This work is distributed as a Discussion Paper by the

STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH

SIEPR Discussion Paper No. 06-24

Separating Psychological Costs from Time Costs: Female Labor Supply and Participation in Food Stamps and WIC

By
Colleen Flaherty
and
Kevin Mumford
Stanford University

February 2007

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

Separating Psychological Costs from Time Costs: Female Labor Supply and Participation in Food Stamps and WIC

Colleen Flaherty† and Kevin Mumford‡

February 2007

Abstract

Many individuals who are eligible for welfare choose not to participate. This well-documented fact suggests that there is a utility cost associated with welfare participation. Previous studies have produced estimates of how large this cost would have to be to explain the observed degree of non-participation. Prior estimates of this utility cost have not differentiated psychological costs of participation from the time and effort required to become eligible and maintain eligibility (time costs). This paper develops a structural model that allows for the separate estimation of these two types of costs associated with welfare participation in dollar terms. The estimation suggests that psychological costs are three times larger than the time costs of welfare participation.

†cflare@stanford.edu; Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305.
‡mumford@stanford.edu; Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305.

Acknowledgments: We would like to thank Giacomo De Giorgi, Tom MaCurdy, John Pencavel, Luigi Pistaferri, and Frank Wolak for their help on this project as well as members of the labor reading group at Stanford University. In addition, we would like to thank Sri Nagavarapu for his very helpful comments and feedback.

1 Introduction

A household that is eligible to participate in a welfare program, but decides to forgo participation suggests a violation of a primary tenet of microeconomic theory, that greater consumption gives higher utility. By refusing to accept benefits offered by welfare programs, some households are essentially leaving money on the table. For example, approximately 50 percent of households eligible for Food Stamps do not participate in the program (Trippe and Doyle, 1992).[1] Hence, the decision to turn down welfare benefits by a substantial fraction of eligible households represents a puzzle and also calls into question the effectiveness of government transfer programs in reaching their target population.

Moffitt (1983) is the first to explicitly address this puzzle by introducing "welfare stigma" – the disutility caused from participating in welfare – into an economic model. He tests how welfare stigma enters the utility function and finds evidence for a sizable utility cost from participation. Several subsequent studies of welfare participation have attempted to measure the effect of stigma on participation; typically estimating the effect of observable characteristics, which the researchers argue are associated with welfare stigma, on the probability of participation using a latent index model. Other studies impose a structural model of labor supply and welfare participation, grouping together all the costs associated with participation into an all-encompassing welfare stigma term.

The estimated magnitude of the cost of welfare stigma informs policymakers about household valuation of welfare benefits. However, being able to distinguish what fraction of this utility cost is attributable to the opportunity costs associated with complying with participation requirements relative to psychological costs conveys important additional information that could have policy implications. For example, if the utility costs of participation are primarily due to time costs, such as paperwork and visits to welfare program offices, policies with the goal of increasing take-up rates among eligibles could focus on streamlining the application process. If the utility costs of participation are primarily due to psychological costs, then take-up rates

[1] Trippe and Doyle (1992) use the Current Population Survey (CPS) from 1976 to 1990. While the mean is about 50 percent, there was some variation in participation rates over the period.

could be increased by efforts to reduce the visibility of welfare participation by using refundable credits in the federal income tax code, like the Earned Income Tax Credit (EITC).[2] Surveys in the sociology and public health literature show that participants in welfare programs report lower self-image as well as negative treatment by neighbors, peers, and program administrators (Stuber and Schlesinger, 2006). Separating time costs from these psychological costs in the estimation allows for more informed public policy discussion. This paper develops a structural model that allows for the separate estimation of these two types of costs associated with welfare participation.

Similar to Keane and Moffitt (1998), our paper models participation in two welfare programs: Food Stamp Programs (FSP) and Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). Important to the identification is that we assume that the psychological cost associated with participation does not increase with the number of programs in which the individual participates. Time costs, however, are specific to the program and thus accrue according to the specific programs in which the individual participates. Using a simulated estimation method, we find that, on average, psychological costs are three times larger than the time costs of welfare participation.

The structural model developed in this paper allows for a more accurate characterization of eligibility for these welfare programs. In the model, welfare program participation decisions are made jointly with labor supply decisions. Therefore, most households are potentially eligible to participate in welfare programs; however, actual eligibility depends the labor supply decision.[3] For example, a household with observed earnings greater than the eligibility cutoff could have received benefits by choosing to earn less. This model seeks to explain not only why eligible households choose not to participate, but also why other households choose to earn more than the eligibility cutoff and thus preclude welfare participation.

[2]This assumes that psychological costs are significantly lower for income received through the tax system than through welfare programs. This is sensible due to the low visibility and the widespread usage of tax credits and deductions, reducing negative peer pressure associated with receiving distributions through the tax system. Hotz and Scholz (2003) estimate that EITC participation among eligibles in 1996 was no greater than 87.2 percent. This is substantially higher than participation in FSP and WIC.

[3]Eligibility for WIC depends primarily on the presence of children in the household. This is a static model, and children are taken as exogenous. The eligibility requirements are explained in Section 2.

Section 2 outlines the benefits and eligibility rules of FSP and WIC, the two programs used in this study. The economic model is defined in Section 3, while Section 4 gives the econometric and functional form specifications. The method of estimation is discussed in Section 5 and the data used in the estimation is described in Section 6. Results from the simulated estimation are given and discussed in Section 7. Section 8 presents a model in which welfare stigma could act as a screening mechanism and discusses the potential for social welfare implications. Section 9 concludes and outlines areas for future research.

2 Program Characteristics and Eligibility

The Food Stamp Program (FSP) and Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) are two of the federal food and nutrition programs which are federally financed and uniform across states. In contrast to AFDC and TANF, most rules for FSP and WIC are set at the federal level.[4]

The national eligibility standards for FSP were established in 1971. Food stamp recipients are clustered into "food stamp units," which can consist of unrelated individuals if they buy and prepare food together. Parents and children must apply together, except if the children are parents themselves. Food stamp units qualify for the program if their gross and net incomes are lower than specified thresholds. In addition, food stamp units must meet an asset limit of $2,000 for households that do not include an elderly individual and $3,000 if an elderly person is present. Food stamp units that do not contain an elderly (or disabled) individual must satisfy two income tests: 1) gross income test, which states that income cannot exceed 130 percent of the poverty threshold for that family size; and 2) net income test, which requires that gross income less deductions for work expenses, child care, and costs of housing, cannot exceed the poverty threshold. Food stamp units that include an elderly (or disabled) individual only have to satisfy the net income test. The maximum benefit level is defined by the household composition and benefits are reduced at a rate of 30 percent for each additional dollar of net income (including

[4]Temporary Aid for Needy Families (TANF) replaced Aid to Families with Dependent Children (AFCD) with the passage of Welfare Reform in 1996.

Table 1: FSP Participation Rates

Year	Percent of Total Population	Percent of Poor Population
1975	7.6	63.0
1980	8.4	65.6
1985	8.3	60.2
1990	8.0	59.2
1995	10.1	73.0
1997	8.5	64.3
2002	6.7	55.2

Source: Greenbook, 2004.
Poor population defined by the poverty threshold

transfers from AFDC or TANF).

Historically, FSP distributed coupons which could be used to purchase any food item at participating stores, excluding alcohol, tobacco, and some prepared foods. In 1993, Maryland instituted an electronic benefit transfer (EBT) system to modernize the process. A mandate was passed in 1996 which required all states to adopt EBT by 2002. The adoption of EBT was slow; by 2000, only twenty states had initiated pilot programs. This paper analyzes participation in the fall of 1997, which is well before the full adoption of EBT. Future work will compare estimates of psychological costs and time costs before and after the adoption of the electronic system.

As referenced in Section 1, FSP participation rates are surprisingly low. The third column in Table 1 shows the percent of households below the poverty threshold that participate in FSP over time. Because families with gross income less than the poverty threshold meet the FSP income eligibility requirements, these participation rates show that a sizable fraction of eligible families choose not to participate. Households that meet the net income test can be eligible for FSP even with gross income greater than the poverty threshold. If these households have a lower participation rate than those at or below the poverty threshold then the participation rates reported in Table 1 are reconcilable with the 50 percent participation rate for eligible individuals reported by Trippe and Doyle (1992) using the CPS. Eligible families with income greater than the poverty threshold qualify for smaller FSP benefit amounts which, for many families, may not be large enough to overcome the time and psychological costs associated with participation.

WIC was established in 1972 as a program to provide nutritional support to women that are pregnant or breast-feeding and to children under age 5. WIC provides paper coupons that specify exactly what and how much food can be purchased. These food items include infant formula, juice, milk, cereal, and protein-rich foods (such as peanut butter and beans). A family is eligible for WIC benefits if its income is less than 185 percent of the federal poverty level. The program stipulates that individuals need to be at risk in terms of nutritional status. However, in practice women and children who meet the income requirement are eligible for WIC benefits because nutritional risk is so broadly defined (Currie, 2003).[5]

Currie (2003) reports that in 1998 approximately 60 percent of eligible families participated. However, Currie finds that the participation rates vary by child's age: the take-up rate for families with an infant (a child under age one) is 73 percent, but drops to 38 percent for families with children between ages one and five. We find very similar results, both in the overall participation rates and in participation by child's age. These statistics are presented in Section 6. Higher participation rates for families with infants is not surprising because the value of WIC benefits for these families is greater than those with only toddlers.[6] Lower WIC participation for families with children between ages one and five compared to families with infants suggests that non-participation is not primarily due to lack of information about the program. Instead, this indicates that families choose to discontinue WIC participation due to the drop in the value of benefits.

3 Model

This paper builds off the work of Moffitt (1983) and Keane and Moffitt (1998) by developing a structural model with an explicit utility cost of welfare participation. Using data on single

[5]Women and children with low-income are classified as being nutritionally at risk.

[6]The difference in the value of WIC benefits is due to the infant formula which WIC provides to families with infants. While, children between ages one and five receive cereal, milk, cheese, and other food items, their combined value is approximately one third the value of the infant formula.

females in the Panel Study on Income Dynamics (PSID), Moffitt (1983) tests whether there is a utility cost associated with participation in AFDC. His specification allows welfare stigma to enter both as a flat cost and a variable cost. A flat cost implies a threshold, given by the level of stigma, which benefits must exceed if the individual is to participate. A variable cost implies that the value of income received from welfare programs is less than that from private income. Empirically, a flat cost implies that participation rates would increase if welfare benefits were to become more generous; a variable cost alone would not have this implication.

Moffitt (1983) finds evidence that stigma enters as a flat cost. Because his estimate of the variable cost component would imply that welfare income is more valuable than private income, he concludes that the variable cost component is unimportant and suggests that stigma enters only as a flat utility cost. This is supported by Smeeting's (1982) observation that food stamp benefits are valued similarly to cash transfers because the average dollar amount of benefits does not typically exceed the family's total food expenditure. The same argument could be extended to WIC benefits because the monetary value of the benefit is relatively low and covers food staples of infants and children (e.g. formula, juice, cereal). Based on Moffitt's finding, our model is designed to only allow for a flat utility cost of welfare participation.

We present a static model of labor supply and welfare program participation in a utility maximizing framework. The individual jointly decides how many hours to work (or to not work at all) and whether or not to participate in multiple welfare programs. Individual i's utility is given by

$$U_i = U\Big(L_i\,,\,C_i\Big). \tag{1}$$

The individual's utility depends on leisure, L_i, and effective consumption, C_i. Effective consumption is the sum of after-tax income (labor and non-labor) and welfare benefits net of the psychological cost of welfare participation. Leisure is time remaining after accounting for hours of work and the time cost of participation in welfare programs. As documented in Section 2, not all eligible individuals participate in welfare. The costs of welfare participation enter the utility function through both the leisure and income channels. The time cost includes time spent traveling to and from the welfare office, filling out forms, and waiting in lines. The psychological cost from participating is the "stigma" of being on welfare, such as disapproval from peers or neighbors. We assume that the psychological cost, ϕ_i, is the same across programs and does not increase in the number of programs. Hence, ϕ_i is the psychological cost from participating in only program 1 or program 2, or both. Keane and Moffit's (1998) result that costs associated with participation are low for additional programs suggests that this structural assumption is plausible. In addition, the assumption is plausible because FSP and WIC are similar in that receiving the benefits requires a transaction at a grocery store involving coupons.

Keane and Moffitt (1998) are the first to model multiple program participation in a structural framework.[7] They attempt to determine whether stigma costs of participation are perfectly additive in the number of programs. They specify a program-specific utility cost of participation, ϕ_A for AFDC and ϕ_F for FSP. The agent's utility is given by:

$$U(leisure, income) - \lambda(\phi_A P_A + \phi_F P_F) + (1-\lambda)\max(\phi_A P_A, \phi_F P_F)$$

where P_A and P_F indicate participation in the welfare programs and ϕ_A, ϕ_F, and λ are parameters to estimate. The value of λ is restricted such that $\lambda \in [0, 1]$. A value of λ close to 1 means that utility costs are perfectly additive, while a λ close to 0 implies that utility costs of participation are zero for additional programs. The estimation reports that $\lambda = 0.05$, which is consistent with a large psychological cost from participating in the first program and only a trivial psychological cost, or purely a time cost, associated with participating in an additional program.

Returning to our model, by substituting for L_i and C_i into equation 1, we get the following utility function:

$$U\Big(TE_i - H_i - P_{1i}\delta_1 - P_{2i}\delta_2\,,\; Y_i + P_{1i}\,(B_{1i} - \phi_i) + P_{2i}\,(B_{2i} - \phi_i) + P_{1i}P_{2i}\phi_i - \mathbb{1}(H_i > 0)FCw_i\Big) \quad (2)$$

[7] Their intent was to model participation in three programs: AFDC, FSP, and subsidized housing, to determine the disincentives created by the interaction of benefit schedules from multiple programs. However, subsidized housing is different from the other two welfare programs in that it is rationed, meaning that even if the individual would optimally choose to participate, she would not necessary be able to due to a queue. Because of this problem, Keane and Moffitt restrict their analysis to participation in AFDC and FSP.

In the utility function, TE_i is the time endowment and H_i is the hours worked. P_{1i} is a binary indicator of the individual's decision to participate in FSP and P_{2i} indicates the individual's decision to participate in WIC. Time costs for each program are given by δ_1 and δ_2. After-tax income is given by Y_i, where $Y_i = w_i H_i + N_i - \tau_i(w_i H_i + N_i)$, the sum of the individual's wage and non-wage income less taxes. The tax function, τ_i, gives the federal income tax liability and depends on i's family characteristics and includes deductions, exemptions, and the EITC. We ignore all state and local taxes. The welfare benefit level that the individual would receive from participating in FSP and WIC are given by B_{1i} and B_{2i}. Workers are assumed to face a fixed cost of working, FC, which is assumed to be proportional to the individual's wage. Introducing a fixed cost of working into the model increases the complexity only slightly but helps the model account for the fact that very few individuals decide to work for less than 5 hours per week.

FSP benefits, B_{1i}, are given by

$$B_{1i} = \bar{B}_{1i} - 0.3\,(0.8\,w_i\,H_i + N_i - \$125\text{ Children})\,. \tag{3}$$

The maximum benefit level, $\bar{B}_{1i}$, depends on the number of persons in the family. FSP benefits are reduced at a rate of 30 percent with net income. Net income is equal to total income less a 20 percent deduction for earned income and a \$125 deduction for each child under five years of age. This formula is a close approximation to the actual FSP benefit formula in which several deductions are allowed from gross income, including a 20 percent deduction for earned income and a deduction for childcare expenditures. We approximate the childcare deduction with a \$125 fixed amount for young children because we do not observe actual childcare expenditures. In addition, we do not include deductions for "excess housing cost" expenses, which will under-predict the level of benefits for those who would have received a housing deduction. We err in this direction to avoid predicting benefits that are greater than actual eligibility.

Consistent with FSP eligibility requirements, B_{1i} is set equal to zero if individuals fail either the gross or net income test. The maximum benefit level, B_{1i}, is set equal to zero for individuals with liquid assets in excess of \$5000 (not including retirement accounts). Assets are defined as liquid if they are held in checking or interest-earning accounts. Assets held in stocks or bonds are not subject to this asset limit because, if these assets are held in pension accounts, they would not be counted against the asset limits by the Food Stamp office. We select an asset cutoff above the FSP level of \$2000 (or \$3000 for families with an elderly individual) because families could opt to spend down or hide assets to meet the asset threshold.[8]

For eligible families, WIC benefits do not decrease in income. Benefits depend on the age and number of children, as well as on whether or not the woman is pregnant.

$$B_{2i} = \begin{cases} 0 & \text{if } i \text{ has no children} < \text{age 5 and is not pregnant} \\ \bar{B}_{2i} & \text{if } w_i H_i + N_i \leq 1.85(\text{poverty}_i) \text{ and } \{\text{children} < \text{age 5 or is pregnant}\} \end{cases} \tag{4}$$

where $\bar{B}_{2i}$ is the dollar value of the food items qualified for based on family characteristics. Benefits are equal to zero if there are no children under age five and the woman is not pregnant or if income exceeds 185 percent of the poverty threshold for that family size.

To solve the individual's problem we substitute equations (3), and (4) into equation (2). The strategy is to then solve for the optimal level of work, H_i^*, for each program-participation combination. For example, assume that the agent participates in both welfare programs and works positive hours. For this case we maximize (2) subject to eligibility constraints for both programs. This gives the optimal value for H_i assuming that the individual is participating in both welfare programs and working positive hours. Substituting the optimal value of H_i into the utility function gives the indirect utility, V_i. The individual's welfare participation decision is determined by comparing V_i for each of the participation combinations.

The incentives created by welfare programs may influence family structure itself; however, studies find that the estimated impact is small in magnitude (Moffitt, 1992). We assume that marital status, number of children, and living arrangement are exogenous and do not depend on benefits levels.

[8]Keane and Moffitt (2002) use a similar FSP asset test limit of \$4,500.

4 Econometric and Functional Form Specification

Several reduced-form analyzes of welfare participation provide insight into which factors are associated welfare stigma. Blundell, Fry and Walker (1988) test for the presence of stigma, or non-monetary costs of participation, using the Standard Housing Benefit in the U.K. They find a positive relationship between benefits levels and take-up rates which is consistent with the presence of stigma in an economic model. They emphasize that public policies which increase the benefit level need to consider the impact on revenues due to increased participation, not just the impact due to higher benefit amounts. Blundell, Fry and Walker find that education and the age of children in the household affect the probability of participation, which suggests that these factors are correlated with welfare stigma because they do not directly determine benefits levels.

Ripahn (2001) also attempts to incorporate stigma in her reduced-form analysis of participation in the German social insurance program. She finds that participation rates are higher for single-parents, for parents with children under the age of 7, and for those living in cities with higher poverty levels. She interprets these findings as indicating that stigma is lower for families with these characteristics. Similar to Blundell, Fry, and Walker (1988), she finds that the probability of participation decreases in education attainment. Our analysis incorporates some of these aforementioned characteristics in the estimation of psychological costs.

The psychological cost incured by an individual from participating in either or both welfare programs, ϕ_i, is given by

$$\phi_i = X_i\,\alpha + \eta_i \tag{5}$$

where X_i is a vector of observed characteristics for individual i and η_i is a normally distributed error term that accounts for heterogeneity in psychological costs across individuals. We assume that the wage for individual i is given by

$$w_i = Z_i\,\beta + \epsilon_i \tag{6}$$

where Z_i is a vector of observed characteristics and includes some of the same elements as X_i.

The error term, ϵ_i, is normally distributed and accounts for heterogeneity in earning ability. The stochastic process differs from that used by Keane and Moffitt (1998); they attach an error term to each participation equation, the hours equation, and the wage equation, while this study only includes two sources of heterogeneity: the wage equation and stigma equation.

The utility function used in this analysis is given by:

$$U = (L_i)(C_i)^{\gamma_1} - \gamma_2 (L_i)^2 \tag{7}$$

This function is not commonly used in economic applications because it does not give closed form analytical expressions for the demand functions. Cobb-Douglas or CES functions are used more often, but have serious limitations, including a constant budget share for each good regardless of income level or relative prices. The utility function given above is more flexible than standard utility function and is still quite simple in form. It has the property that the marginal utility of leisure depends on the level of consumption. This is desirable because it allows leisure time to have a high marginal value for those with high consumption and a low value for those with low consumption. In fact, at very low levels of consumption, it is possible for an individual to actually be satiated in leisure. Similarly, individuals with low levels of leisure have a lower marginal utility of consumption.

We estimate the utility function parameters, γ_1 and γ_2, as well as the fixed cost of working and time endowment parameters. The parameters of the wage equation, β, and σ_ϵ (the standard deviation of the mean-zero normally distributed error term, ϵ) are also estimated. The primary focus of the analysis is estimating the time cost parameters for each program, δ_1 and δ_2, and parameters of the psychological cost equation, α and σ_η (the standard deviation of the mean-zero normally distributed error term, η).

Identification of α, δ_1 and δ_2 is obtained both through exclusion restrictions (different elements in X_i and Z_i) and the assumed functional form. The assumptions that psychological cost does not increase in the number of programs and that time cost of participation is program specific are the primary structural features that provide identification along with the assumption of how these two costs separately enter the utility function through consumption and leisure.

In writing down the likelihood function, there are eight possible program participation/labor supply combinations. The likelihood for individual i is given by the following:

$$\begin{aligned}
\ell_i = \ & \Pr\left[H_i = 0, P_{1i} = 0, P_{2i} = 0 | X_i\right]^{(1-P_{1i})(1-P_{2i})\,\mathbb{1}(H_i=0)} \cdot \Pr\left[H_i = 0, P_{1i} = 1, P_{2i} = 0 | X_i\right]^{(P_{1i})(1-P_{2i})\,\mathbb{1}(H_i=0)} \\
& \cdot \Pr\left[H_i = 0, P_{1i} = 0, P_{2i} = 1 | X_i\right]^{(1-P_{1i})(P_{2i})\,\mathbb{1}(H_i=0)} \cdot \Pr\left[H_i = 0, P_{1i} = 1, P_{2i} = 1 | X_i\right]^{(P_{1i})(P_{2i})\,\mathbb{1}(H_i=0)} \\
& \cdot \Pr\left[H_i > 0, P_{1i} = 0, P_{2i} = 0 | X_i\right]^{(1-P_{1i})(1-P_{2i})\,\mathbb{1}(H_i>0)} \ f(H_i, w_i | X_i, H_i > 0, P_{1i} = 0, P_{2i} = 0) \\
& \cdot \Pr\left[H_i > 0, P_{1i} = 1, P_{2i} = 0 | X_i\right]^{(P_{1i})(1-P_{2i})\,\mathbb{1}(H_i>0)} \ f(H_i, w_i | X_i, H_i > 0, P_{1i} = 1, P_{2i} = 0) \\
& \cdot \Pr\left[H_i > 0, P_{1i} = 0, P_{2i} = 1 | X_i\right]^{(1-P_{1i})(P_{2i})\,\mathbb{1}(H_i>0)} \ f(H_i, w_i | X_i, H_i > 0, P_{1i} = 0, P_{2i} = 1) \\
& \cdot \Pr\left[H_i > 0, P_{1i} = 1, P_{2i} = 1 | X_i\right]^{(P_{1i})(P_{2i})\,\mathbb{1}(H_i>0)} \ f(H_i, w_i | X_i, H_i > 0, P_{1i} = 1, P_{2i} = 1).
\end{aligned}$$

The likelihood function is given by:

$$L = \prod_{i=1}^{N} \ell_i \tag{8}$$

When $H_i > 0$, the individual's hours and wage are observable; a joint conditional density captures that information. Let $f(H_i, w_i | X_i, H_i > 0, P_{1i}, P_{2i})$ be the joint conditional density of wages and hours for positive hours. The set of parameters, θ, in the above likelihood function is suppressed for ease of notation.

5 Estimation

The individual's budget set is non-convex and complicated because we use realistic tax functions, FSP benefit functions, and WIC eligibility cutoffs. This makes it difficult to derive a closed-form labor supply function or to use stepwise-linear techniques. Instead, we compartmentalize hours of work into 16 discrete bins. The bin is denoted by h_i.[9] The agent then has two program participation choices and chooses one of the 16 hours options, making for a discrete choice problem with 64 choice bundles. This treatment of hours of work implies a slightly different

[9]The bin values are: {0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100}, which correspond to h_i = {1,2,...,16}. Observed hours are assigned to each bin by creating a range between bins that spans half the distance to the next bin. This procedure is quite common in estimating structural models, for example, Keane and Moffitt (1998) consider 3 hours choices: 0, 20, 40.

log-likelihood function given by:

$$\ln L = \sum_{i=1}^{n} \ln \ell_i. \tag{9}$$

where the log-likelihood for individual i is given by:

$$\begin{aligned}
\ln \ell_i = {} & (1-P_{1i})(1-P_{2i}) \ln\Big(\Pr\left[h_i = 1, P_{1i} = 0, P_{2i} = 0 | X_i, \theta\right]\Big) \\
& + (P_{1i})(1-P_{2i}) \ln\Big(\Pr\left[h_i = 1, P_{1i} = 1, P_{2i} = 0 | X_i, \theta\right]\Big) \\
& + (1-P_{1i})(P_{2i}) \ln\Big(\Pr\left[h_i = 1, P_{1i} = 0, P_{2i} = 1 | X_i, \theta\right]\Big) \\
& + (P_{1i})(P_{2i}) \ln\Big(\Pr\left[h_i = 1, P_{1i} = 1, P_{2i} = 1 | X_i, \theta\right]\Big) \\
+ \sum_{k=2}^{16} \Bigg(& (1-P_{1i})(1-P_{2i}) \ln\Big(\Pr\left[h_i = k, P_{1i} = 0, P_{2i} = 0 | X_i, \theta\right] \; f(w_i | X_i, \theta, P_{1i} = 0, P_{2i} = 0)\Big) \\
& + (P_{1i})(1-P_{2i}) \ln\Big(\Pr\left[h_i = k, P_{1i} = 1, P_{2i} = 0 | X_i, \theta\right] \; f(w_i | X_i, \theta, P_{1i} = 1, P_{2i} = 0)\Big) \\
& + (1-P_{1i})(P_{2i}) \ln\Big(\Pr\left[h_i = k, P_{1i} = 0, P_{2i} = 1 | X_i, \theta\right] \; f(w_i | X_i, \theta, P_{1i} = 0, P_{2i} = 1)\Big) \\
& + (P_{1i})(P_{2i}) \ln\Big(\Pr\left[h_i = k, P_{1i} = 1, P_{2i} = 1 | X_i, \theta\right] \; f(w_i | X_i, \theta, P_{1i} = 1, P_{2i} = 1)\Big)\Bigg)
\end{aligned}$$

where $k \in \{1, 2, \ldots, 16\}$ represents the hours of work choices $\{0, 5, \ldots, 100\}$.

Simulation methods can be used to estimate otherwise intractable structural models. Because they are difficult to derive in this framework, the probabilities and conditional wage densities in the log-likelihood equation above are computed using simulated methods. A large number of draws (D total draws) for the error terms in the psychological cost and wage equations, η and ϵ, are each taken from a standard normal distribution. Given the vector of parameter values, θ, the error terms are scaled by σ_η and σ_ϵ respectively. The simulated probability $\Pr_S\left[h_i, P_{1i}, P_{2i}\right]$ is given by:

$$\Pr_S\left[h_i, P_{1i}, P_{2i}\right] = \frac{1}{D} \sum_{d=1}^{D} \mathbb{1}\left(h_{id} = h_i, P_{1id} = P_{1i}, P_{2id} = P_{2i}\right) \tag{10}$$

where d indicates a simulation draw for η and ϵ. The conditional density of wages, $f(w_i|X_i, \theta_i, P_{1i}, P_{2i})$, is estimated using a simple nonparametric density estimator.[10] The log-likelihood is evaluated

[10] We use a simple histogram nonparametric density estimator where $f(w_i) = \frac{1}{D\,v} \sum_{d=1}^{D} \mathbb{1}\left(-\frac{1}{2} \le \frac{w_{id} - w_i}{v} \le \frac{1}{2}\right)$. This estimator simply constructs a histogram that is based on all w_{id} within $\frac{v}{2}$ of w_i, where v is the width of this histogram interval. A drawback of the simple histogram estimator is that predicted wage values w_{id} that lie outside of the $\frac{v}{2}$ range contribute nothing to the density estimation. We alter this simple method by allowing

given a vector of parameter values, θ, and then a simplex method is used to update θ in order to improve the log-likelihood value. A simplex method is used rather than standard quasi-Newton or conjugate gradient methods because the non-convexity of the budget set makes these standard methods less reliable than simplex methods.

The results presented in Section 7 were computed using $D = 2000$ simulation draws.[11] Various initial values for θ were tried and each eventually converged to the reported parameter estimates, although this does not guarantee that a global maximum was found. However, the robustness of the parameter estimates to different initial parameter values and the fact that the estimates are economically sensible suggest that the estimation procedure is reliable.

The simulated log-likelihood parameter estimates are asymptotically unbiased as the number of simulation draws grows large. The standard errors are computed as the inverse of the outer-product of the simulated scores. This procedure requires calculating the matrix of contribution to the gradient, $G(\theta)$, but does not require computation of the full Hessian. Calculating the Hessian is computationally difficult because the derivatives of the likelihood function must be found numerically.[12] The matrix of contribution to the gradient is a N x J matrix where N is the number of observations and J is the dimension of the vector of parameters, θ. The elements of $G(\theta)$ are given by:

$$G_{ij}(\theta) = \frac{\partial \ln \ell_i(\theta)}{\partial \theta_j}. \tag{11}$$

The variance-covariance matrix is computed as the inverse of the outer-product of $G(\hat{\theta})$:

$$V(\hat{\theta}) = \left[G'(\hat{\theta})G(\hat{\theta})\right]^{-1}. \tag{12}$$

predicted wages outside of the bin to contribute to the density in proportion to the percent deviation as given by $e^{(-\frac{1}{v}(\frac{w_i - w_{id}}{w_{id}})^2)}$.

[11] With each restart of the estimation procedure, both the tolerance level and the number of simulation draws D were increased.

[12] The Hessian matrix is often computed as part of the estimation procedure. However, this is not the case when only a simplex method is used. Because the simplex method does not rely on derivatives of the log-likelihood function, these derivatives must be computed numerically once the estimation procedure is completed.

6 Data

The data used for the estimation is a sample of female household heads from the Survey of Income and Program Participation (SIPP). Our sample consists of non-married females over 18 years of age who are in households where there is a clear decision-maker. Households with multiple agents of working age were eliminated to alleviate concerns about joint labor supply and welfare participation decisions within a household. For example, a household with two females of working age would not be included in our sample. In addition, any female-headed household with a male over age 18 who has income greater than the female head is dropped from the sample. Households in which an adult female is living with her parent or an older working-age adult are also excluded. For each household in the selected sample, the female household head is the clear decision-maker.

The selected sample is only representative of households with a female household head and no other working-age adults. However, this group represents a large fraction of welfare participants. In 1997, 60 percent of households that participated in food stamp and 40 percent of households that participated in WIC had an unmarried female household head. Our sample imposes the additional restriction that there are no other potential decision-makers in the household. In 1997, 44 percent of households that participated in food stamp and 28 percent of households that participated in WIC met this additional restriction. While the selected sample does not represent the full welfare-eligible population, it does represent a substantial part of that population.

The household decision-maker is the unit of analysis. Within our sample, determining who is the head of household is usually straightforward because of the sample selection criteria listed above. For more ambiguous family arrangements, the assignment of household head status is based on earned income, age, whether the women is a mother, and who owns the welfare benefits (when applicable). In total, we have 6,270 households. We only include households consisting of individuals or families; we did not allow for unrelated secondary individuals or subfamilies (as classified in SIPP). Because we limit our sample to households with a single decision-maker and do not include households with unrelated individuals, our households closely correspond to a food stamp unit.

The Welfare Reform of 1996 presents a problem for imposing universal benefits and eligibility rules because it introduces a large amount of state discretion in the determination of program specifics and time limits on receipt of welfare. To mitigate this problem, we choose to analyze participation choices in FSP and WIC, which were less affected by the reform compared to TANF (formerly AFDC). Additionally, we choose to analyze data from the fall of 1997, which was before the transition to state-determined welfare was complete to limit confusion regarding time limits by the eligible population. The family composition was defined as of September 1997, with pregnancy being imputed using later waves of the SIPP. Participation in FSP and WIC was taken from two months, September and October, to allow for a longer time window to observe participation. This means that a family is considered a participant if any member participated in FSP or WIC during either of these two months. Earnings and hours data were averaged over four months, July through October, in order to smooth over shocks and give a more accurate measure of labor supply. For each individual, if the hourly wage implied by observed earnings and labor supply was below minimum wage, the individual's wage was set equal to the federal minimum wage of $4.75.

Unlike FSP, WIC benefits are specified in ounces of food, not as a dollar value. For this analysis, we convert the food items into dollar amounts using inflation-adjusted prices of these goods. The food items covered by WIC depend on family characteristics, hence the value of benefits depends on the family's composition. Table 2 shows the value for mothers by age of child in 1997 dollars. Prices were computed using 2006 prices per ounce of food product and deflated using the CPI-U. Prices per ounce were selected from large-size packages to use the lowest available price to err on the side of undervaluing the benefits to avoid overestimating the role of psychological and time costs in the participation decision.

The descriptive statistics for the sample are given in Table 3 and were computed using sample weights. The average age of female head of households is 43 and the youngest female is 18 years old. Roughly one-third of these women are of minority status (black or Hispanic). Over 30 percent have a post-secondary degree and 29 percent have only a high school diploma. Most of these women live in an urban area and roughly one-third live in Southern states. Nearly 20

Table 2: Value of WIC Benefits

Family Member	Monthly Value of Food Items ($1997)
Infant: 0 to 3 months	$97.66
Infant: 4 to 12 months	$105.41
Child: 1 to 5 years	$31.26
Mother: Pregnant or Breast-feeding	$33.59

Sources: Food items from www.fns.usda.wic/benefitsandservices/foodpkgtable.htm
Prices: www.giantfood.com and prices deflated using CPI-U: www.bls.ogv/cpi

percent have children under the age of 18 living with them and 15 percent have teenagers living with them. Only 13 percent have a child young enough to meet the eligibility requirement of WIC (under age 5).

Non-labor income includes earnings from assets, royalties, and property, as well as earned income from other members in the family and government transfers; most of these women have positive non-labor income. While the majority of women have positive assets, the distribution is skewed to the right. Three-fourths of the women had positive weekly hours at some point over the four month window (July 1997 to October 1997) and the average weekly hours was just over 30.[13] Average weekly wage was computed using the reported work history from the four-month window. However, final wages that were below minimum wage were set to $4.75.[14]

Table 4 displays the participation rates and benefit values for FSP and WIC. Not controlling for eligibility, roughly 6 percent of the sample participates in WIC and 18 percent participate in FSP; 4 percent participate in both programs. Of those females who meet the WIC requirement based on ages of children (under age 5), 41 percent participate in WIC and 29 percent participate in both FSP and WIC. For WIC, participation rates by child's age allow for a comparison to the rates reported by Currie (2003). Table 4 reports that participation rates in WIC are highest for households with an infant (67 percent) and drop substantially for those with children between ages one and five (35 percent); these numbers correspond closely to Currie's finding of 73 percent

[13] Those women who report working more than 100 hours per week were top-coded to 100.
[14] The minimum wage floor was imposed on less than 5 percent of the sample.

Table 3: Descriptive Statistics

Demographic Characteristics	Mean	St. Error	Minimum	Maximum
Age	43.0	14.2	18	85
Black	23.28%	42.26%	0	1
Hispanic	7.39%	26.16%	0	1
Master's Degree or higher	7.46%	26.27%	0	1
Bachelor's Degree	14.58%	35.29%	0	1
Associate's Degree	11.51%	31.92%	0	1
Some College	21.01%	40.74%	0	1
High School Graduate	28.31%	45.05%	0	1
High School Dropout	9.00%	28.63%	0	1
Junior High Dropout	8.13%	27.33%	0	1
Live in City	83.05%	37.52%	0	1
South	34.68%	47.60%	0	1
Family Size	1.8	1.2	1	13
Children in Family (*under age 18*)	19.08%	39.30%	0	1
Number of Children (*under age 18*)	0.7	1.1	0	10
Infant in household (*0 to 12 months*)	2.24%	14.79%	0	1
Child under age 5 (*WIC eligible*)	12.16%	32.68%	0	1
Teen in Household (*ages 13 to 17*)	15.19%	35.89%	0	1

Labor Force Participation and Income	Mean	St. Error	Minimum	Maximum
Non-labor Income (*weekly*)	$144	$270	$0	$11,258
Positive Non-Labor Income	83.33%	37.27%	0	1
Assets (*median = $232*)	$11,155	$62,670	$0	$1,876,205
Positive Assets	61.58%	48.64%	0	1
Weekly Hours of Work	30.4	21.5	0.0	100
Positive Hours	74.92%	43.35%	0	1
Hourly Wage (*minimum wage imposed*)	$11.70	$8.55	$4.75	$259.04

Table 4: Welfare Participation and Benefits

Program Participation	Mean	St. Error	Minimum	Maximum
WIC	5.58%	22.94%	0	1
FSP	17.77%	38.23%	0	1
WIC and FSP	3.81%	19.14%	0	1
WIC (with infant)	67.15%	47.14%	0	1
FSP and WIC (with infant)	47.48%	50.11%	0	1
WIC (with Child age 1 to 5)	35.11%	47.73%	0	1
FSP and WIC (with Child age 1 to 5)	24.30%	42.89%	0	1
WIC (with Child under age 5)	41.01%	49.18%	0	1
FSP and WIC (with Child under age 5)	28.57%	45.17%	0	1

Monthly Benefit	Mean	St. Error	Minimum	Maximum
Maximum FSP Benefits	$202.89	$107.53	$121	$1,180
Value of WIC Benefits (Child < 5 years old)	$54.56	$37.46	$31.26	$242.08

and 38 percent, respectively.

Because of the panel nature of the SIPP, it is possible to observe subsequent WIC participation decisions of households that have an infant in 1997. These results are reported in Table 5. Of those households with an infant that participated in WIC in 1997, only 50 percent continued to participate during 1998 (when the child was age 1), and only 44 percent continued to participate during 1999 (when the child was age 2). As shown in Table 2, the value of benefits for an infant are three times larger than for a child between ages one and five. For the 67 percent of households with an infant (not controlling for income) that chose to participate in WIC while the child is under age one, the value of these benefits outweigh the costs of participation. However, only half of those families continued to participate during the next year when the child is one and benefits are much lower, indicating that the costs of participation outweigh the benefits for non-participating families. These results show that the drop in participation in WIC by these households is not due to a lack of information about the program.

In addition, nearly 70 percent of households in the sample with an infant participate in WIC, regardless of income. Given this widespread participation, it is likely that most non-participating mothers have come in contact with program users. While a lack of information could prevent

Table 5: Participation in WIC by Families with Infants in 1997
(Conditional on participation in 1997)

Year	Obs.	Mean	St.Dev
1997	88	1	0
1998	88	0.5016	0.5028
1999	88	0.4380	0.4989

calculated using SIPP weights

some households from participating in welfare programs, these results for WIC indicate that a lack of information is not a major cause of non-participation.

Returning to Table 4, the last two rows of the table under the monthly benefit heading report summary statistics for the maximum welfare benefits. Maximum monthly benefits for FSP were computed using family size and are equal to the value of benefits at zero dollars of net income. The value of WIC benefits was computed based on the price of the bundle of goods covered for each family member (see Table 2). This maximum benefit value, and not the observed level of benefits, is relevant to the model because it gives the information necessary to determine what the benefit level would be for any potential labor supply.

Before we estimate the complete model, we estimate simple wage and participation equations separately as a baseline for comparison. Table 6 shows the estimates from the wage equation using ordinary least squares (OLS) with hourly wage as the dependent variable (for those women with positive hours of work). The estimates correspond to those typically found in the literature: wage is concave in age, increasing in education, higher for women who live in urban areas, and lower for women with children. Wages are also lower for women identifying themselves as black or Hispanic relative to white (excluded group) and for those living in Southern states.

The decision to participate in welfare can be examined using a latent index model. Abstracting from eligibility constraints, an individual participates if her utility from participation is greater than her utility from not participating. Let the latent index, $P^*_{i,M}$, be the difference

Table 6: Wage Regression - OLS (with weights)

Characteristic	Coefficient	St. Error	P-value
Age	0.561	0.075	0.000
Age - Squared	-0.006	0.001	0.000
Masters or higher	8.397	0.705	0.000
Bachelor's Degree	5.781	0.575	0.000
Associate's Degree	2.77	0.307	0.000
Some College	1.437	0.225	0.000
High School Dropout	-1.157	0.276	0.000
Junior High Dropout	-1.563	0.354	0.000
Black	-0.485	0.244	0.046
Hispanic	-1.068	0.355	0.003
South	-0.896	0.297	0.003
Live in City	1.552	0.397	0.000
Presence of Children	-0.320	0.109	0.003
Constant	-3.915	1.414	0.006
Observations	4618		
R-Squared	0.1778		

of the two utilities. A linear specification is given by the following:

$$P^*_{i.,M} = X_i \gamma + \nu_i \tag{13}$$

where,

$$P_{i,M} = \begin{cases} 1 & \text{if } P^*_{i,M} \geq 0 \Leftrightarrow X_i \xi \geq -\nu_i \\ 0 & \text{if } P^*_{i,M} \leq 0 \Leftrightarrow X_i \xi \leq -\nu_i \end{cases}.$$

The vector X_i is a vector of observable individual characteristics and M denotes the program ($M = \{WIC, FSP\}$). If we assume that ν_i is distributed standard normal, we can estimate how individual characteristics relate to the probability of participating in FSP and WIC using a probit model. The first set of results in Table 7 shows the marginal effect of individual characteristics on the probability of participating in FSP. The probability of participation decreases in educational attainment and is higher for women identifying themselves as black or Hispanic. The relationship between age and the probability of participating in FSP appears to be a "Ushape": the probability of participation is highest for the youngest (born after 1960) and oldest women (born before 1940).

Table 7: Welfare Participation - Probit (with weights)

	Participation in FSP			Participation in WIC		
Characteristic	dF/dX	St. Error	P-value	dF/dX	St. Error	P-value
Born before 1940	-0.039	0.012	0.003	-0.371	0.056	0.011
Born in 1940s	-0.091	0.010	0.000	-0.335	0.059	0.001
Born in 1950s	-0.060	0.011	0.000	-0.076	0.059	0.213
Born in 1960s	-0.004	0.013	0.747	-0.097	0.041	0.021
Masters or higher	-0.131	0.007	0.000	-	-	-
Bachelor's Degree	-0.143	0.007	0.000	-0.246	0.072	0.007
Associate's Degree	-0.065	0.011	0.000	-0.135	0.067	0.060
Some College	-0.062	0.009	0.000	-0.071	0.051	0.175
High School Dropout	0.145	0.020	0.000	0.013	0.056	0.816
Junior High Dropout	0.254	0.026	0.000	0.203	0.074	0.007
Black	0.163	0.014	0.000	0.156	0.044	0.000
Hispanic	0.093	0.021	0.000	0.121	0.060	0.004
South	-0.038	0.009	0.000	-0.022	0.042	0.600
Urban	-0.052	0.013	0.000	-0.095	0.001	0.087
Observations	6270			750	(with Child under 5)	
Likelihood	-2307.2			-476.4		

In estimating the probability of participation in WIC, the sample is restricted to only those women with children under age five (750 women). Similar to participation in FSP, the probability of participating in WIC is decreasing in educational attainment, as shown in the second set of results in Table 7. In fact, women with a Master's Degree or higher were dropped from the estimation due to perfect co-linearity with the constant term. Women identifying themselves of a minority race are more likely to participate in WIC: black women are nearly 16 percentage points more likely, and Hispanic women are 12 percentage points more likely to participate in WIC relative to white women. Unlike FSP, the probability of participating in WIC is lowest for older women. This could be due to the uniqueness of this family arrangement: these women are likely grandmothers or elder relatives taking on a guardian role. In both FSP and WIC, women living in a city are less likely to participate.

While the simple participation regressions provide some indication of the relative importance of the various factors that may be associated with the utility costs of participation, they are unable to distinguish psychological costs from time costs in explaining non-participation. For instance, that those with more education are less likely to participate may be because those with more education have higher psychological costs or it may be that the primary reason is that they have higher wages and thus would have higher time costs and are less likely to choose a level of labor supply that would make them eligible. In fact, the above estimation ignores earnings-based eligibility rules. One could include the wage (using a procedure to impute a wage for those with no observable wage) and a measure of income, but this would not allow for the joint determination of labor supply and welfare participation. The ability to do this is a clear advantage of the structural model developed in Section 3.

7 Results

The estimation procedure computes estimates of the structural parameters from the model developed in Section 3. The parameter estimates are given in Tables 8 and 9. The parameter estimates of the wage equation are consistent with other estimates in the literature: wages are concave in age, decreasing in the number of children, and increasing in education (obtaining only a high school degree is the excluded education category). Wages are also lower for workers in the South and those of minority status.

The lower half of Table 8 gives the estimates from the psychological cost equation. Because the psychological costs enter the utility function through consumption, the parameter estimates are measured in dollars per week. The first interesting result is that females born before 1940—those likely affected the Great Depression—have much lower psychological costs than other cohorts, which could be due to the emergence of entitlement programs following the Depression. Psychological costs of welfare participation are increasing in educational attainment, increasing dramatically for those obtaining a Bachelor's degree or higher. This finding is consistent with results from previous studies which found that the probability of welfare participation is decreasing in educational attainment. The last result worth noting is that females of

Table 8: Wage and Stigma Parameter Estimates

Wage Equation

Variable	Estimate	St. Error	p-value
constant	0.774	0.003	0.0000
Age	0.382	0.000	0.0000
Age-squared	-0.004	0.000	0.0000
Master's Degree or Higher	4.298	0.019	0.0000
Bachelor's Degree	3.284	0.040	0.0000
Associate's Degree	2.291	0.071	0.0000
Some College	0.974	0.023	0.0000
High School Dropout	-0.085	0.049	0.0822
Junior High Dropout	-4.610	0.073	0.0000
Urban Area	0.110	0.002	0.0000
Number of Kids	-0.472	0.019	0.0000
South	-1.208	0.010	0.0000
Hispanic	-0.111	0.059	0.0637
Black	-0.065	0.009	0.0000
σ_ϵ	3.147	0.002	0.0000

Psychological Cost Equation

Variable	Estimate	St. Error	p-value
constant	122.072	1.039	0.0000
Born before 1940	-115.157	10.889	0.0000
Born in 1940s	-0.294	4.725	0.9506
Born in 1950s	-0.087	1.718	0.9598
Born in 1960s	5.457	0.996	0.0000
Master's Degree or Higher	362.212	69.881	0.0000
Bachelor's Degree	236.773	18.519	0.0000
Associate's Degree	86.891	4.080	0.0000
Some College	80.610	0.896	0.0000
High School Dropout	-29.745	12.060	0.0153
Junior High Dropout	-127.924	13.540	0.0000
Urban	-6.362	0.804	0.0000
Teen in Household	2.348	1.383	0.0926
South	48.167	0.956	0.0000
Hispanic	-56.319	13.712	0.0001
Black	-92.411	2.531	0.0000
σ_η	300.122	0.558	0.0000

Table 9: Time Costs Results

Variable	Estimate	St. Error	p-value
FSP Time Cost (δ_1)	1.212	0.067	0.0000
WIC Time Cost (δ_2)	6.392	0.034	0.0000
δ_3	0.826	0.110	0.0000
δ_4	-14.324	0.051	0.0000
δ_5	2.185	0.150	0.0000
Time Endowment	109.631	0.070	0.0000
Fixed Cost of Work	7.363	0.012	0.0000
Utility parameter (γ_1)	0.669	0.000	0.0000
Utility parameter (γ_2)	0.002	0.000	0.0000

minority status face lower psychological costs on average.

To give an idea of the magnitude of these costs, we can compute the costs for a set of observable characteristics: for a white female born in the 1960s with a Bachelor's degree who lives in a urban environment in a non-Southern state with no children, the psychological cost associated with participating in welfare is \$357.94 per week. However, much of the variation in psychological costs is attributable to unobserved heterogeneity; the standard deviation of the psychological costs is approximately \$300 dollars.

Table 9 shows the estimates for δ_1, the time cost of FSP, and δ_2, the time cost associated with WIC. These estimates imply that receiving benefits from FSP requires 1.2 hours a week on average. The estimated time cost associated with obtaining benefits through WIC is much higher: 6.4 hours a week. Receipt of benefits through WIC involves doctors visits, nutritional education, and more frequent office visits, which explains the greater time requirement.

Valuing the time costs by the predicted wage for each participant, we can compare the relative magnitude of time costs to psychological costs. Table 10 shows average time costs for participants and non-participants. As expected, time costs and psychological costs are higher for participants relative to non-participants. For participants, average psychological costs are three times the size of average time costs. This implies that psychological costs are more important that time costs for participants in these welfare programs. This also implies that there are

Table 10: Costs Associated with Participation

	Average		Median	
	Participants	Non-Participants	Participants	Non-Participants
Time cost (FSP)	$8.57	$11.74	$8.58	$11.70
Time cost (WIC)	$43.33	$59.91	$45.26	$61.69
Psychological cost	$55.13	$190.97	$64.83	$171.69
For Participants Only				
Combined time cost	$20.24		$10.33	
Psychological cost	$55.13		$64.83	

greater potential gains from policies that reduce the psychological costs associated with welfare participation–increasing the level of transfer payments in the income tax system–relative to policies that streamline the benefits process.

The variation in the time costs of welfare participation is entirely due to variation in the predicted wage rates. In contrast, there is a large degree of variation in the estimated psychological costs. The correlation of the estimated psychological cost and other characteristics provides valuable information about the distribution of psychological costs. Figure 1 shows the average psychological cost by the observed wage rate and Figure 2 shows the average psychological cost by the observed hours of work. Average psychological cost is increasing in both observed wages and hours of work.

Figure 1: Average Psychological Cost by Observed Wage Rates

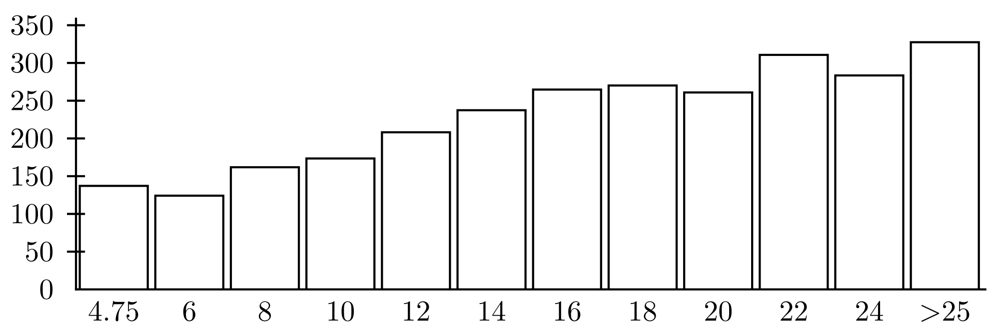


Figure 2: Average Psychological Cost by Observed Hours of Work

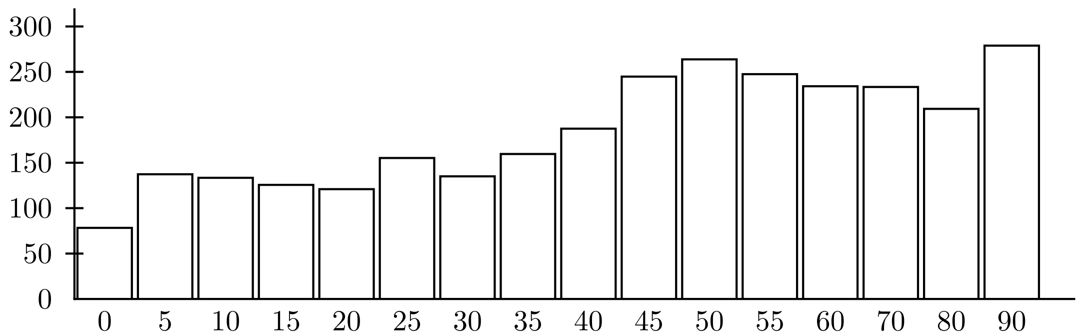


Table 9 also reports the estimated time endowment and fixed costs of work.[15] Recall that fixed costs of work are assumed to be proportional to the individuals wage, which gives them an interpretation in hours. The estimated value implies that working costs are equivalent to seven and a half hours of earnings per week. The utility parameter γ_1 measures the degree to which the marginal utility of consumption declines with additional consumption. Similarly, the utility parameter γ_2 measures the degree to which the marginal utility of leisure declines with additional leisure.

Comparing the participation decision predicted by the estimation to actual participation behavior allows us to evaluate the accuracy of our model and empirical specification. Tables 11 and 12 show actual versus predicted participation behavior for FSP and WIC. The predicted participation choice for each individual is calculated as the participation combination that yields the highest utility given a value of zero for both error terms. If the observable characteristics in the empirical specification perfectly predicted participation, there would be no weight in the "off-diagonal" elements of the tables. For FSP, the observable characteristics are able to correctly predict participation for nearly 85 percent of individuals; these characteristics correctly predict WIC participation for nearly 96 percent of individuals. The substantial fraction of incorrect predictions is not surprising given the importance of unobserved heterogeneity in determining welfare participation.

[15] The parameters δ_3, δ_4, and δ_5 allow the time endowment to adjust depending on the composition of the family.

Table 11: FSP Participation, Actual and Predicted Percentages

	Predicted Non-Participant	Predicted Participant	Total
Actual Non-Participant	78.38	3.85	82.23
Actual Participant	11.89	5.88	17.77
Total	90.27	9.73	100

Table 12: WIC Participation, Actual and Predicted Percentages

	Predicted Non-Participant	Predicted Participant	Total
Actual Non-Participant	93.49	0.94	94.42
Actual Participant	3.73	1.85	5.58
Total	97.21	2.79	100

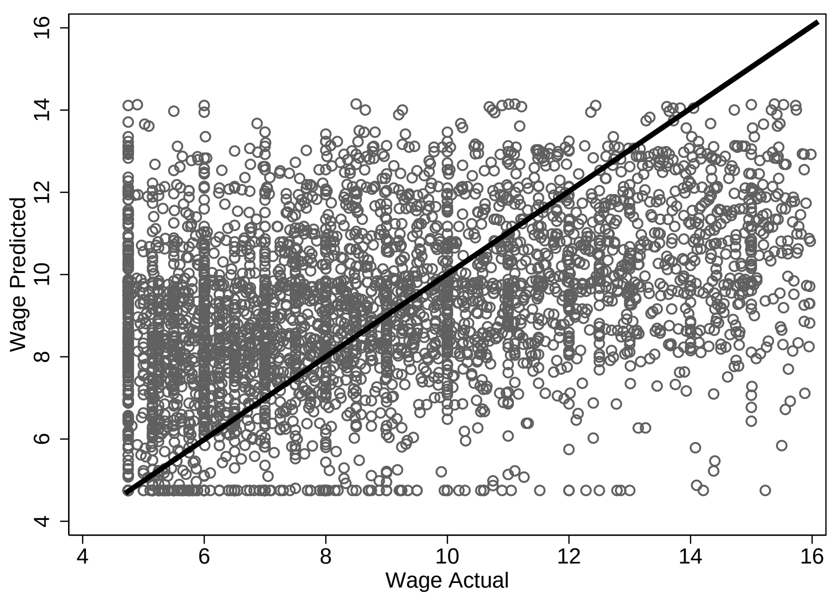
An additional method for evaluating the model and empirical specification is to compare actual to predicted hours of work as shown in Table 13. Overall, predicted hours match actual hours in terms of predicting a large mass of individuals at zero and at forty hours. However, the estimation under-predicts the number of non-workers (zero hours) and the number working precisely 40 hours; instead, it over-predicts the number of females working 30 to 35 hours and 45 to 50 hours. In addition, there are no predicted hours above 60, while 4 percent of the sample is in this range. A comparison of predicted wage to actual wage for individuals is shown in Figure 7. In general, the estimation tends to slightly over-predict wages, as shown by the mass of wages just above the 45 degree line. However, the highest wage predicted by observable characteristics is under \$15 per hour. The mass of predicted wages at \$4.75 is due to imposing the minimum wage on both the data and in the estimation.

It is important to note that these results rely not only on the data selected for the exercise but also on the specification of the model. In particular, we achieve identification from the assumption that psychological costs do not increase in the number of programs and the specifi-

Table 13: Hours of Work Actual and Predicted Factions

Hours	Actual	Predicted
0	25.08	16.97
5	0.88	0.00
10	1.02	0.02
15	1.54	0.62
20	2.78	1.06
25	2.40	2.31
30	4.09	8.75
35	5.93	12.12
40	34.15	15.48
45	6.37	20.48
50	5.36	20.20
55	2.99	1.53
60	3.30	0.46
70	2.05	0.00
80	1.48	0.00
90	0.39	0.00
100	0.17	0.00

Figure 3: Actual and Predicted Wage Rate



cation of the utility function (i.e. how time and psychological costs enter the utility function). The relative importance of these assumptions in achieving identification is unclear.

8 Stigma as a Screening Mechanism

While this paper's goal has been to empirically separate time costs from psychological costs associated with welfare participation, they are not necessarily a pure cost to society. The utility cost of welfare participation might be a useful way of distinguishing high-ability from low-ability individuals who both have low income. In the context of asymmetric information (the government only observes income, not ability) a screening mechanism can be used to encourage individuals to self-select. If the government wants to only provide income transfers to those individuals with low ability, without some selection mechanism, it will provide transfers to both high- and low-ability individuals because it can only provide benefits based on income. Because welfare benefits are available to all individuals with low income, some high-ability individuals could choose to expend less work effort to qualify.

Here we develop a simple model to show the conditions under which utility costs could be used as a screening mechanism. There are two types of agents, high and low ability types $\{\theta_H, \theta_L\}$. The agent chooses work effort and participation in welfare ($P = \{0, 1\}$). For convenience, we assume that there are only two possible income levels, high income (I_H) and low income (I_L). If the agent chooses work effort such that she has high income, she is not eligible for welfare participation. Thus the possible choice set is $\{(I_H, P = 0), (I_L, P = 0), (I_L, P = 1)\}$.

Welfare participants receive benefit level B and incur utility cost $C(\theta)$. Individuals also bear a utility cost of work effort, $g(I, \theta)$. By assumption, the cost of work effort is such that high-ability agents would choose to be high income in the absence of a welfare program, $I_H - g(I_H, \theta_H) > I_L - g(I_L, \theta_H)$. The cost obtaining high income for low-ability agents is prohibitive, $I_L - g(I_L, \theta_L) > I_H - g(I_H, \theta_L)$; they will always choose low income. The agent's utility function is given by:

$$U(I, P|B, \theta) = I - g(I, \theta) + \mathbb{1}(P = 1)(B - C(\theta)). \tag{14}$$

Because there is a discrete choice set, the maximization problem can be expressed as:

$$U_H = \max\big\{I_H - g\left(I_H, \theta_H\right), \, I_L - g\left(I_L, \theta_H\right), \, I_L - g\left(I_L, \theta_H\right) + B - C\left(\theta_H\right)\big\} \tag{15}$$

$$U_L = \max\big\{I_H - g\left(I_H, \theta_L\right), \, I_L - g\left(I_L, \theta_L\right), \, I_L - g\left(I_L, \theta_L\right) + B - C\left(\theta_L\right)\big\} \tag{16}$$

A effective screening mechanism is one that induces high-ability types to not participate and low-ability types to participate. Given the assumptions on the cost of work effort, we can reduce the maximization problem to the following:

$$U_H = \max\big\{I_H - g\left(I_H, \theta_H\right), \, I_L - g\left(I_L, \theta_H\right) + B - C\left(\theta_H\right)\big\} \tag{17}$$

$$U_L = \max\big\{I_L - g\left(I_L, \theta_L\right), \, I_L - g\left(I_L, \theta_L\right) + B - C\left(\theta_L\right)\big\} \tag{18}$$

Therefore, if $C(\theta_L) < B < C(\theta_H)$ then the high-ability type will choose to not participate and the low-ability type will choose to participate. Hence, utility costs of welfare participation can result in a separating equilibrium if the costs are increasing in ability. The time costs are increasing in the wage rate because they reflect the opportunity cost of the individual's time. Additionally, the earlier estimation indicates that psychological costs are higher for those women with higher educational attainment and wages, which suggests that utility costs could acting as screening mechanism. While the government may not be generating welfare stigma, it could be using it to induce sorting to avoid welfare participation by high-ability types.

This paper finds that psychological costs are positively correlated with indicators of high ability, however, other researcher have shown that these costs are higher for those in poor health and for minorities in predominately white communities (Stuber and Schlesinger, 2006). This diminishes the effectiveness of using psychological costs as a screening mechanism since the government might not want to preclude participation by these individuals. To determine the impact of stigma on social welfare, the benefit from screening needs to be compared to utility costs of welfare participation. Future work could develop a model in which the government has two instruments for transferring income - high-stigma welfare program and a low-stigma tax transfer. The low-stigma tax transfer would not provide the benefits of screening, but would also not impose stigma on recipients. The high-stigma welfare program would provide the benefit of screening as described above. The optimal transfer policy would likely involve both instruments; future work will determine the appropriate mix.

9 Conclusion

This paper differs from the previous studies that seek to estimate the utility costs of welfare participation because it distinguishes psychological costs from time costs. We develop a model of labor supply and participation in multiple welfare programs that we estimate using data on participation in FSP and WIC by female household heads in the SIPP. We estimate the model using a simulated maximum likelihood procedure. To identify psychological costs and time costs, we assume that the psychological cost does not increase in the number of programs in which they participate. However, time costs depend on the number and type of programs. Another source of identification is the way in which these costs enter the utility function: psychological costs are assumed to decrease utility by reducing effective consumption, while time costs of participation reduce available leisure time.

We find that the time cost associated with participation in FSP is 1.2 hours a week. This implies that the time required to participate in FSP is approximately 5 hours a month. The estimated time cost associated with WIC is much higher: 6.4 hours a week. This difference in time cost is consistent with the more time-intensive requirement of WIC, including doctor visits, nutritional education, and more frequent office visits.

Psychological costs are estimated in dollars because they enter the utility function through effective consumption. We estimate that the average psychological cost for a participant is approximately \$55 per week, which is lower than the average of \$164 for the population. Psychological costs are lower for those women born before 1940, which could be related to the introduction of entitlement programs following the Great Depression. These costs strongly increase in educational attainment and are lower for women who identify themselves as black or Hispanic. By using the predicted wage to compute the dollar value of time costs associated with participation, we find that psychological costs are three times larger than the time costs of welfare participation. Separately identifying the components of the utility cost associated with participation in welfare is important to welfare reform and polices designed to more effectively reach the target population.

Future work could extend this model by incorporating participation in other welfare programs, namely AFDC or TANF. In addition, it would be desirable to test the robustness of the results using different assumptions regarding the way in which costs enter the utility function. While this paper presented some evidence that a lack of information was not a major determinant of non-participation in WIC, further exploration of this explanation is also needed. Another shortcoming of the model is that it is static and does not consider the duration or timing of participation. Future work could develop a dynamic model of welfare participation that would fully incorporate the participation time limits imposed during the Welfare Reform of 1996.

The estimated model could be used to assess the utility implications of different distribution policies, such as tightening welfare program eligibility requirements while expanding the EITC program in a way that preserves existing expenditures levels. Given the large estimates of the psychological costs of welfare participation, if a policy of reducing the role of traditional welfare programs and replacing them with less visible transfer programs is able to decrease the psychological costs, such a policy could significantly increase social welfare. Estimating this same model using more recent data from the 2004 SIPP could provide evidence of whether the welfare reform, including the adoption of EBT, increased or decreased the relative importance of psychological costs.

References

Blundell, R., V. Fry, and I. Walker (1988): "Modelling the Take-Up of Means-Tested Benefits: The Case of Housing Benefits in the United Kingdom," *The Economic Journal*, 98(1), 58–74.

Currie, J. (2003): "U.S. Food and Nutrition Programs," in *Means-Tested Transfer Programs in the United States*, ed. by R. A. Moffitt. Chicago: University of Chicago Press.

Greenbook (2004): *U.S. House of Representatives Green Book, 108th Congress.* Washington, D.C.: Ways and Means Committee.

Hotz, V. J., and J. K. Scholz (2003): "The Earned Income Tax Credit," in *Means-Tested Transfer Programs in the United States*, ed. by R. A. Moffitt. Chicago: The University of Chigao Press.

Keane, M., and R. Moffitt (1998): "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, 39(3), 553–589.

Moffitt, R. (1983): "An Economic Model of Welfare Stigma," *American Economic Review*, 73(5), 1023–1035.

——— (1992): "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature*, 30(1), 1–61.

Riphahn, R. T. (2001): "Rational Poverty of Poor Rationality? The Take-up of Social Assistance Benefits," *Review of Income and Wealth Series*, 47(3), 379–398.

Smeeting, T. (1982): "Alternative Methods for Valuing Selected In-Kind Benefits and Measuring Their Effect on Poverty," Technical Paper Number 50, Washington D.C.: Bureau of the Census.

Stuber, J., and M. Schlesinger (2006): "Sources of stigma for means-tested government programs," *Social Science & Medicine*, 63, 933–945.

Trippe, C., and P. Doyle (1992): *Food Stamp Program: Eligibility and Participation.* Washington, D.C.: U.S. Congress, Congressional Budget Office.