year dummies in the transaction costs or stigma of welfare participation, ρ, to reflect the increased difficulty of obtaining welfare benefits and the less favorable public attitudes toward welfare receipt.

The Components of Income and the Budget Constraint

Income is calculated as pre-tax earnings minus net taxes paid, plus AFDC and Food Stamps, plus Medicaid benefits, plus the value of employer provided health coverage, plus other income. We calculate the earnings, taxes, and benefits for a given individual incorporating family composition (number and ages of children), and characteristics of state and federal policies at the time. We allow the coefficients on the different components of income to differ, since income from different sources may be valued differently and details of the tax and benefit system may not be well known to taxpayers and recipients. In equations, income is calculated as

$$Y(H, P) = WH + \alpha_1 \text{ net taxes paid} + \alpha_2 \text{AFDC and Food Stamp benefits}$$
$$+ \alpha_3 \text{Medicaid coverage valued at cost}$$
$$+ \alpha_4 \text{employer provided health coverage valued at cost} + \text{other income.}$$

We expect that α_1 will be negative and note that the magnitude of the α coefficients would all be 1 if all income sources were valued the same and if w and other income are not correlated with

the taste for leisure. We will also estimate some models splitting taxes into state and federal taxes.[7] We should emphasize that net taxes paid, AFDC and Food Stamp benefits, and Medicaid coverage are a function of WH. AFDC plus Food Stamp benefits are also a function of P, i.e. they are zero if P=0. We calculate the values of the variables net taxes paid, AFDC and Food Stamp Benefits, and Medicaid coverage for all the earned income levels equal to each possible WH. These values are stored for each observation, speeding computation.

Implicit in this formulation of the choice problem of single women is that consumption equals income. While this is a questionable approximation for many demographic groups, for single mothers this assumption is fairly reasonable. Meyer and Sullivan (2003) emphasize that assets and liabilities of single mothers tend to be extremely low.

We calculate real income and benefits across states using a cost of living index which depends on state housing costs. Work and hours decisions should depend on the real return to work, not the nominal return. For example, if $300 a month can be used to rent two extra bedrooms in Texas, but cannot even get you one more bedroom in California, the value of an extra dollar of income provided by a tax credit is much lower in California than in Texas, under the plausible assumption that location amenities do not change with the number of bedrooms.

Likelihood Functions

With a few more definitions, we can construct a log-likelihood function to allow the estimation of our parameters. Our data consist of a sample of observations indexed by i=1,...,M

[7] We do this to examine the estimation of the coefficient on federal taxes in special situations where it is identified only through the point in time shape of the tax schedule .

with observed hours worked per year $H_i = h$, participation indicator $P_i = p$, and, if $h>0$, hourly wages $W_i = w$. Then the log-likelihood is

(1) $\log L = \Sigma_{i|Hi=0} \log\{\text{Prob}[H_i =0, P_i =p]\}$

$+ \Sigma_{i|Hi>0} \log\{\text{Prob}[H_i =h, P_i =p| W_i=w] \text{Prob}[W_i =w]\}$.

Let hours worked per year, $h \in \{H^1, H^2,..., H^J\}$, where J is the number of possible hours values. Let $p \in \{0,1\}$, 1 if a person participates in welfare and 0 otherwise. Let $w \in \{w^1, w^2,..., w^K\}$, where K is the number of possible wage levels. For computation reasons we parameterize $\text{Prob}[W_i =w^k]$ as $\exp\{\omega^k\} /(1 + \Sigma_{k=1}^{K-1} \exp\{\omega^k\})$, k=1,..., K-1 and $\text{Prob}[W_i =w^K]$ as $1/(1 + \Sigma_{k=1}^{K-1} \exp\{\omega^k\})$.

We numerically integrate over the distribution of ϵ by taking random draws from its distribution. This method is usually called simulated maximum likelihood estimation. Let N be the number of draws of ϵ for each observation. Then we can write the zero hour probability terms in the log-likelihood function above as

$\text{Prob}[H_i =0, P_i =p] = \Sigma_{n=1}^{N} N^{-1} \Sigma_k \text{Prob}[W_i =w^k]\{\text{Prob}[U(Y, 0, p, Z, \theta| \epsilon_i^n, W_i =w^k)$

$>U(Y, h', p', Z, \theta | \epsilon_i^n, W_i =w^k)$ for all h', p' not equal to 0, p}.

For H_i not equal to zero we have

$\text{Prob}[H_i =h, P_i =p |W_i =w] = \Sigma_{n=1}^{N} N^{-1}\{\text{Prob}[U(Y, h, p, Z, \theta | \epsilon_i^n, w)$

$>U(Y, h', p', Z, \theta | \epsilon_i^n, w)]$ for all h', p' not equal to h, p}.

Likelihood Approximation

One difficulty with the above formulation is that the above probabilities are either 0 or 1 and thus not continuously differentiable in their parameters. A standard computational trick used by McFadden (1989), Keane and Moffitt (1998), Brewer et al. (2005) is to approximate Prob[$U(Y, h, p, Z, \theta \mid \epsilon_i^n, w) > U(Y, h', p', Z, \theta \mid \epsilon_i^n, w)$] for all h', p' not equal to h, p as $\exp[U(Y, h, p, Z, \theta \mid \epsilon_i^n, w)/\tau]/\{\Sigma_j \Sigma_m \exp[U(Y, H^j, m, Z, \theta \mid \epsilon_i^n, w)/\tau]\}$ for small τ. As $\tau \to 0$ this expression goes to 0 or 1. This modification is equivalent to adding an extreme value error term to the utility of each alternative. This computational trick smooths the likelihood function and thus speeds convergence. Now, we can calculate analytic derivatives to speed convergence (and check our programming). Alternatively, one can consider the additional error terms as optimization error and take the model estimated to be the true model.

We include dummies for each wage value in z to control in the most flexible way for any relationship between wages and preferences. While differences in wages still enter through the budget constraint, preference parameters are not identified through differences in wages across individuals. Thus, this approach provides a way of examining estimates when the role of wages in the estimation process is circumscribed. One can think of this step as analogous to controlling for the main effect in difference-in-difference studies where the key coefficient is on an interaction of this main effect and another variable. The full parameter vector to be estimated is $\theta=(\alpha_1,..., \alpha_4, \beta_0, \beta_1, \gamma, \rho, \phi, \mu, \eta, \omega^1,..., \omega^{K-1})$, where γ is a vector.

Under-reporting of Welfare Receipt

We also modify this model to account for the underreporting of welfare receipt.

Underreporting is likely to bias upwards estimates of ρ, the transaction costs and stigma of welfare participation. We can account for underreporting by reinterpreting P_i to be the observed participation status, while P_i^* is the true participation status. Let R_t be the fraction of those receiving welfare in year t that truly report welfare receipt. Then, the earlier expressions given above are for the true probabilities P_i^* rather than observed probabilities P_i, and in the likelihood function one replaces

$\text{Prob}[H_i = 0, P_i = 1]$ with $\text{Prob}[H_i = 0, P_i^* = 1]\ R_t$ and

$\text{Prob}[H_i = 0, P_i = 0]$ with $\text{Prob}[H_i = 0, P_i^* = 0] + \text{Prob}[H_i = 0, P_i^* = 1][1 - R_t]$.

in the terms where h=0. Analogously, for h not equal to zero, one replaces

$\text{Prob}[H_i = h, P_i = 1 | W_i = w]$ with $\text{Prob}[H_i = h, P_i^* = 1 | W_i = w]\ R_t$ and

$\text{Prob}[H_i = h, P_i = 0 | W_i = w]$ with

$\text{Prob}[H_i = h, P_i^* = 0 | W_i = w] + \text{Prob}[H_i = h, P_i^* = 1 | W_i = w][1 - R_t]$.

This form of the probabilities has been simplified by assuming that the probability that someone who is not receiving welfare reports receiving welfare is sufficiently small that it can be set to 0. We also set $P_i = 0$ for those observations reporting welfare receipt despite wage and hours such that welfare income is zero. Lastly, we assume that those with annual income below $3,000 are under-reporting income. We then set other income to the amount which brings their total income up to $3,000 to avoid the estimation being driven by likely mis-measured values.

Samples and Implicit Comparisons Giving Identification

In our base sample of single mothers, the state dummies remove attitudes toward welfare participation and work, since these behaviors may depend on norms which vary across state. We also include year dummies in preferences because attitudes toward work may depend on norms which vary over time. We may consider including year dummies in the expression for the transactions costs and stigma of welfare participation to allow the costs of welfare participation to vary over time. In our single women sample that also includes single women without children, the inclusions of state dummies accounts for attitudes toward work common to both single women with and without children (and also implicitly accounts for state labor market conditions).

We also estimate on some additional subsamples. We use a sample of single women without children, where the EITC plays little role in identifying the tax parameter because the EITC for nonparents is very small. As a result, the tax coefficient is identified in this case through the shape of the tax schedule and state tax differences. Finally, we also use a sample of single mothers with just one child over the years 1988-1990, since tax schedules were constant in real terms over this period. In this sample, the tax coefficient is identified by the shape of tax schedule rather than differences in taxes by the number of children or changes in taxes over time.

Given that, to identify preference parameters, we are using variation in how the EITC affects the after-tax wages of women over time, and of different subsets of women at a point in time, a question worth raising at this point is whether one should be worried about the effects of policies on pre-tax wages in this setting, particularly since there is some research examining the effects of the EITC on pre-tax wages.

Three considerations suggest that this concern may be minor in the present study. First, there is disagreement on the presence and sign of the effect of the EITC on pre-tax wages (Leigh (2004) finds a sizable negative effect, but Rothstein (2005) finds a positive effect, and finds "no evidence that the effect [of the EITC] was to reduce wages.") Second, many of our analyses can be thought of as partial equilibrium analyses with the implicit comparison to single women without children who are in the same labor markets. As long as the groups being compared are in the same labor market but face different tax or benefit rates, then the estimates should not be greatly affected. This is due to the fact that when we include state and year dummies we are netting out tendencies to work that are common to both groups, but vary with state and year. Thus differential treatment of single mothers relative to single women with children over time and across states should provide a valid identification source even if wages change somewhat, as long as the wage changes are similar for single mothers and other single women.[8]
Indeed, Leigh (2004) finds that the effect of the EITC did not differ depending on whether an individual had children. Finally, for the single mother only samples, the effect of the EITC on wages may be more relevant. However, here year dummies are included in the specification so most of the identification is coming through differences in taxes across different numbers of children, and again, Leigh did not find any difference in the effect of the EITC by presence of children.

[8] This assumption is probably more compelling within narrowly defined education groups or when one appropriately controls for education level, especially since Leigh (2004) found differing effects of the EITC by education level.

7. Initial Results

We have estimated simplified versions of the above model. These estimates provide some very striking and intriguing initial results. For example, the source of identifying variation and the features of labor supply on which one focuses have important effects on the estimates and the correction for welfare under-reporting matters. We should emphasize that these results are preliminary and for a simplified version of the model. In particular, these estimates do not allow for unobserved heterogeneity ($\epsilon \equiv 0$), or work costs, nor do we incorporate the value of Medicaid and private health insurance (though we have created these variables). However, based on the small effects that these variables had on other coefficients in difference-in-difference type analyses previously done by one of the authors, we do not expect this last omission to have a substantial effect. In most of the specifications we allow three positive annual hours values (800, 1700 and 2200 hours) and eight hourly wage values (4, 6, 8, 10, 12, 14, 18, and 24 dollars). The assumed wage offer distributions do not vary with individual characteristics, though this will be one of the first simplifications we relax. The models include in the key preference parameter γ controls for the variables that would be in a difference in difference analysis: year and state dummies, as well as demographics.

In Table 1 we report some of these preliminary estimates for single mothers, while in Table 2 we report estimates using the all single women sample. The welfare benefit coefficient α_2 always has the expected sign and the order of magnitude is reasonable, though a bit smaller than expected. The estimates indicate that higher welfare benefits increase welfare participation. There is a substantial stigma or transaction cost of welfare participation ρ, though the magnitude

seems to depend on the estimate of maximum hours μ. Despite small standard errors μ does not seemed to be narrowly pinned down by the data and varies dramatically across specification. As a consequence, in column (4) of the tables we have restricted it to be 4000. The other policy variable coefficients generally accord with expectation. More childcare and training and time limits on welfare increase employment and decrease welfare participation. The coefficients on demographic variables are mostly as expected with older women, whites, those with more education, fewer children, fewer young children all putting a greater weight on income rather than leisure.[9]

The key result that has puzzled us is the coefficient on taxes in the model with choice of hours. Looking at Table 1, column (1), one will see that the coefficient on taxes has the opposite sign from the case where all sources of income are valued the same. However, when one examines just the employment decision by having only one hours choice (see columns (2) and (3) of Table 1) taxes have an even larger coefficient but of the correct sign. In all of these cases the tax coefficient is very precisely estimated so that imprecision of the estimates is not the explanation for this sign change.

These results are not inconsistent with the EITC difference in difference literature much of which has looked at the same data and time period. These studies found that hours increased among single mothers as the EITC was expanded even though marginal tax rates rose for this population. Our result may indicate that single mothers do not understand the effects of the

[9] Some of these results may be due to effects that work through wages rather than preferences, and thus need to be checked after we allow wage distributions which differ by some of these same dimensions.

EITC on marginal tax rates. Further evidence of the how the hours response of single mothers to EITC stands out comes from column (4) of Table 2. Here we estimate using only a sample of single women without children and allow and hours choice as well as the work/nonwork choice. Now the tax rate has the expected sign. The unexpected sign only appears for single mothers and only when we examine hours as well as the work/nonwork decision. This result also suggest that a surprising degree of the identification of the tax coefficient comes from the hours choice as opposed to the work/nonwork decision. The importance of the hours choice is at least somewhat of a surprise given that the standard errors are not that much smaller in the specifications that include the hours choice (compare columns (1) and (2) of Table 2).

We found a number of other interesting results:

- Allowing the mean of tastes for income v. leisure to vary by wage level seems to push the coefficients in sensible directions, for example tax coefficients that are closer to -1 and more similar in magnitude to the welfare coefficient. Compare columns (2) and (3) of Table 1.

- While not reported in the table, allowing for underreporting of welfare receipt tends to make the estimate of stigma or transaction costs lower. This result makes sense since the estimate of transaction costs is in large part driven by women who look like they should participate and yet do not. The number of such women is reduced by the correction for under-reporting which recognizes that many of those not reporting welfare receipt are nevertheless receiving it.

• We obtain fairly similar estimates when we use single women without children as an implicit comparison group for single mothers in the first three columns of Table 2.

• When we rely more on reports of hourly wages when they are available and restrict the sample to rotation groups 4 and 8 so that most women have an hourly wage, there is a much lower probability of the lowest wage category both for workers and the full sample.

• The nonparametric selection correction works well, i.e. the probabilities of the different wage values are estimated precisely. The selected wage distribution of workers generally stochastically dominates the distribution for all women, working and non-working.

8. Conclusions

We expect to have substantive results on the effects of the EITC, marginal tax rates and welfare benefits on work, hours and welfare participation. The estimates will also yield overall labor supply elasticities for single mothers. The puzzling hours response to the EITC seems to be confirmed with a structural approach that is well suited to addressing the question. In the past, this result could be potentially dismissed as coming from difference in difference analyses not obviously suited to analyze the hours question. Our study will also be informative about the sources of identification in structural labor supply models. We also hope that the methods in this paper may find some use in other areas of research where structural models are estimated.

References

Blank, Rebecca M. (2001). "What Causes Public Assistance Caseloads to Grow?" *Journal of Human Resources* 36 (1) 85-118.

Blank, Rebecca M. and Patricia Ruggles (1996). "When Do Women Use Aid to Families with Dependent Children and Food Stamps?" *Journal of Human Resources*, Winter 1996, 31(1), pp. 57-89.

Blundell, Richard, Mike Brewer and Andrew Shephard. (2005). "Evaluating the labour market impact of Working Families' Tax Credit using differences-in-differences." Inland Revenue Working Paper 4.

Blundell, Richard, Alan Duncan, Julian McCrae and Costas Meghir (2000). "The Labour Market Impact of the Working Families Tax Credit." *Fiscal Studies*. 21(1), pp. 75-104.

Bollinger, Christopher R. And Martin H. David (1997). "Modeling Discrete Choice with Response Error: Food Stamp Participation," *Journal of the American Statistical Association* 92, 827-835.

Bound, John, Charles Brown, and Nancy Mathiowetz (2001). "Measurement Error in Survey Data." In *Handbook of Econometrics, Volume 5*, eds. James J. Heckman and Edward Leamer, 3705-3843. Amsterdam: Elsevier.

Brewer, Mike, Alan Duncan, Andrew Shephard, and Maria Jose Suarez. (2005). "Did Working Families Tax Credit work? The final evaluation of the impact of in-work support on parents' labour supply and take-up behaviour in the UK. Mimeo, Institute for Fiscal Studies.

Danziger, Sheldon, Robert Haveman, and Robert Plotnick (1981). "How Income Transfers Affect Work, Savings, and the Income Distribution: A Critical Review." *Journal of Economic Literature*, September 1981, 19(3), pp. 975-1028.

Dickert, Stacy, Scott Houser, and John Karl Scholz (1995). "The Earned Income Tax Credit and Transfer Programs: A Study of Labor Market and Program Participation." In James M. Poterba, ed. *Tax Policy and the Economy* 9. Cambridge, MA: MIT Press, 1995, pp. 1-50.

Eissa, Nada and Jeffrey B. Liebman (1996). "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics*, May 1996, 112(2), pp. 605-637.

Ellwood, David T. "The Impact of the Earned Income Tax Credit and Social Policy Reforms on Work, Marriage and Living Arrangements." *National Tax Journal*, December 2000, 53(4 Part 2) pp. 1063-1105.

Grogger, Jeffrey (2003). "The Effects of Time Limits, the EITC and Other Policy Changes on Welfare Use, Work, and Income Among Female-Headed Families."*Review of Economics and Statistics* 85, 394-408.

Grogger, Jeffrey (2004). "Welfare Transitions in the 1990s: The Economy, Welfare Policy, and the EITC."*Journal of Policy Analysis and Management* 23, 671-698.

Grogger, Jeffrey and Lynn A. Karoly (2005). *Welfare Reform: Effects of a Decade of Change*. Cambridge, MA: Harvard University Press.

Hausman, Jerry. (1981). "Labor Supply." in *How Taxes Affect Economic Behavior*, H. Aaron and J. Pechman. Washington, D.C.: Brookings Institution.

Hotz, V. Joseph and John Karl Scholz (2003). "The Earned Income Tax Credit" in *Means-Tested Transfer Programs in the United States*, edited by Robert A. Moffitt, University of Chicago Press.

Hotz, V. Joseph; Mullin, Charles H. and John Karl Scholz (2005). "The Earned Income Tax Credit and the Labor Market Participation of Families on Welfare." Mimeo, University of Wisconsin.

Heckman, James J. (1993). "What Has Been Learned about Labor Supply in the Past Twenty Years?" *American Economic Review*, May 1993, 83(2), pp.116-121.

Hotz, V. Joseph; Mullin, Charles H. and John Karl Scholz. (2005). "The Earned Income Tax Credit and the Labor Market Participation of Families on Welfare." Mimeo, University of Wisconsin, 2005.

Hoynes, Hilary W. (1996). "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP, *Econometrica*, 64, 295-332.

Keane, Michael and Robert Moffitt (1998). "A Structural Model of Multiple Program Participation and Labor Supply," *International Economic Review*, August 1998, 38, pp. 553-589.

Leigh, Andrew (2004). "Who Benefits from the Earned Income Tax Credit? Incidence Among Recipients, Coworkers and Firms," Working Paper, Australian National University.

Levine, Phillip B. and Diane M. Whitmore (1998). "The Impact of Welfare Reform on the AFDC Caseload." Washington: National Tax Association Proceedings, Ninetieth Annual Conference, 1998.

Liebman, Jeffrey B. (1998). "The Impact of the Earned Income Tax Credit on Incentives and Income Distribution." In *Tax Policy and the Economy 12*, James M. Poterba, ed. Cambridge, MA: MIT Press, 1998, pp. 83-120.

Marquis, Kent H. and Jeffrey C. Moore. 1990. "Measurement Errors in SIPP Program Reports." In Proceedings of the 1990 Annual Research Conference, 721-745. Washington, DC.: U.S. Bureau of the Census.

McFadden, D. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration", *Econometrica*, 57(5), 995-1026.

Meyer, Bruce D. (2002). "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare and Hours Worked" *American Economic Review, Papers and Proceedings*, 92, 373-379.

Meyer, Bruce D. and Rosenbaum, Dan T. (2001). "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics*, August 2001, 116(3), pp. 1063-1114.

Meyer, Bruce D. and Rosenbaum, Dan T. (2000). "Making Single Mothers Work: Recent Changes in Tax and Welfare Policy and their Effects." *National Tax Journal*, December 2000, 53(4 Pt. 2), pp.1027-1062.

Meyer, Bruce D. and Rosenbaum, Dan T. (1999). "Medicaid, Private Health Insurance and Labor Supply," Unpublished Manuscript, July 1999.

Meyer, Bruce D. and James X. Sullivan (2003). "Measuring the Well-Being of the Poor Using Income and Consumption." *Journal of Human Resources*, 38:S, 1180-1220.

Moffitt, Robert (1983). "An Economic Model of Welfare Stigma." *American Economic Review*, December 1983, 75(3), pp. 1023-35.

Moffitt, Robert (1992). "Incentive Effects of the U.S. Welfare System." *Journal of Economic Literature*, March 1992, 30(1), pp. 1-61.

Moore, Jeffrey C., Linda L. Stinson and Edward J. Welniak, Jr. (1997). "Income Measurement Error in Surveys: A Review." Statistical Research Report. Washington D.C.: U.S. Census Bureau.

Rothstein, Jesse (2005). "The Mid-1990s EITC Expansion: Aggregate Labor Supply Effects and Economic Incidence," Working Paper, Princeton University.

van Soest, Arthur. (1993). "Structural Models of Family Labor Supply: A Discrete Choice Approach." Journal of Human Resources. 30, pp. 63-88.

TABLE 1
Single Mothers, Estimates from Various Specifications

	(1)	(2)	(3)	(4)
tax coefficient α_1	1.182 (0.035)	-4.800 (0.031)	-1.466 (0.157)	1.244 (0.153)
welfare coefficient α_2	0.855 (0.018)	0.410 (0.012)	0.307 (0.010)	0.877 (0.054)
maximum hours μ (0079)	4958 (0062)	2626 (0017)	1738	4000
minimum income η (in 1000s of $)	11.3 (0.4)	12.7 (0.4)	17.3 (0.4)	20.4 (1.7)
cost of participation in welfare ρ (in hours)	1027 (0019)	0346 (0016)	0037 (0017)	1085 (0058)
hours choices (annual hours)	800,1700, 2200	1700	1700	800,1700, 2200
wage dummies in income share	no	no	yes	no
additional sample restrictions	none	none	none	1-child 1988-90
sample size	42474	42474	42474	4828

Note: Standard errors are in parentheses. The smoothing parameter τ has been set to 0.1.

TABLE 2
Single Women, Estimates from Various Specifications

	(1)	(2)	(3)	(4)
tax coefficient α_1	-3.870 (0.055)	1.814 (0.043)	1.660 (0.080)	-3.288 (0.033)
welfare coefficient α_2	0.419 (0.010)	0.825 (0.017)	0.911 (0.037)	
maximum hours μ (0130)	8107 (0153)	8914 (0206)	6132	4000
minimum income η (in 1000s of $)	0.25 (0.01)	6.22 (0.16)	14.5 (0.7)	21.1 (1.7)
cost of participation in welfare ρ (in hours)	0893 (0017)	1005 (0022)	0942 (0041)	
hours choices (annual hours)	1700	800,1700, 2200	800,1700, 2200	800,1700, 2200
wage dummies in income share	no	no	no	no
additional sample restrictions	none	none	rotation groups 4,8	no children 1988-96
sample size	117902	117902	29563	51815

Note: Standard errors are in parentheses. The smoothing parameter τ has been set to 0.1.