AN EMPIRICAL ANALYSIS OF WELFARE DEPENDENCE

By

Mark W. Plant

University of California, Los Angeles

Discussion Paper #237

March 1982

* This paper was taken from research done for my dissertation at Princeton University. I am grateful to Orley Ashenfelter, James Brown, George Cave, and Richard Quandt for useful comments and suggestions. Financial support was provided by the National Science Foundation, the Industrial Relations Section of Princeton University and the Institute of Industrial Relations at U.C.L.A.

Introduction

Recipients of the benefits of social welfare programs often receive these benefits for prolonged periods. For some observers this fact is evidence for Michael Harrington's classic statement that the culture of poverty has produced a two class economic system with some members continuously in need of assistance (Harrington, 1962). For others this fact is evidence that, like a narcotic, social programs create their own dependence. Once potential recipients have benefited from a social program, it is asserted, they are less likely to become non-recipients.

This essay presents the results of an effort to sort out the issues involved in the time pattern of receipt of benefits in an experimental negative income tax program. The intial empirical results use the data from the Seattle-Denver Income Maintenance Experiments, the largest and most comprehensive experiments to date. The basic pattern of welfare dependence is reflected in the Seattle-Denver data. In the first three years of the experiment 33.4 of the potential recipients received benefits in all three years while 30.2 received no benefits in any of the three years. Only 36.4 had mixed patterns of benefit receipt in the first three experimental years. From these data alone one can only infer that a variable indicating whether a family receives benefits from a negative income tax (NIT) is serially correlated. The intent of this paper is to determine the cause of this serial correlation by formulating an econometric model of participation behavior that allows us to distinguish among three economic phenomena that may result in persistent welfare participation.

The correlation of earnings over time is the first reason one might observe welfare dependence. Earnings patterns are a result of education, family size, race, ethnic heritage among a host of other components. If earnings in society are structured in such a way that poor families stay poor, then those families will be chronically dependent on welfare. The welfare system will affect participation only if it results in a change in the structural characteristics of the population.[1] The predictable components of the time pattern of earnings can be deduced from a longitudinal data set using some well known econometric tehniques. (See, for example, Lillard and Willis (1978), Heckman (1980), Heckman and Willis (1977), Lillard and Weiss (1979), Weiss and Lillard (1978)). The components can be summarized in an earnings function that enables one to predict the distribution of earnings in any given time period. When coupled with the demographic profile of any population, the estimated earnings function allows the prediction of the time path of the distribution of earnings for that population. In particular one could predict the fraction of the population that would be eligible for welfare payments in any given year. A population that exhibited a highly correlated earnings structure would exhibit a high correlation in welfare participation behavior. To remedy such chronic welfare participation, the government would have to change the fundamental earnings structure of the population. Any change in the welfare system would have no short term effects on the correlation in participation behavior.

The purely statistical analysis of welfare behavior based on an earnings function ignores economic behavior that may result in persistent receipt of welfare benefits. For example, a family might adjust its work effort to become eligible for payments. This potential change in work behavior is the second phenomenon that may result in chronic welfare participation.[2/] It is necessary to incorporate into the statistical model an economic model of behavior to guage correctly the effect a welfare system is having on work effort.

Finally, we might observe welfare dependence because families, once they become recipients of welfare benefits in one period, become dependent on these payments in later periods. The family which has a temporary need for welfare support becomes "addicted" to those benefits. Statistically, this notion of a welfare trap means that the chance of receiving welfare benefits this year depends on whether you received them in the past, ceteris paribus. If indeed this trap is the reason for welfare dependence, then the appropriate policy response is not to allow persistent participation to occur, that is not to encourage bad habits.

The remainder of this essay presents a model of participation in NIT programs, that will enable us to determine empirically the causes of persistent welfare participation. In Section I, we present a simple model of participation based on a stochastic earnings function. In Section II, we incorporate into the model the potential for behavioral response and welfare dependence. Section III presents the basic empirical results and Section IV presents some alternative estimates of the paramteters of the model. Concluding remarks are made in Section V.

Section I: The Statistical Approach to Participation Behavior

The fact that a family's earnings are correlated over time implies that any observable variable determined by earnings will be serially correlated. In particular, since the receipt of benefits from any type of welfare program depends on earnings, welfare participation patterns in the eligible population should exhibit some serial correlation. The extent of such correlation depends on the stochastic nature of earnings. In this section, we present a model of participation behavior that relies only on a statistical characterization of earnings in the population. This model is to be used as a benchmark for analyzing actual participation behavior. Since we know some persistence in the receipt of benefits is simply due to serially correlated earnings and bears no short term relationship to the disincentive effects of welfare, measurement of this basic immobility is necessary before trying to discern the disincentive effects of welfare programs. In this section we show how we can characterize this basic underlying correlation in an NIT program and present the basic data we will use throughout this study.

In an NIT, a family receives a subsidy if its income is less than the specified breakeven level.[4/] Rather than spelling out an explicit model of earnings determination, let us view income as a stochastic function of a family's characteristics (for example: size, education, location). The pattern of benefit receipts over several years depends on the family's "draws" from the income distribution. In any year that a family's income is below the breakeven level the family reveives a subsidy. Since income draws are correlated over time, subsidy receipts will be serially correlated. If the multi-period income distribution for the population of eligible NIT recipients is known, the aggregate patterns of participation can be predicted.

Using these predictions we can determine if persistent welfare participation in the population is due to the structure of earnings in the population, or if other explanations, such as the work disincentives of welfare, should be considered.

To make this notion of a statistical characterization of participation behavior more precise, consider a family indexed by i whose income in period t, Y_{it}, is stochastically determined.[5/] Suppose this family is exposed to an NIT program with guarantee level G and tax rate τ.[6/] In period t this family participates if and only if $Y_{it} \leq \frac{G}{\tau}$. Let d_{it} be defined as the dichotomous variable

$$d_{it} = \begin{cases} -1 \text{ if family i does not participate in period t} \\ 1 \text{ if family i does participate in period t.} \end{cases}$$

Then the probability of observing d_{it} is

$$\text{(1)} \qquad Pr[d_{it}] = Pr\,[Y_{it}\, d_{it} \leq \frac{G}{\tau}\, d_{it}]$$

Define the participation pattern $d_i = (d_1, d_2, \ldots, d_{iT})$,

and define $Y_i = (Y_{i1}, Y_{i2}, \ldots, Y_{iT})$, then

$$\text{(2)} \qquad Pr[d_i] = Pr[Y_i \circledast d_i \leq \frac{G}{\tau} \times d_i]$$

where $\circledast$ indicates element by element product operation[7/] and x indicates scalar multiplication into a vector. If Y_i follows some known T-variate density function, denoted as f (•) , the probability (2) can be evaluated as a multivariate integral of f(•).[8/] If f(•) is known for the eligible population and if the NIT is not affecting the income distribution then the fraction of eligible families exhibiting any particular participation pattern should be close to the probability calculated from integrating f(•) appropriately.

Any statistically significant deviations of the actual patterns of participation from the predicted patterns derived from f(•) would indicate that one cannot explain participation patterns in the NIT simply on the basis of earnings correlation, but instead the NIT is effecting some behavioral response that is altering the distribution of income in the population.

To predict the fraction of families that will participate in an NIT program, we must have some information on the income distribution which the statistician would characterize by f(•). The NIT experiments were in theory designed to give us the required distribution information.[9] The data gathered include observations on families not eligible for the NIT payments whose labor force and earnings behavior was simply recorded. This set of families constitutes the control portion of the sample, and provides us with the statistical benchmark for analyzing the behavior of the experimental families who were eligible for NIT subsidies.

Assume that both the experimental and control families were drawn from the same population. If there were no change in the income distribution induced by the NIT, unbiased and consistent estimates of the fraction of experimental families who would receive NIT subsidies can be calculated using the data on the control families. To do this one need only calculate the percentage of control families whose income is below the breakeven income. We can then compare our predicted participation patterns with the actual patterns exhibited by the experimental families. If this statistical model of participation behavior is adequate, then the predictions should be close to the actual realizations. If there is a behavioral reaction to the NIT, the realizations will not square with the predictions. The appealing feature of this comparison is that no explicit specification of the distribution f(·) needs to be made. The only maintained hypothesis necessary is that the control and experimental families are chosen from the same population. We do not need to characterize that population's earnings behavior mathematically.

The comparison between predicted and actual behavior for the Seattle-Denver Income Maintenance Experiment is presented in Table I. These estimates are based on data gathered on 958 white families with two parents present. Of these families, 591 were exposed to an NIT program for three years and the remaining 367 were observed as a control group.[10] In any given year experimental families were said to have participated in the NIT program if the payment made to them in either the first six months or the last six months exceeded $120.[11] Experimental families who left the program were dubbed non-participants from the time they left.[12] Each experimental family was initially assigned to one of eleven NIT programs which differed in their guarantee levels and tax rates. The only characterization of the individual NIT programs necessary for this analysis is the breakeven level of which there are ten. Some experimental families were assigned to three year NIT programs and some to five year NIT programs.[13] Any family, control or experimental, who changed experimental status, NIT program or length of NIT program was deleted from the sample.

In Table I, the first column lists the eight possible participation patterns in the three years of the experiment, 0 denoting nonparticipant and 1 denoting participant. The ten subsequent columns show the predicted and actual percentage of the sample following a given participation pattern for a given breakeven.[14] For example, for the breakeven income level of $5430,

TABLE I: Predicted and Actual Participation Patterns of Families in the Seattle-Denver Income Maintenance Program

Breakeven in Dollars

Participation Pattern Period 3 2 1	5430 Pred. %	5430 Act. %	5801 Pred. %	5801 Act. %	6650 Pred. %	6650 Act. %	7366 Pred. %	7366 Act. %	7600 Pred. %	7600 Act. %	8000* Pred. %	8000* Act. %
0 0 0	58.7	66.7	43.1	38.7	51.3	32.4	25.2	21.1	31.1	20.6	35.8	34.4
0 0 1	11.5	13.3	20.1	6.5	12.0	20.3	19.4	7.9	15.0	17.6	15.7	9.8
0 1 0	1.3	0.0	2.4	3.2	1.8	4.1	2.3	0.0	1.5	5.9	2.2	4.9
0 1 1	3.6	6.7	4.4	6.5	5.0	4.1	6.7	18.4	6.2	8.8	6.5	4.9
1 0 0	6.1	3.3	5.1	9.7	5.4	6.8	4.2	5.3	3.7	5.9	5.4	3.3
1 0 1	2.1	3.3	2.0	3.2	4.4	2.7	6.6	2.6	6.7	5.9	3.7	1.6
1 1 0	3.6	3.3	2.5	6.5	3.2	5.4	5.7	7.9	4.1	4.4	2.4	6.6
1 1 1	13.2	3.3	20.5	25.8	17.0	24.3	29.8	36.8	29.1	30.9	28.1	34.4
Number of Observations	30		31		74		38		68		61	
$\chi^2(7)$	4.40		7.26		15.87		13.43		13.12		10.09	

Participation Pattern Period 3 2 1	8000** Pred. %	8000** Act. %	9600 Pred. %	9600 Act. %	10343 Pred. %	10343 Act. %	11200 Pred. %	11200 Act. %	12000 Pred. %	12000 Act. %
0 0 0	29.5	25.6	22.7	17.4	17.2	22.4	12.1	12.5	7.8	11.6
0 0 1	16.6	20.9	14.3	8.7	11.6	9.2	8.6	6.3	5.7	10.1
0 1 0	2.1	7.0	3.3	2.9	2.1	3.9	2.3	0.0	0.3	1.4
0 1 1	5.7	11.6	6.2	7.2	8.8	5.3	8.8	6.3	7.8	15.9
1 0 0	4.7	4.7	2.6	4.3	3.8	2.6	2.7	6.3	2.0	7.2
1 0 1	4.9	7.0	6.2	1.4	5.4	1.3	4.4	0.0	3.5	2.9
1 1 0	2.5	0.0	2.5	13.0	2.6	5.3	2.1	3.1	2.4	2.9
1 1 1	34.0	23.3	40.8	44.9	48.5	50.0	59.0	65.6	70.5	47.8
Number of Observations	43		69		76		32		69	
$\chi^2(7)$	11.13		37.60		8.55		4.44		27.13	

*This program does not have a declining tax rate ($G = 5600$, $\tau = 0.7$)

**This program does have a declining tax rate ($G = 4800$, $\tau = 0.8$, rate of decline = 0.000025) using the control data 58.7% of the sample would be expected not to participate in any of the three periods, and 66.7% of the experimental sample actually followed that pattern. Beneath each column, the number of experimentals in a program with the appropriate breakeven level is given followed by X^2 goodness of fit statistic. The X^2 statistic provides a test of the null hypothesis that the actual participation pattern frequencies were generated from a multinomial distribution with the predicted participation percentages as the true probabilities.[15] Again considering Table I, there were 30 families exposed to an NIT program with breakeven $5430. One cannot reject the hypothesis that the observed actual percentages were drawn from a multinomial distribution characterized by the predicted percentages since the X^2 statistic of 4.40 does not exceed the critical values at the 95% confidence level.

Table I does not allow us to make any firm conclusion regarding the statistical model of participation behavior. For three of the eleven breakevens we can reject the pure statistical model at the 95% confidence level. For both low and high breakeven levels the predicted rates of participation tend to be overestimates of the actual participation rates, especially in the $12,000 breakeven program where 70.5% of the families are predicted to participate in all three years, and only 47.8% of families do so. However, the programs with intermediate breakeven levels show underpredictions of participation. This underprediction for the "111" pattern ranges from 1.8% for the 7600 breakeven program to 7.3% for the 6850 breakeven program. This phenomenon is also demonstrated if actual participation rates are compared with predictions year by year as is done in Table II. For each breakeven the yearly

Table II

Yearly Participation in the Income Maintenance Experiments
for White Families

Breakeven	Year 1		Year 2		Year 3	
	Pred %	Act %	Pred %	Act %	Pred %	Act %
5430	30.4	26.6	21.7	13.3	25.0	13.2
5801	47.0	42.0	29.8	42.0	30.1	45.2
6850	38.4	51.4	27.0	37.8	30.0	39.2
7366	62.5	65.7	44.5	63.1	46.3	52.6
7600	58.0	63.2	40.9	50.0	43.6	47.1
8000*	54.0	50.7	39.2	50.8	39.6	45.8
8000**	61.2	62.8	44.3	41.9	46.1	35.0
9600	67.5	62.2	52.8	68.0	52.1	63.6
10343	74.3	65.8	62.0	64.5	60.3	59.2
11200	80.8	78.2	72.2	75.0	68.2	75.0
12000	87.5	76.7	81.0	68.0	78.4	60.8

*This program does not have a declining tax rate ($G = 5600$, $\tau = 0.7$)

**This program does have a declining tax rate ($G = 4800$, $\tau = 0.8$, rate of decline = 0.000025) predicted and actual participation rate are given. Again note the persistent underprediction in the mid-range breakeven programs. There seems to be an indication that the statistical model cannot completely explain the observed participation patterns. To make any definitive conclusion an economic model of participation behavior needs to be formulated so we can test the statistical model against some specific alternative. In the next section we present such a model that allows for an economic reaction to the NIT program as well as the possibility of welfare dependence.

Section II: Economic Effects of the NIT on Participation Behavior

It is widely recognized that any type of welfare subsidy program provides a potential work disincentive to the recipient population. The statistical model of participation behavior does not account for any such economic response. The use of participation patterns inferred from the control population to predict the actual response to a welfare program is predicated on a hypothesis that the income distribution does not change as a result of the NIT. In this section we develop a model of participation behavior of a utility maximizing economic agent with an eye towards empirical estimation.[16/]

The canonical indifference curve diagram for the NIT is shown in Figure 1. ACD is the budget constraint with no NIT, and BCD the budget constraint given an NIT. If the family initially chooses OM hours of leisure, and if leisure is a normal good the imposition of the NIT will result in the choice of OM' hours of leisure, some amount greater than OM. This simple diagram clearly indicates that any family with income less than the breakeven level of income OY* will choose to participate and will decrease hours of work if leisure is a normal good. This participation choice agrees with our statistical model; however, the decrease in earnings resulting from the increased consumption of leisure indicates the subsidy the family actually receives is greater than would be predicted based on the preexperimental labor force behavior.[17/]

The statistical participation rule is not sufficient when analyzing participation behavior in an economic framework. Although those families who receive income less than the breakeven level do participate as predicted, families whose income in the absence of the NIT would be greater than the break-

FIGURE 1

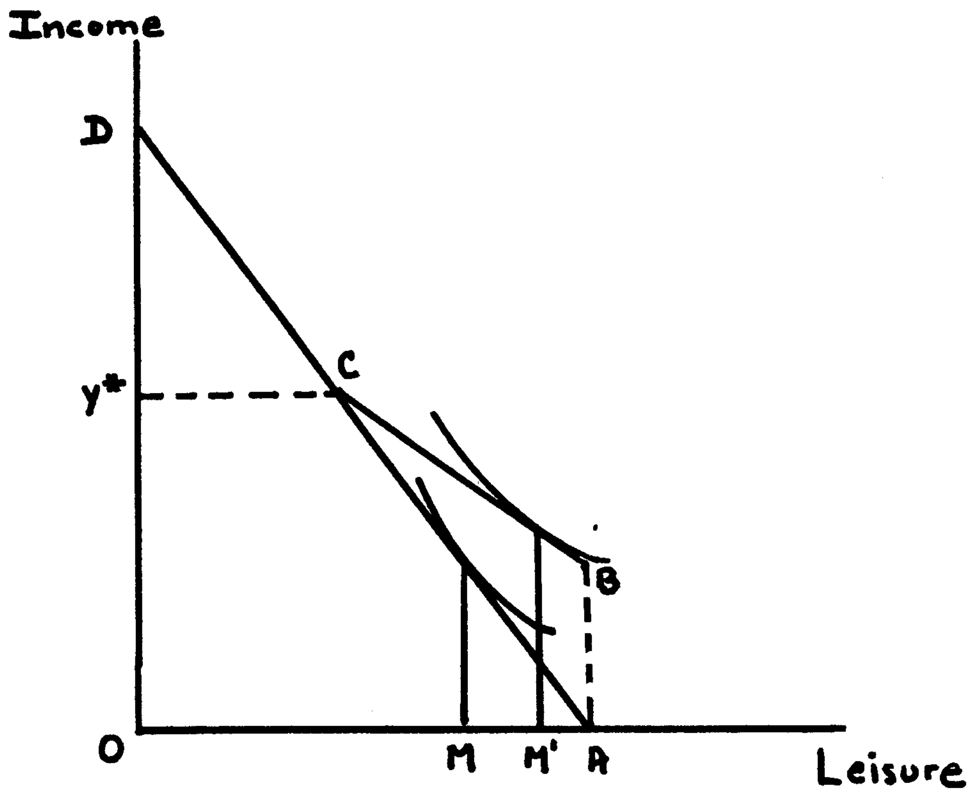
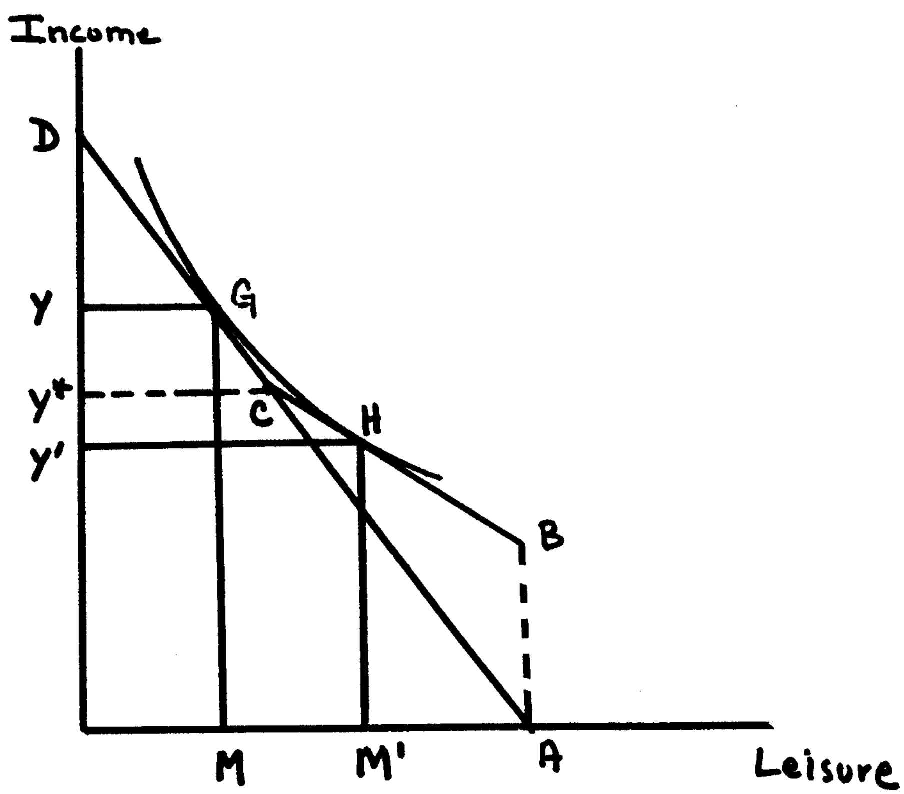


FIGURE 2



even level may also participate. Consider Figure 2. The family is indifferent between points G and H. If the family chooses G it receives no subsidy, and consumes OM hours of leisure and OY income. If the family chooses H, it consumes OM' hours of leisure and has earned income and subsidy which total OY'. Using a purely statistical model of participation the family whose choice problem is illustrated in Figure 2 would not be counted among those predicted to participate because its income is greater than the breakeven. The family acts in an economic fashion, substituting increased leisure for income and thus the predicted behavior based on income in the absence of an NIT does not coincide with the actual behavior when the family is eligible for the subsidy.

The extent to which there will be a significant economic effect on participation patterns depends on preferences and on the implicit tax rate on earnings for those families eligible to receive an NIT subsidy. Consider the family whose preferences are depicted in Figures 3 and 4. In Figure 3, we have depicted an NIT program with a very low tax rate.[18] In Figure 4, we have reproduced the family's indifference map, but we have depicted an NIT program with a higher tax rate holding the breakeven level constant. From the two figures, it is clear that the family would not change its behavior to receive a subsidy in the low tax rate program, but would do so in the high tax rate program. If there is an economic effect that induces participation, participation rates will be positively related to program tax rates, holding the breakeven level constant. The magnitude of this tax rate effect will depend on the extent to which leisure and income are substitutes. The flatter the family's indifference curves, the more willing they will be to substitute

FIGURE 3

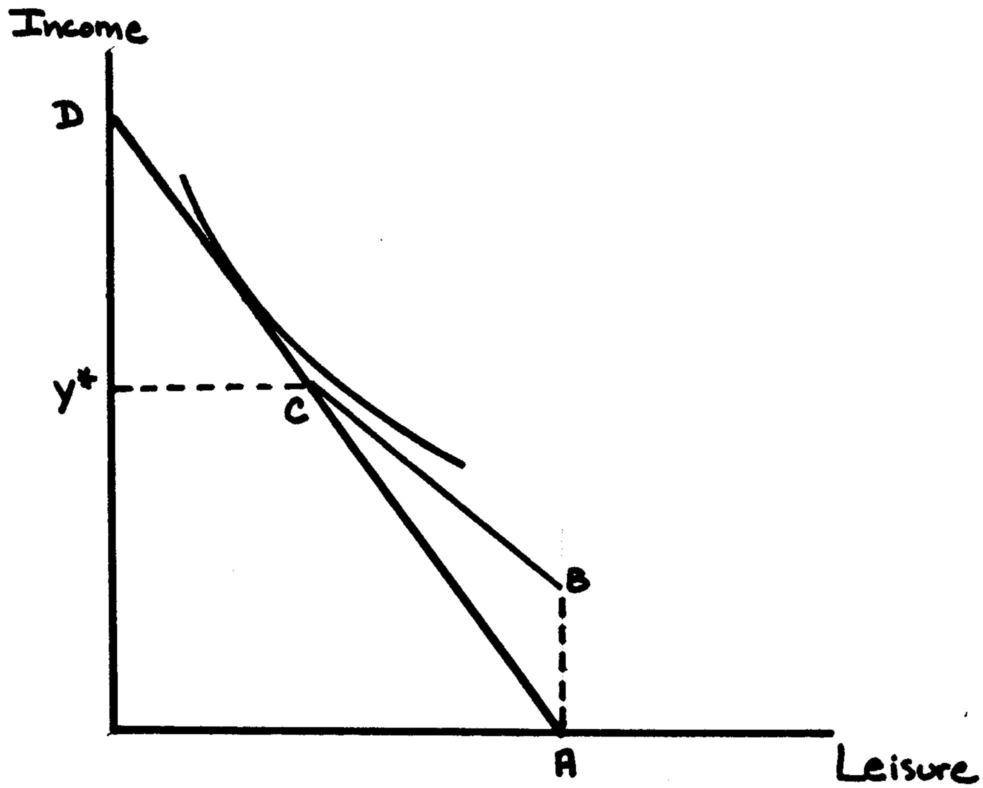
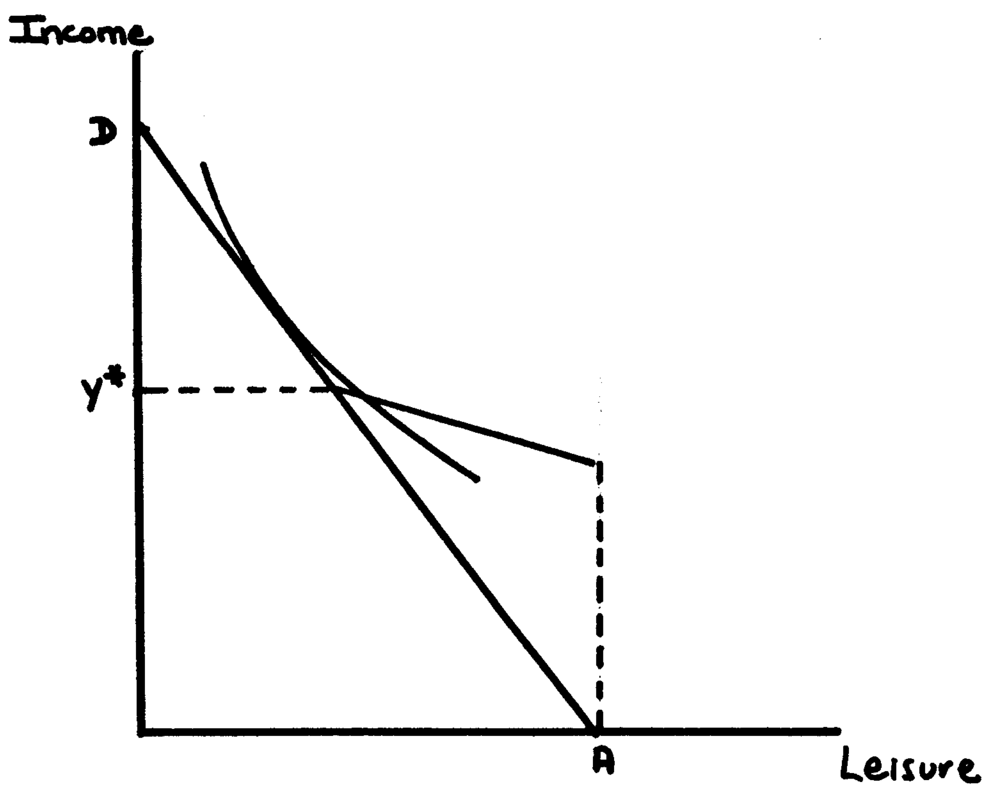


FIGURE 4



leisure for income and the more sensitive they will be to the implicit NIT tax on earnings. We would suspect that the magnitude of the tax rate on participation rates would give an empirical handle on the nature of the family's income/leisure preferences. In fact, Ashenfelter (1980) shows that a family participates in an NIT program if and only if

$$\ln Y_t \lessgtr \ln \left(\frac{G}{\tau}\right) + 1/2\, e\, \tau \tag{2}$$

where e denotes the compensated elasticity of labor supply. Although this is a local result, derived from a second order approximation to the expenditure function, it gives us some foundation for examing discrete changes. The essence of equation (2) is that the critical level of the logarithm of earnings at which families begin to participate is the log of the breakeven plus a term that increases as the indifference curves become flat. The more easily income can be substituted for leisure, the larger is e, and the larger the critical earnings level. Throughout the remainder of the paper, we will use e as the summary measure of the "taste for leisure." There are two implicit assumptions made in characterizing the utility function by this elasticity. First, we asusme e is constant over all earnings levels for an individual family and secondly, we assume e is constant across families. Both assumptions are somewhat unrealistic. The first implies we are using a local result globally, and the second implies all families have the same shaped indifference curves, although the location may vary.[19/] Any estimate of e then must be interpreted cautiously. We are attempting to characterize families' sensitivity to the tax rate which depends on preferences. Large tax rate effects imply considerable substitution between income and leisure.

Therefore the empirical measure of those tax effects which we are calling e is at least a cousin to the true compensated elasticity of labor supply.

The other economic behavior that affects participation patterns is what many casual analysts refer to as "the welfare trap." In statistical terms, the welfare trap hypothesis posits that the probability of any family receiving a subsidy depends on its past history of subsidy receipts. In particular, previous participation should increase the probability of current participation, ceteris paribus. Some analysts would be tempted to say that once a family receives an initial subsidy from a welfare program, its preferences shift towards leisure. The critical earnings level at which the family switches into program participation would be higher because the family has developed a taste for leisure. Although this taste argument may be appealing to the non-economist, it really has no economic context. As Becker [1976] notes, any behavioral change can be explained in a regime of changing preferences and all economic modelling is wasted effort. Plant [1982] shows that the welfare trap can arise when there are fixed costs associated with the initial receipt of benefits. Alternatively there may be fixed costs of leaving the welfare program and re-entering the workforce. The presence of either type of fixed cost means that once a family is in one state or another, it tends to stay there unless earnings shift so dramatically that it is worth incurring the fixed costs to change states. The notion of state dependence behavior is precisely what is meant by the term "welfare trap." Formally, if we let d_{t-1} equal 1 if the family is a previous participant and -1 if not, the family participates in the welfare program if

$$\ln Y_t \leq \ln \frac{G}{\tau} + \frac{1}{2} e\tau + \left(\frac{d_{t-1}-1}{2}\right) \frac{F}{G} \qquad (3)$$

where F denotes the monetary fixed cost of movement between states. (For a derivation of (3) see Appendix A.) Equation (3) demonstrates that a family who has previously participated is more likely to participate than a previous nonparticipant (ceteris paribus) since the critical earnings level is higher. Our **examination** of the data will allow us to detect any state dependence in participation behavior, but we will not be able to distinsuish among competing explanation of why the state dependence exists. Our goal in this study is to uncover the presence of the behavior indicating a wlefare trap. We cannot determine the reasons for the trap. Equation (3) demonstrates that fixed cost may be one reason, but changing information structures or even changing preferences may also generate state dependent behavior.

Let us recast the simple statistical model of participation behavior developed in Section I by incorporating the complete economic model developed in this section. If earnings are stochastically determined then participation patterns are stochastic. Equation (2) tells us that for a family to participate, the earnings it would have received had there been no welfare program would have to fall below a critical level that is a function of the breakeven level, the tax rate and previous participation behavior. We do not observe these hypothetical earnings that would have occured without the welfare program, but we do observe families' participation behavior. From this behavior we can estimate the parameters that determine participation patterns and thereby guage the effects of welfare programs on families' behavior. In the next section we develop the estimation method and present the basic empirical results of the study. The fourth section presents some alternative econometric specifications and the corresponding results.

Figure 1

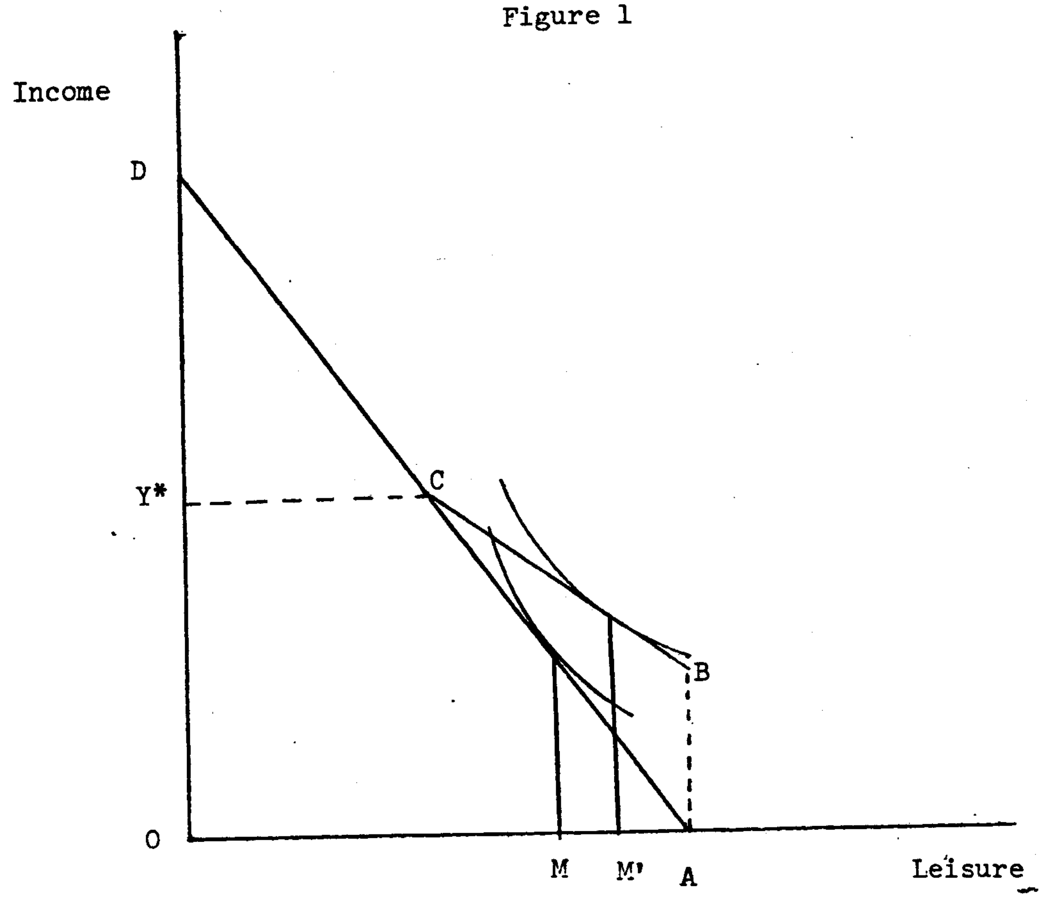
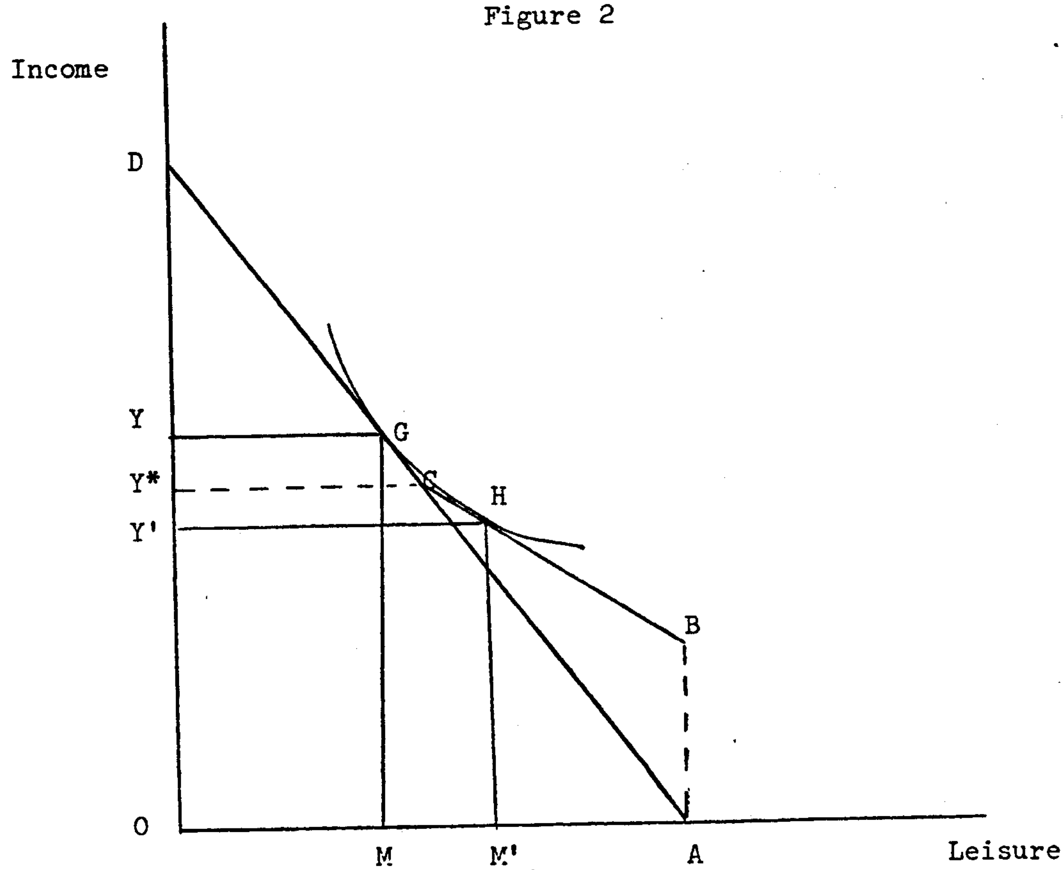


Figure 2



Section III: Empirical Strategy and Results

The economic theory of participation behavior presented in Section II lends itself to a very simple estimation procedure which describes the basic data presented in Table I and allows us to guage the extent of the behavioral effects of the NIT. In this section, we develop that estimation method, as well as the basis for a more involved maximum likelihood estimation (MLE). The results of the simple technique are presented, but the complete exposition of the MLE procedure is contained in Section IV.

Suppose the distribution of the logarithm of family income in T periods, $\tilde{Y}_i$ = $(\ln Y_{i1}, \ln Y_{i2} \ldots \ln Y_{iT})$ in the population is characterized by the T-variate density function f(•). If an NIT program is offered to the population, then the percentage of families participating in any period t is that percentage whose log-incomes would be less than the critical level of earnings in equation (3). The percentage of families following any particular pattern of participation over time, $d_i = (d_{i1}, d_{i2}, \ldots, d_{iT})$ is thus determining the rule implicit in equation (3) for each period. Let Pr (d_i) denote the probability of observing any given family that has participation pattern d_i. Then

(4) $$\Pr(d_i) = \Pr\,[\tilde{Y}_i \circledast d_i \leq (\ln(\frac{G}{\tau}) + \frac{1}{2} e\,\tau + (\frac{d_{t-1}^{-1}}{2})\,\frac{F}{G}) \circledast d_i].$$

This expression is simply the fraction of families who wuld follow pattern d_i given the stochastic nature of earnings. This probability is calculated by taking a multivariate integral of f(•) where the bounds of the integral are the critical values of log-income.[20/] If f(•) were known these integrals could be calculated and a probit-like maximum likelihood estimation of the unknown parameters, e and $\frac{F}{G}$ could be performed. (See Section IV). In this section however we take a simpler approach. In Table I, we compared the actual participation patterns of the experimental families to the predicted patterns of the controls. These predicted patterns were generated assuming that the rule for participation was based only on the breakeven level. In particular a control family was said to be a predicted participant in period t if $\ln Y_t \leq \ln (\frac{G}{\tau})$. In essence we constrained the coefficients on the economic variables in the full economically based criterion to be zero; that is, in Table I we assumed $e = 0$ and $\frac{F}{G} = 0$. We could have calculated predictions using any values of e and $\frac{F}{G}$. The proposed estimators of these parameters are those that provide the closest fit between predicted and actual participation patterns.[21/] Some criterion must be chosen to characterize the closeness of fit. One obvious criterion is the χ^2 statistics summed over all eleven programs. Unfortunately the predicted percentages are often zero for some participation patterns and therefore the χ^2 statistic becomes infinite.[22/] In fact for the white families in Denver, every possible value of e and γ resulted in an infinite χ^2 statistic. Therefore we considered two other measures of closeness of fit: first, the sum of the squared deviations of the actual from the predicted percentages and secondly, the sum of the absolute deviations. Formally, let P_{ij} denote the predicted probability for observing the i^{th} pattern in program j, where $i = 1,8$ and $j = 1,11$. These P_{ij} are determined by applying the participation rule to the control families earnings patterns. Thus the P_{ij} are functions of e and γ. Let n_{ij} be the actual number of experimental families in the j_{th} program observed following pattern i and let n_j be the number of families exposed to the j^{th} program, thus

$n_j = \sum_{i=1}^{8} n_{ij}$. Then choose $\hat{e}$ and $\hat{\gamma}$ to minimize

$$\sum_{j=1}^{11} \sum_{i=1}^{8} \left(P_{ij} - \frac{n_{ij}}{n_j}\right)^2 .$$

Alternatively, $\hat{e}$ and $\hat{\gamma}$ can be chosen to minimize

$$\sum_{j=1}^{11} \sum_{i=1}^{8} \left| P_{ij} - \frac{n_{ij}}{n_j} \right| .$$

We call the former estimates the minimum squared deviation estimates and the latter are dubbed the minimum absolute deviation estimates. The statistical properties of these functions and the resultant estimates are unknown. In particular, the p_{ij} and the n_{ij} are random variables and thus there is a dual source of variation. In usual χ^2 fitting procedures the p_{ij} are taken to be fixed under the null hypothesis. Furthermore, unlike the χ^2 statistic, a deviation of a fixed percentage magnitude is weighted equally no matter what proportion of the sample is in the cell. For example, p_{ij} = 3% and $\frac{n_{ij}}{n_j}$ = 6% receives in equal weight in the criterion function as a p_{ij} = 50% and corresponding $\frac{n_{ij}}{n_j}$ = 53%. In the first case twice as many people as predicted are observed in the ij^{th} cell. In the latter case only 6% more families than predicted to follow pattern i are observed doing so. The former case seems to be a more serious deviation, yet the two deviations receive equal weight. Nevertheless, because of the large number of zero predicted percentages, we cannot remedy this problem.

This estimation procedure is intuitively appealing. We choose some value of e and $\frac{F}{G}$ and calculate the percentage of experimental families we would have expected to participate if indeed e correctly characterized the compensated elasticity of labor supply and $\frac{F}{G}$ were the correct measure of fixed. We then measure the closeness of fit between our predictions and the actual behavior of the experimentals. We choose the values of e and $\frac{F}{G}$ that give us the "best" fit. Under the assumption that the two samples were drawn from the same population, we observe in the control sample how the experimental families would have behaved if there had been no NIT. Thus the controls give us a characterization of earnings patterns without the NIT and basis for comparison of the behavior of families who were eligible for subsidies. This comparison allows us to guage the extent to which behavior is due solely to earnings correlation, how sensitive families are to the implicit NIT wage and how big the welfare trap might be.

The results of the estimation are presented in Table III.[23/] The first column lists the subset of the sample for which the estimation was done. The corresponding number of families is listed in the second column. The third and fourth columns show the result of the minimum squared deviation estimations and the last two columns are the estimates based on the absolute deviations. For example, the white families have an estimated compensated elasticity of labor supply of 11.3% with a very weak state dependence effect of only 1.7%, when the minimum squared deviation criterion is used. When the minimum absolute deviation criterion is used the estimated elasticity increases to 13.9% and the state dependence parameter increases to 2.7%. Whites show the largest sensitivity to marginal tax rates and blacks show the most pronounced state dependence effect. The negative values of e are somewhat disturbing because they imply concave indifference curves. The Chicanos show a negative state dependence effect indicating they move in and

TABLE III

Minimized Deviation Estimates of e and γ

Sample	Number of Observations	Minimum Square Estimates		Minimum Absolute Value Estimates	
		e	γ	e	γ
Whites	958	0.113	0.017	0.139	0.027
Blacks	646	-0.181	0.022	-0.022	0.078
Chicanos	396	-0.313	-0.147	-0.092	-0.132
White Families in Seattle	445	-0.128	-0.162	-0.127	-0.134
White Families in Denver	513	0.223	0.120	0.226	0.096
Three Year White Families	776	0.083	0.006	0.092	0.036
Five Year White Families	549	0.027	0.029	0.207	0.001

out of subsidization with ease. The white families in Denver show the largest behavioral reactions to the program. Families on the long exposure program (five years) have a smaller value of the elasticity and a larger state dependence effect. The point estimates presented in Table III describe the behavior one would conjecture on the basis of Table I alone: a small sensitivity to tax rates indicating an underprediction of participation and a small welfare trap leading to more persistence than would otherwise be predicted. The conclusion is that the important reason for persistent participation is correlation in earnings.

In order to guage the precision of these minimum deviation estimates standard errors were computed for the white families using a jackknife procedure described by Mostellar and Tukey (1977).[24/] The experimental sample was broken into 20 groups and the pseudovalues were computed by deleting each twentieth of the sample in turn. The estimated variance is the sample variance of those pseudovalues.[25/] This procedure also yields jackknifed estimates of the parameters of interest. These estimates using the minimum squared deviation criterion are $\hat{e}_J = 0.253$ and $\hat{\sigma}_J = 0.068$ where the subscripts denote jackknifing. The corresponding standard errors are $S_{\hat{e}} = 0.082$ and $S_{\hat{\sigma}} = 0.041$. The jackknife procedure indicates that the actual point estimates in Table III are sensitive to outlying observations and the precision of estimation is not great. Again this squares with the conclusions one might intuit from an examination of Table I.

In this section, we have presented a series of estimates based on a minimum χ^2-like criteria. The estimates of the key behavioral parameters in our model of participation behavior lead us to the conclusion that earnings correlation is the culprit in the persistent participation of families in these NIT programs. There is a fair amount of evidence for a small positive compensated elasticity of labor supply but the "welfare trap" does not seem to be an empirically compelling explanation for persistent participation.

Section IV: Extended Empirical Results

The estimation procedure implemented in Section III made no assumption regarding the functional form of the distribution of earnings. The estimation procedure relied instead on a non-parametric test of the goodness of fit between two discrete distributions. In this section, we explain a maximum likelihood procedure for estimating the parameters of the participation model. We present estimates from that model and make some critical remarks regarding this whole approach to estimating discrete choice behavior.

The maximum likelihood approach to estimation is a straightforward extension of the theory developed in the previous sections. The likelihood of observing any particular pattern of participation is given in Section III in equation (4) which is reproduced here for convenience:

$$(4) \quad \Pr(d_i) = \Pr[\tilde{Y}_i \otimes d_i \leq (\ln(\frac{G}{\tau}) + \frac{1}{2} e\tau + (\frac{d_{t-1}^{-1}}{2}) \frac{F}{G}) \otimes d_i].$$

If the multivariate distribution of income were known, this probability would be a multivariate integral of the appropriate density function. In T periods, there are 2^T possible patterns of participation behavior. Let the 2^T corresponding probabilities be denoted as P_j, $j = 1, 2, \ldots 2^T$. Let the number of families observed following pattern j be denoted as N_j. Then the likelihood of observing the experimental sample is

$$(5) \qquad \mathcal{L}_E = \prod_{j=1}^{2^T} (P_j)^{N_j} \quad .$$

Let the multivariate density function of the logarithm of earnings be denoted as $f(\cdot \mid \mu, \Sigma)$ where u and Σ are unknown parameters of the density. Equation (5) then gives us the basic framework for estimation. Given the observed participation pattern for each experimental family, maximum likelihood estimates of μ, Σ, e and $\frac{F}{G}$ can in principle be obtained. If $f(\cdot \mid \mu, \Sigma)$ is Normal this is simply a multivariate probit computation where the bounds of integration of the multivariate distribution depend on parameters being estimated. This estimation, however, does not use all the information in the sample. We also have observations on families who were not exposed to the NIT (the control subsample). Assuming controls and experimentals were initially chosen from the same population we know that the earnings patterns are characterized by the distribution $f(\cdot|\mu,\Sigma)$. The likelihood of observing the control sample of N_c families with earnings patterns Y_i is

$$\mathcal{L}_c = \prod_{i=1}^{N_c} f(\ln Y_i|\mu,\Sigma).$$

Thus, the likelihood of the entire sample is

$$= \mathcal{L}_E \times \mathcal{L}_c$$

$$= \left[\prod_{J=1}^{2^T} (P_j)^{N_j}\right]\left[\prod_{i=1}^{N_c} f(\ln Y_i|\mu,\Sigma).\right]$$

Intuitively, we include the controls to help parameterize the earnings distribution.

There are several ways in which the specification can be generalized to make it more intuitively appealing. It seems unreasonable to restrict mean income in a given period to be constant across families. We can modify the likelihood function by replacing μ by a vector of linear combinations of of family dependent characteristics. That is, let $\mu_{it} = X_{it}'\beta_t$ where X is a vector of characteristics of family i in period t and β_t is a vector of coefficients on X_{it}. Thus

$$\mu_i = (X_{i1}\beta_1 \; X_{i2}\beta_2 \; X_{i3}\beta_3 \cdot \cdot \cdot X_{iT}\beta_T).$$

The X vector might contain such characteristics as family size, race, or location, and would certainly contain a time varying constant. For the experimental families the normal income assignment variables were also included. The coefficients on these characteristics can be constrained to be the same across periods if convenient. In this modified likelihood function the parameters being estimated are $\beta_1, \ldots, \beta_T$, Σ, e and $\frac{F}{G}$.

Table IV presents the important maximum likelihood parameter estimates for various subsets of the sample. (In Appendix B, a detailed explanation of the estimation procedure, is given along with a full set of parameter estimates. Further information is contained in Plant (1982)). The estimates in Table IV are startling in light of those calculated in Section III. For example, the first line of Table IV shows that for white families the maximum likelihood compensated elasticity is estimated at 0.968, as opposed to 0.113 using minimum squared deviations. The state dependence effect increases for white families from 1.7% to 35.2%. The maximum likelihood estimate would lead us to the conclusion that these families are showing a strong reaction to the marginal tax rates and there is a pronounced welfare trap, which is contrary to what we concluded in Section III. Which set of estimates are we to believe?

The maximum likelihood estimates depend on a specific functional form of the density function, $f(\cdot|\mu,\Sigma)$. The assumption of the lognormality of earnings is arbitrary and was made to allow an arbitrary covariance structure

Table IV

Labor Supply and State Dependence Results for Various Demographic Groups (Asymptotic t-statistics in parentheses)

Demographic Group	Number of Observations	Compensated Elasticity of Labor Supply	State Dependence Parameter
Whites	958	0.968 (4.56)	0.352 (5.50)
Blacks	646	0.284 (1.22)	0.432 (5.86)
Chicanos	396	0.429 (1.27)	0.080 (0.82)
Denver residents-White families	513	1.367 (4.15)	0.561 (5.47)
Seattle residents-White families	445	0.434 (1.75)	0.069 (0.97)
Three year white experimental and all controls	776	0.900 (3.62)	0.392 (5.25)
Five year white experimental and all controls	549	1.219 (3.71)	0.343 (3.14)

and for computational convenience. Using methods devised by White (1979), the normality of the log-earnings structure was tested accounting for the truncation of the sample on the basis of pre-experimental log-earnings. (See Plant (1982)). Testing this assumption using two different methods the log-earnings failed the normality tests at the 90% confidence level. White indicates that if the test is at all biased it is in favor of the normality assumption. Other estimates were computed using alternative distributional assumptions. Using the logistic function, the compensated elasticity was found to be 0.485 for white families and significantly different from zero; the state dependence parameter was 0.155. (See Plant, 1982). The maximum likelihood estimates from a multivariate discrete choice model are very sensitive to distributional assumptions. In this instance, specifying a particular functional form has radically altered our conclusion. Because we have experimental data, we can check our distribution assumption which is a luxury not given to many economic investigators. The White tests lead us to reject normality, but do not lead to any reasonable alternative specification. The search for the correct maximum likelihood estimates become a very expensive guessing game.

To check the explanatory power of the ML estimates in Table IV, we recomputed the predicted participation patterns for the controls using the point estimates for white families, e = 0.968 and $\frac{F}{G} = 0.352$. These are presented in Table V and put to rest any suspicion we have that perhaps these parameter estimates are the best we can do. The predictions generated using these estimates do not mimic actual participation patterns at all. The large behavioral parameters result in considerable overprediction of participation.

TABLE V

Predicted and Actual Participation Patterns of Families in the Seattle/Denver Income Maintenance Programs: Minimum Squared Deviation Estimates
(=0.11 , γ=0.014)

Breakeven in Dollars

Participation Pattern Period 3 2 1	5430 Pred. %	5430 Act. %	5801 Pred. %	5801 Act. %	6650 Pred. %	6650 Act. %	7366 Pred. %	7366 Act. %	7600 Pred. %	7600 Act. %	6000* Pred. %	6000* Act. %	6000** Pred. %	6000** Act. %	9600 Pred. %	9600 Act. %	10343 Pred. %	10343 Act. %	11200 Pred. %	11200 Act. %	12000 Pred. %	12000 Act. %
0 0 0	57.5	66.7	42.6	38.7	48.3	32.4	24.4	21.1	11.0	20.6	32.8	34.4	27.2	25.6	22.1	17.4	16.5	22.4	12.1	12.5	7.8	11.6
0 0 1	10.7	13.3	20.3	6.5	12.1	20.3	19.9	7.9	15.3	17.6	15.2	9.8	16.6	20.9	12.8	8.7	10.9	9.2	7.5	6.3	5.1	10.1
0 1 0	1.9	0.0	2.6	3.2	2.3	4.1	2.3	0.0	1.5	5.9	1.8	4.9	1.3	7.0	1.8	2.9	1.4	3.9	2.2	0.0	0.3	1.4
0 1 1	4.4	6.7	4.2	6.5	5.4	4.1	6.7	18.4	5.7	8.8	6.6	4.9	5.7	11.6	7.1	7.2	6.7	5.3	7.9	6.3	7.1	15.9
1 0 0	6.0	3.3	5.1	9.7	5.6	6.8	3.7	5.3	3.7	5.9	5.1	3.3	4.5	4.7	2.6	4.3	3.8	2.6	2.7	6.3	2.0	7.2
1 0 1	2.1	3.3	1.7	3.2	4.2	2.7	6.6	2.6	4.1	5.9	3.8	1.6	4.9	7.0	4.4	1.4	4.0	1.3	4.1	0.0	2.7	2.9
1 1 0	3.6	3.3	2.5	6.5	3.0	5.4	5.7	7.9	4.1	4.4	2.2	6.6	1.9	0.0	3.8	13.0	2.9	5.3	2.1	3.1	2.2	2.9
1 1 1	14.0	3.3	21.0	25.8	19.1	24.3	30.7	36.8	33.1	30.9	32.6	34.4	37.6	23.3	43.9	44.9	53.8	50.0	61.3	65.6	72.8	47.8
Number of Observations	30		31		74		38		68		61		43		69		76		32		69	
S.S.E.	19.97		26.41		196.69		51.66		71.47		27.69		55.54		65.87		42.96		6.01		366.90	

*This program does not have a declining tax rate (G = 5600, t = 0.7)

**This program does have a declining tax rate (G = 4800, t = 0.8, rate of decline = 0.00005)

In essence the non-normality of log-earnings is being translated into inaccurate estimates of the behavioral parameters.

The lesson of this section is important. Using the widely touted very costly state-of-the-art techniques for analyzing discrete decisions over time, the parameter estimates derived do not result in good predictions of actual behavior. This is because the underlying distributional assumption is wrong. The simple estimation procedure in Section III leads to estimates that explain observed behavior well. By letting the control data characterize the stochastic process generating the experimental behavior, we impose no mathematical constraints on the estimates of the parameters of the model. Such results are a clear warning that estimates of the parameters of models of discrete choice which depend heavily on distributional assumptions should be suspect unless some clear empirical justification for the assumption is given. We reject the estimates in Table IV because they lead to bad predictions. Their complex derivation does not justify their acceptance. Maximum likelihood is not magic. It, too, must be evaluated in light of how well it describes the observed data.

Section V: Conclusion

In attempting to explain participation in welfare programs, this study has made several important observations, both empirical and methodological. In this section we review those observations.

First, we noted that participation in NIT programs is determined by the pattern of earnings in the population. Persistent participation will be observed among families if earnings are highly correlated. We formulated a model that accounted for this reason for persistent participation, but also incorporated potential economic reaction to the subsidy program. We noted that there was the potential for substitution into leisure because of the implicit wage tax and that there might be a state dependence effect that leads to increased participation over time. This is the first interesting observation. The idea of state dependence has been discussed widely in the econometric and labor economics literature. To the best of my knowledge, this is the first study where the existence of state dependence is an important question for policy purposes. If there is state dependence, that is if a welfare trap does exist, it means that economic incentives to work are steadily declining because of welfare programs. We would conclude that the government should limit the length of families' eligibility so they do not become entrapped into dependence on the government's generosity. If there is no welfare trap and we are concerned with persistence in participation we must look to other remedies for the problem.

The important empirical finding is that persistence is explained by correlation in earnings and the evidence pointing towards a welfare trap is at best weak. There is some indication that there is tax rate sensitivity, but the compensated elasticities are imprecisely estimated. The magnitudes of the elasticities are similar to those found in other studies. (for example, Ashenfelter, 1980). We should worry about the social structures that entrap the poor into persistent low earnings, rather making short-term disincentives to work provided by welfare the culprit.

Finally, for the empirical labor economist, an important methodological point has been made. We contrasted a simple econometric method for estimating a model of discrete choice to one which relies on a vast array of econometric and computational machinery. The simple method explained the data well. The complex model did not. This is not an invective against complex econometrics. Instead we need to consider carefully the methods we use, the implicit assumptions contained in the implementation of those methods and the ability of the resultant model to describe the observed behavior. In this case, the multinomial probit model did violence to the data. Acceptance of its estimates just because they were the maximum likelihood estimates would have lead us astray.

This study has been a first attempt at explaining welfare participation behavior. Further elaboration is contained in Plant (1982), but two avenues for future research should be clear. First, a full model of labor force behavior should be incorporated into the model of discrete behavior. In no way do we incorporate the extent of the family's substitution into leisure the estimation procedure. Secondly the econometric methodology must be studied more carefully if future researchers are to handle problems like this in a careful and correct way.

FOOTNOTES

1 For example, AFDC may encourage welfare mothers to have more children. These long-term effects of welfare will not be the focus of this paper.

2 The imposition of an experimental welfare program may also result in intertemporal substitution of labor and leisure. If indeed a program is known to be temporary and if the net wage to the welfare family has decreased, the family may substitute into current consumption of leisure. This may result in an observed persistence solely due to the experimental nature of the program.

3 The breakeven level is the income at which the subsidy, D, is equal to zero. Since, in the simple NIT program, $D = G - \tau Y$ where G is the guaranteed income level, τ the implicit tax rate and Y the income, $D = 0$ if $Y = \frac{G}{\tau}$.

4 The breakeven level is the income at which the subsidy, D, is equal to zero. In the simple NIT program $D = G - \tau Y$ where G is the guaranteed income level and τ is the implicit tax rate on earnings, and Y is earnings. Obviously, $D = 0$ if and only if $Y = \frac{G}{\tau}$. Therefore the breakeven level is $\frac{G}{\tau}$.

5 We can make the distribution of Y_{it} dependent on any characteristics of family i at time t, so this specification can incorporate any type of earnings determination function.

6 We assume that G and τ are not time varying parameter. With some notational inconvenience this assumption is easily relaxed.

7 Let $a = (a_1 \cdot \cdot \cdot a_n)$ and $b = (b_1 \cdot \cdot \cdot b_n)$ then $a \circledast b = (a_1 \cdot b_1, a_2 \cdot b_2, \ldots a_n \cdot b_n)$.

8 As an example suppose $d_i = (-1, 1, -1)$ then $Pr(d_i) = Pr[(-1,1,-1)]$

$$= Pr\left[Y_{i1}(-1) \leq \frac{G}{\tau}(-1),\ Y_{i2} \leq \frac{G}{\tau},\ Y_{i3}(-1) \leq \frac{G}{\tau}(-1)\right]$$

$$= Pr\left[Y_{i1} \geq \frac{G}{\tau},\ Y_{i2} \leq \frac{G}{\tau},\ Y_{i3} \geq \frac{G}{\tau}\right]$$

$$= \int_{\frac{G}{\tau}}^{\infty} \int_{-\infty}^{\frac{G}{\tau}} \int_{\frac{G}{\tau}}^{\infty} f(y_1, y_2, y_3)\, dy_1\, dy_2\, dy_3$$

where $f(\cdot)$ is a trivariate density function.

9 The experimental design was sufficiently complicated to cast doubt on whether the data gathered are adequate for such analysis. Further comment will be made in a later section.

10 Although more families were enrolled initially, we have restricted our sample in order to ensure a complete earnings history. To be included in the control sample, the family had to have a complete earnings history for all four years. Some families were deleted from the data set because their pre-experimental income exceeded the established truncation points. See later sections for more details.

11 Each family was paid $20 per month just for reporting their income and employment status. Thus no welfare subsidy is being given unless the payment in any month exceeds $20.

12 The implicit assumption is that families who left the program did so because it was not beneficial to continue their eligibility for payments. Thus, they are behaving like non-participants.

13 Our analysis is restricted to the first three periods in this essay.

14 It is important to note that the predicted percentages were computed by a somewhat more complicated procedure than previously indicated. Assignment of experimental families to the eleven different NIT programs was not made randomly. Each family was assigned to a normal income category, of which there were seven for experimental families. These categories were $0-$999, $1000-$2999, $3000-$4999, $5000-$6999, $7000-$8000, $9000-$10999 and denoted as E1 to E6. E7 was a category for control families with normal income of $11999 to $13000. E0 is a residual category for families not otherwise assigned. For a detailed description of the assignment process see Keeley and Robins (1980). Those with low normal incomes tended to be concentrated in the programs with small breakevens. Simple predictions of the experimental sample's behavior from the control sample would not take into account this non-random assignment. Since control families were also assigned to normal income levels, predictions of participation patterns were made for families in each normal level and then a weighted average of those prediction was taken, the weights reflecting the distribution of experimental families over normal income levels for each NIT program.

15 The statistic computed is only distributed as X_7^2 asymptotically. A rule of thumb is that this approximation is sufficient if the predicated number of observations in each cell is at least 1.5. (Degroot, 1975).

16 This exposition relies heavily on Ashenfelter (1978) and Ashenfelter (1980).

17 In Figure 1 if the family consumed leisure OM as anticipated the subsidy would be the distance from ACD to BCD at OM. Since the family actually chooses more leisure the transfer is the distance from ACD to BCD at M' which is clearly greater.

18 The slope of the budget constraint below the breakeven Y* is equal to $w(1-\tau)$ where w is the wage rate. The higher the tax rate, τ, the flatter the below breakeven portion of the budget constraint.

19 In Plant [1982] the second assumption is relaxed by parameterizing e as a function of family specific characteristics.

20 For example, assume for notational simplicity that $\frac{F}{G} = 0$ and thus there is no state dependence. Let $d_i = (1, -1, 1)$:

$$\Pr[d_i] = \Pr(1, -1, 1)$$

$$= pr[\ln Y_1 \leq \ln \frac{G}{\tau} + 1/2\ e\ \tau, \ln Y_2 \geq \ln \frac{G}{\tau} + 1/2\ e\ \tau, \ln Y_3 \leq \ln \frac{G}{\tau} + 1/2\ e\ \tau]$$

$$= \int_{-\infty}^{\ln \frac{G}{\tau} + 1/2\ e\ \tau} \int_{\ln \frac{G}{\tau} + 1/2\ e\tau}^{\infty} \int_{-\infty}^{\ln \frac{G}{\tau} + 1/2\ e\ \tau} f(\cdot) d\ln Y_1\ d\ln Y_2\ d\ln Y_3$$

where $f(\cdot)$ denotes the trivariate distribution of log income.

21 I am grateful to Hal White for initially suggesting this general method.

22 $\chi^2 = \Sigma \frac{(np_i - n_i)^2}{np_i}$ where p_i = predicted probability for pattern i, n = number in sample, n_i = actual number following pattern i. Thus is $p_i = 0$, $\chi^2 = \infty$.

23 The estimates were calculated using the simplex search method and a method due to Powell that does not rely on a first derivatives of the criterion function. The simplex method is described by Nelder and Mead (1964) and the Powell method in Powell (1964). The actual routines used are contained in the maximization package GQOPT.

These criteria functions were not easy to minimize because sufficiently small changes in e and $\frac{F}{G}$ would result in no change in functional value because

23 (cont'd)

there would be no change in predicted behavior. Therefore the algorithm tended to get "stuck" in areas. To remedy this problem, starting values for the search algorithm were found using a grid search over the intervals -1 to +1 for both e and $\frac{F}{G}$. The contours of the function were sufficiently regular that any possibility of a minimum outside this area could be detected.

24 Similar estimates for other demographic groups have not yet been made due to the expense of estimation.

25 For any coefficient β_1, the parameter estimates is the average of five pseudovalues computed from the maximum likelihood routines:

$$\beta_i = \frac{1}{5} \sum_{j=1}^{5} \beta^*_{ij}$$

where β^*_{ij} is the j^{th} pseudovalue of the i^{th} coefficient. Furthermore the standard error, s_* is computed as

$$s^2_* = \frac{s^2}{5}$$

where

$$s^2 = \frac{\Sigma \beta_{ij}{}^2 - \frac{1}{5} (\Sigma \beta^*_{ij})^2}{4}$$

See Mosteller and Tukey (1977).

Appendix A: Proof of Equation (3).

Suppose the family has a utility function $U(\ell,c)$ where ℓ denote leisure and a consumption which a corresponding expenditure function $E(w,v)$. Denote the fixed cost of entrance as F. The family's problem is to maximize $U(T-h,Y)$ where h = number of hours of work and T = total available time. The following constraints are imposed on the maximization:

(a) If the family has participated in the previous period(s) then

$$Y = \begin{cases} (wh + Y^{o}) + G - \tau(wh + Y^{o}) & \text{if } G - \tau(wh + Y^{o}) \geq 0 \\ wh + Y^{o} & \text{if } G - \tau(wh + Y^{o}) < 0 \end{cases}$$

(b) but, if the family has not participated in the previous periods(s) then

$$Y = \begin{cases} (wh + Y^{o}) + G - \tau(wh + Y^{o}) - F & \text{if } G - \tau(wh + Y^{o}) \geq 0 \\ wh + Y^{o} & \text{if } G - \tau(wh + Y^{o}) < 0. \end{cases}$$

The prior non participant will choose to participate if

$E(w,v) - E(w(1-\tau),v) \leq G - F$, that is if the lump sum minus the fixed costs exceeds the value of utility lost from the wage reduction. Following Ashenfelter (1980), this means the family participates if

$$\ell n(wh) \leq \ell n \left(\frac{G-F}{\tau}\right) + \frac{1}{2} e\tau \quad \text{a/}$$

$$\ell n\ Y \leq \ell n \left[\frac{G}{\tau}\left(1 - \frac{F}{G}\right)\right] + \frac{1}{2} e\tau$$

$$\ell n\ Y \leq \ell n\ \frac{G}{\tau}\left(1 - \frac{F}{G}\right) + \frac{1}{2} e\tau$$

and using the approximation that $\ell n\ (1 + z) \approx z$ the participation rule becomes

$$\ell n\ Y \leq \ell n\ \frac{G}{\tau} - \frac{F}{G} + \frac{1}{2} e\tau$$

for those families who have not participated previously. Let d_{t-1} equal 1 if the family participated previously and -1 if not. Then the participation rule is

$$\ln Y \leq \ln \frac{G}{\tau} + \left(\frac{d_{t-1}-1}{2}\right) \frac{F}{G} + \frac{1}{2} e\tau \quad .$$

a/ This expression comes from taking a second order approximation to the expenditure function.

Appendix B: Maximum Likelihood Estimation

The maximum likelihood estimates of the model in Section III were computed using the likelihood function described on page 23. Because of the method of sample selection, some additional elaboration is required. In particular, families with normal income, as determined by program administrators, exceeding $13,000 in the year preceeding the experiment were deleted from the sample. Furthermore, families whose income was between $11,000 and $13,000 were included only in the control sample. Plant (1982) shows that we can account for this truncation by including pre-experimental log-earnings in the vector of variables that forms the mean of the distribution $f(\cdot)$. The resulting estimates of the parameters of the density function are conditional means and variances.

Assignment of the experimental families to the eleven different NIT programs tested was made on the basis of the normal income category to which families were assigned. (See Keeley and Robins, (1980)). Therefore it is necessary to make the mean conditional on dummy variables indicating the family's normal income category.

All computations use an earnings measure that is normalized for family size. The NIT programs are parameterized for a family of four, and then adjustment is made for larger or smaller families.(See Keeley and Robins (1980)). Since the estimation is based on normalized income, it is unnecessary to include a measure of family size in the conditional mean.

The multivariate normal integrals were computed by using the Clark approximation proposed by Daganzo et al (1977). Plant (1982) criticizes this approximation but cannot find any reasonable alternative.

Table B.1 gives the results of the full maximum likelihood estimation for white families. The full set of results for other demographic groups can be found in Plant (1982).

Column 1 contains the estimates of the stochastic earnings function using the control sample only. The likelihood function is simply that of a multivariate regression with the coefficients on the normal income dummies constrained to be the same in all three periods. The standard deviation of log-earnings is precisely estimated in all three periods. There is evidence of non-stationarity: the standard deviation decreases from 0.840 in the first period to 0.777 in the second and increases to 0.932 in the third year. The covariances between log-earnings seem small, but recall these are the covariances conditional on pre-experimental log-earnings. The coefficients on pre-experimental log-earnings decrease over time as would be expected. The coefficients on the normal income categories are not surprising. The omitted category is E5, ($7,000-$8999) and thus the negative signs on the category dummies for normal income less (greater) than $7000 (9000) are negative (positive). Only the E2 dummy is significantly different from zero. The conclusion to be drawn from this first estimation is that there are no striking anomalies in the data, and the stochastic earnings structure seems to be quite well estimated.

The remaining estimations use the entire sample of families-controls and experimentals. Recall that the likelihood function for these estimations is a multivariate regression likelihood function for the controls multiplied by a trivariate probit for the experimental families. Unlike the usual trivariate probit scheme, the variance-covariance structure and the time varying mean are identified because of the presence of the control sample. Without the sample, this detailed model would not be fully identified. Therefore, a separate estimation for experimental families alone is not presented.

The second column of Table III presents the estimates of the model using the entire sample, but not including any potential for an economic reaction to the NIT by the experimental families. Overall, the stochastic earnings function estimated for the entire sample seems much the same as that for the controls.

The set of estimates in the third column account for the possibility of an economic response to the NIT. The estimation differs from that in column 2 by including the compensated elasticity term in the probit integral bounds. The estimated compensated elasticity of labor supply is 0.191 and has an asymptotic t-statistic of 1.20. The remaining parametive estimates differ vary little from those in column 2. This elasticity is quite high, but it is so imprecisely estimated there seems to be no reason to conclude there is a strong response to the implicit wage tax.

The results presented in the fourth column of Table III include the state dependence parameter in the model, but exclude the compensated labor supply elasticity. The state dependence parameter is positive and significant and, since earnings are measured in logarithms, indicates that the critical earnings level is about 15% lower for previous nonparticipants than for previous participants.

In column 5 we present the estimates of the full model that includes the state dependence term in the probit limits of integration and the compensated elasticity term (see equation (15)). The estimates of the variances and covariances are closer to those for the control sample alone. The elasticity of labor supply is 0.968 which is extremely large and is significantly different from zero. The state dependence parameter is large and significantly different from zero as well. This parameter implies earnings must be 35% lower to induce a family to participate for the first time than they need be to induce participation in subsequent years. Further discussion and variants on this basic model can be found in Plant (1982).

TABLE B-1

Maximum Likelihood Parameter Estimates of Participation Behavior Model for White Families (Asymptotic t-statistics in Parentheses)

	(1)	(2)	(3)	(4)	(5)
Standard Deviation of Log Earnings:					
Period 1	0.840 (26.5)	0.897 (25.2)	0.899 (25.1)	0.892 (26.0)	0.879 (25.9)
Period 2	0.777 (26.2)	0.874 (24.1)	0.874 (24.0)	0.857 (25.2)	0.842 (25.0)
Period 3	0.932 (26.5)	1.031 (24.8)	1.033 (24.7)	1.007 (25.2)	0.991 (24.3)
Correlation Coefficient of Log Earnings					
Period 1 & 2	0.320 (6.77)	0.517 (14.7)	0.518 (14.4)	0.497 (14.0)	0.464 (12.3)
Period 1 & 3	0.155 (3.06)	0.388 (9.50)	0.391 (9.19)	0.369 (9.14)	0.346 (8.73)
Period 2 & 3	0.363 (8.07)	0.581 (19.2)	0.582 (18.7)	0.557 (17.0)	0.530 (16.1)
Compensated Elasticity of Labor Supply			0.191 (1.20)		0.968 (4.56)
State Dependence Parameter - one lag				0.146 (3.06)	0.352 (5.50)
Coefficients on Preexperimental Log Earnings					
Period 1	0.752 (14.2)	0.766 (15.8)	0.764 (15.5)	0.769 (15.8)	0.761 (15.9)
Period 2	0.501 (10.2)	0.515 (11.6)	0.512 (11.4)	0.509 (11.6)	0.484 (10.9)
Period 3	0.436 (7.60)	0.438 (8.41)	0.434 (8.25)	0.433 (8.34)	0.409 (7.49)
Constant					
Period 1	2.247 (4.75)	2.103 (4.85)	2.126 (4.85)	2.049 (4.71)	2.129 (4.96)
Period 2	4.609 (10.7)	4.465 (11.3)	4.495 (11.2)	4.514 (11.5)	4.738 (11.9)
Period 3	5.178 (10.1)	5.156 (11.1)	5.189 (11.1)	5.194 (11.3)	5.417 (11.8)
Coefficient on Normal Earnings Category					
E0	0.189 (0.86)	0.150 (0.65)	0.147 (0.63)	0.157 (0.69)	0.152 (0.66)
E1	-- —	0.015 (0.024)	0.052 (0.07)	-0.032 (0.05)	0.093 (0.15)
E2	-0.305 (3.02)	-0.314 (3.20)	-0.311 (3.15)	-0.312 (3.22)	-0.305 (3.17)
E3	-0.012 (0.21)	-0.039 (0.70)	-0.033 (0.59)	-0.049 (0.90)	-0.033 (0.60)
E4	-0.071 (1.61)	-0.035 (0.83)	-0.030 (0.69)	-0.046 (1.09)	-0.032 (0.75)
E6	0.059 (1.35)	0.058 (1.32)	0.057 (1.30)	0.059 (1.34)	0.055 (1.26)
E7	0.068 (0.52)	0.082 (0.59)	0.077 (0.55)	0.095 (1.34)	0.087 (0.64)
Log of Likelihood Function	-1333.57	-2324.93	-2324.18	-2320.32	-2309.90
Number of Observations	367	958	958	958	958

REFERENCES

Ashenfelter, Orley, "Discrete Choice in Labor Supply: The Determinants of Participation in the Seattle-Denver Income Maintenance Experiments." Industrial Relations Section Working Paper No. 136, Princeton University, March 1980.

__________, "The Labor Supply Response of Wage Earners," in Welfare in Rural Areas: The North Carolina-Iowa Income Maintenance Experiment, John Palmer and J.A. Pechman, eds., Brookings Institution, Washington, D.C., 1978, pp. 109-38.

Becker, Gary S., "The Economic Approach to Human Behavior, (Chicago, 1976).

Burtless, Gary and Hausman, Jerry A., "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment," Journal of Political Economy, Vol. 86, No. 6, 1978, pp. 1103-1130.

Clark, Charles E., "The Greatest of a Finite Set of Random Variables," Operations Research, March-April 1961, pp. 145-162.

Daganzo, Carlos F., Bouthelier, F. and Sheffi, Y., "Multinomial Probit and Qualitative Choice: A Computationally Efficient Algorithm," Transportation Science, Vol. 11, No. 4, November 1977, pp. 338-358.

Dutt, J.E., "Numerical Aspects of Multivariate Normal Probabilities in Econometric Models," Annals of Economic and Social Measurement, Vol. 5, No. 4, 1976, pp. 547-561.

Flinn, C.J. and Heckman, J.J., "Models for the Analysis of Labor Force Dynamics," Unpublished Working Paper, June 1980.

Goldfeld, S.M. and Quandt, R.E., "Econometric Modelling with Nonnormal Disturbances," Econometric Research Program Research Memorandum No. 265, Princeton University, June 1980.

__________, and __________, Nonlinear Methods in Econometrics, North Holland Publishing Co., Amsterdam, 1972.

Harrington, Michael, The Other America: Poverty in the United States, New York, Macmillan, 1962.

Hart, P.E., "The Dynamics of Earnings," The Economic Journal, September 1976, Vol. 86, pp. 551-565.

Hausman, Jerry A., and Wise, David A., "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," Econometrica 46 (1978), pp. 403-426.

Heckman, James J., "Statistical Models for Discrete Panel Data," 1980, Unpublished paper.

__________, and Willis R., "A Beta Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," Journal of Political Economy, 85, pp. 27-58.

Horowitz, J., Sparman, J.M., and Daganzo, C.F., "An Investigation of the Accuracy of the Clark Approximation for the Multinomial Probit Model," Unpublished Working Paper, March 1981.

Johnson, N.L. and Kotz, S., Distributions in Statistics: Continuous Multivariate Distributions, John Wiley & Sons Publishing Co., New York, 1972.

Keeley, Michael, C. and Robins, Philip D., "Experimental Design, the Conlisk-Watts Assignment Model, and the Proper Estimation of Behavioral Response," Journal of Human Resources, 15 (1980), pp. 480-98.

Lillard, L. and Weiss, Y., "Components of Variation in Panel Earnings Data: American Scientists 1960-70," Econometrica, 47, pp. 437-454.

__________, and Willis, Robert J., "Dynamic Aspects of Earnings Mobility," Econometrica, Vol. 46, No. 5, September 1978, pp. 985-1012.

Moffitt, Robert, "The Effect of a Negative Income Tax on Work Effort: A Summary of the Experimental Results," April 1980.

__________, "Limited-Duration NIT Experiments and Life Cycle Labor Supply," Rutgers University September 1980.

Mostellar, F. and Tukey, J.W., Data Analysis and Regression, Addison-Wesley Publishing Co., Reading, Massachusetts, 1977.

Nelder, J.A. and Mead, R., "A Simplex Method for Minimization," Computer Journal, 7 (1964), pp. 308-313.

Owen, D.B., "Tables for Computing Bivariate Normal Probabilities." Annals of Mathematical Statistics, Volume 27, (1956), pp. 1075-1090.

Parks, J., "Work Incentive and the Negative Income Tax," Unpublished Paper, 1979.

Pearson, E.S., D'Agostino, R.B., and Bowman, K.O., "Tests for Departure from Normality: Comparison of Powers," Biometrika, 64 (1977), pp. 231-246.

Plant, Mark W., "A Note on the Trivariate Probit Model as a One-Factor Scheme," Unpublished Paper, September 1980.

Plant, Mark W., Participation Over Time in the Seattle-Denver Maintenance Programs, Ph.D. Dissertation, Princeton University, January 1982.

Powell, M.J.D., An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," Computer Journal, 7 (1964), pp. 155-162.

Rees, Albert, "An Overview of the Labor Supply Results," The Journal of Human Resources, Spring 1974, Vol. 9, No. 2, pp. 158-80.

Robins, Philip K. and West, Richard W., "Labor Supply Response Over Time," Journal of Human Resources, 15 (1980), pp. 524-44.

__________, and West, Richard W., "Program Participation and Labor Supply Response," Journal of Human Resources, 15 (1980), pp. 499-523.

Weiss, Y. and Lillard, L. "Experience, Vintage and Time Effects in the Growth of Earnings: American Scientists, 1960-70," Journal of Political Economy, 86, pp. 427-447.

White, Halbert, "Specification Tests for the Tobit Model," Discussion Paper 79-22, University of California, San Diego, 1979.

__________, and Olson, Lawrence, "Conditional Distributions of Earnings, Wages and Hours for Blacks and Whites," Unpublished paper, 1981.