



CHICAGO JOURNALS



The Society of Labor Economists

NORC at the University of Chicago

The University of Chicago

---

Nonparametric Estimates of the Labor-Supply Effects of Negative Income Tax Programs

Author(s): Orley Ashenfelter and Mark W. Plant

Source: *Journal of Labor Economics*, Vol. 8, No. 1, Part 2: Essays in Honor of Albert Rees (Jan., 1990), pp. S396-S415

Published by: [The University of Chicago Press](#) on behalf of the [Society of Labor Economists](#) and the [NORC at the University of Chicago](#)

Stable URL: <http://www.jstor.org/stable/2535218>

Accessed: 17/01/2015 17:20

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago, The University of Chicago are collaborating with JSTOR to digitize, preserve and extend access to *Journal of Labor Economics*.

<http://www.jstor.org>

# Nonparametric Estimates of the Labor-Supply Effects of Negative Income Tax Programs

Orley Ashenfelter, *Princeton University*

Mark W. Plant, *Bureau of the Census,  
U.S. Department of Commerce*

This article reports nonparametric estimates of the effect of labor-supply behavior on the payments to families enrolled in the Seattle/Denver Income Maintenance Experiment. The randomized assignment of families to the treatment groups in this experiment was designed to permit the calculation of these nonparametric estimates. However, the nonparametric estimates have never been reported, even though they are easy to construct using a simple weighting procedure. Unfortunately, responses to the data collection instrument (which depended on costly surveys) were not random, and this opens up some ambiguity in the results.

## I. Introduction

In the 1970s, the Department of Health and Human Services funded the Seattle-Denver Income Maintenance Experiment SIME/DIME. It was the last and largest of four negative income tax (NIT) experiments in

This article was completed while Ashenfelter was a fellow at the Center for Advanced Study in the Behavioral Sciences. He is grateful for financial support from the Alfred P. Sloan and the Russell Sage foundations. The authors are also grateful for helpful discussions with David Card. The views expressed in this paper do not necessarily represent the views of the U.S. Department of Commerce.

[*Journal of Labor Economics*, 1990, vol. 8, no. 1, pt. 2]  
© 1990 by The University of Chicago. All rights reserved.  
0734-306X/90/0801-0021\$01.50

which assignments were made to the experimental programs by classical randomization.<sup>1</sup> The expressed purpose of the experiment was “to measure the disincentive effects of cash transfers on the labor force participation of those eligible for them” by simple comparisons of mean earnings between treatment and control groups (SRI International 1983, p. v.)<sup>2</sup> Ironically, the simple nonparametric results for which the experiment was designed have never been reported.<sup>3</sup>

In this article, we present the simple nonparametric estimates of the effects of the Seattle/Denver negative income tax program on work effort. It seems natural that the starting point of any experimental study of labor-supply effects should be an econometric analysis that is not contaminated by arbitrary assumptions. This, after all, is the benefit we expect to gain from the great expense of using experimental, as opposed to nonexperimental, methods. Frankly, when we began this research, we were surprised that these basic empirical results had never been tabulated. By the time we completed our research, we understood some of the reasons why that had not been done. We discovered that the basic problems with the SIME/DIME experiment are a result of its design and that the promised nonparametric estimates simply cannot be computed. In short, there is a great deal that can be learned for future experimental designs from the flaws in the design of the SIME/DIME.

The layout of the article is as follows. We first set out the economic effects of an NIT and how these may be studied with a randomized experiment. An apparent complication of all the actual NIT experiments is that, in order to reduce transfer costs, families with higher preexperimental incomes were systematically assigned to more generous programs, and

<sup>1</sup> Albert Rees was a principal investigator and articulate advocate for the SIME/DIME's precursor, the New Jersey Negative Income Tax Experiment. Rees (1974) reported his view of the findings of this experiment in a paper still well worth reading for its characteristic clarity and style. We hope our own article shows the same clarity.

<sup>2</sup> The experimental strategy underlying these NIT studies was clearly defined in an early paper by Orcutt and Orcutt (1968, p. 758): “Direct measurement of change in the earned income of any group, due solely to the introduction of a tax law, is out of the question. It is impossible simultaneously to apply, and not apply, a tax law to the same group. However, change in earned income of a group could be estimated by randomly drawing a sample from the group, and then using randomization in the assignment of tax treatments to members of the sample. . . . Differences between the earned incomes of the two sample groups then would be due solely to sampling variability and to differences in the tax treatment.”

<sup>3</sup> The final report of the experiment, SRI International (1983), goes a long way toward distilling the various studies into a coherent whole, but it does not present any of the simple two-way contrasts for which the experiment was designed. Whether this is because the experimenters were not aware of the simple method for constructing nonparametric estimations we describe below, or whether it is because of some other reason, we do not know.

families with lower preexperimental incomes were systematically assigned to less generous programs. At first blush, this type of design may appear to compromise the experiment, but it is easy to show, as we do below, that a simple weighting scheme will remedy any harm done. We next discuss the details of the design of the SIME/DIME and present the basic nonparametric results from it. As it turns out, a critical problem with the SIME/DIME experiment is that the data collection apparatus set up to analyze it could not operate independently of the treatments. In short, attrition from the experiment was closely related to the type of program in which an experimental family was enrolled. This fact complicates very considerably any attempt to draw inferences from the SIME/DIME experiment. We offer some conclusions about the experiment's design and results in a final section of the article.

## II. Behavioral Implications of an NIT Program

The basic setup of a negative income tax program is well known. If a family has no income, then it is guaranteed a subsidy of  $G$ . The subsidy,  $D$ , given to families with positive income,  $Y$ , is a decreasing function of income, with  $\tau$  being the rate of decrease. Thus,

$$D = G - \tau Y \quad \text{if } Y < G/\tau,$$

and

$$D = 0 \quad \text{if } Y > G/\tau.$$

The quantity  $B = G/\tau$  is called the break-even income level and is the income at which the subsidy is zero. If income is equal to the wage rate multiplied by hours worked, then under the negative income tax the implicit wage rate for low income workers is  $w(1 - \tau)$  since the individual earns  $w$  for an extra hour of work but the grant is reduced by  $w\tau$ . Thus  $\tau$  is referred to as the implicit tax rate. Current welfare programs are generally characterized by high implicit tax rates—often 100%.

The total transfer cost of an NIT program is determined by the number of participants in the program as well as the size of the subsidies paid to the eligible families. Both of these components depend on the earnings of families eligible for an NIT program. If earnings were fixed, the potential cost of an NIT program could be calculated by applying the formula above to the earnings of the eligible population. Since there is an implicit tax on wages, and since families whose incomes were initially below the break-even level will receive income transfers, the income and substitution effects in labor-supply functions may cause these families to decrease their hours of work and, thus, decrease their earnings in response to the presence of the NIT. Furthermore, families who initially would not have qualified for an NIT payment because their income was above the break-even level may

decrease their earnings sufficiently to become eligible. Thus, any preexperimental calculation of the cost of a welfare program may underestimate the total cost because the earnings of participants may decrease and because more families may participate than would be expected.

Attempts to measure the size of these economic reactions are what stimulated the interest in the NIT experiments. The experiments allow welfare program designers to gauge how much the program will actually cost once the eligible population adjusts their behavior to the incentives of the cash-transfer program.

### III. Basic Econometric Method

The total payment made in any NIT program can be broken into three components. The first is the total subsidy that would be predicted from the distribution of incomes in the eligible population before the NIT was instituted, assuming no behavioral response to the NIT. The second component is the additional amount paid as a result of a change in work effort to those whose income in the absence of the program would have been below the break-even level. The third component of the total cost is those payments to families whose earnings in the presence of a potential subsidy decrease sufficiently to make them eligible.

Let  $N$  be the total number of families who receive the NIT payment, and let  $D$  be the average amount of the subsidy. The total payment made is  $ND$ . Suppose that, in the absence of any labor supply effect,  $N^0$  families would be predicted to receive a mean NIT payment of  $D^0$ . Let  $\Delta E$  be the average change in earnings as the result of labor supply behavior among the  $N^0$  families. Suppose that  $N'$  families change their labor-supply behavior in such a way that they "become" participants, with average payment  $D'$ . Letting  $\tau$  be the average tax rate on earnings,

$$ND = N^0 D^0 + N^0 \tau \Delta E + N' D'. \quad (1)$$

Thus, the total cost of any NIT program is equal to the predicted payment assuming no labor-supply reaction,  $N^0 D^0$ , plus the payment due to changes in labor supply. The labor-supply effect can be broken into an increment due to a change in work effort among predicted participants,  $N^0 \tau \Delta E$ , and an induced participation effect,  $N' D'$ . The latter two components cannot be estimated without resort to some kind of experimental or economic inference.

The experimental design of the SIME/DIME is, in theory, ideally suited to estimation of an equation like (1). The average payment absent any labor supply effects,  $D^0$ , can be estimated using the earnings of the control families. The difference between the actual average payment made,  $D$ , and the payment predicted from using the control families,  $D^0$ , is an estimate of the labor-supply effect of the program. If there were no labor-supply

effects, the control families would have the same distribution of earnings as the experimental families and the same distribution of NIT payments. A comparison of the experimental families with the control families allows us to measure the extent to which the NIT changes earnings and payments. Thus, we can estimate directly the potential inaccuracy of any NIT cost calculation that ignores the endogeneity of work effort. No complex parametric characterization of labor supply is necessary for making these calculations.

The SIME/DIME study provides 11 different NIT programs varying by tax rate ( $\tau$ ) and generosity of guarantee ( $G$ ). The multiplicity of programs requires that we make several estimates of equation (1), one for each program. Variation in  $\tau$  and  $G$  is designed to permit the measurement of the incentive effects due to different programs. Specifically, we can measure the extent to which incentive effects vary with  $G$  and  $\tau$ .

The difficulty in using the SIME/DIME data to estimate (1) is that assignment to the various treatments was not random. Thus, given non-random assignment, the control families do not correctly characterize the distribution of earnings absent the experiment, and the simple comparison suggested above does not provide an estimate of the labor-supply effect of the experiment. At first blush, this design decision seems to have compromised the entire experiment, but, in fact, that is not the case. In order to understand why simple nonparametric estimates of labor-supply effects are still feasible, it is necessary to understand how the experiment was designed.

Prior to the experiment, each family (control or experimental) was assigned to one of six "normal-income" categories. Families were assigned to one of the 11 NIT programs on the basis of this categorization. A higher fraction of families in the high normal-income categories was assigned to the NIT programs with high break-even levels so as to reduce transfer-payment costs. However, within a normal-income category, assignment to control or experimental status was random. Only the sampling fraction varied across normal income category. Thus, for each normal-income category, unbiased estimates of what mean NIT payments would have been in the absence of incentive effects can be calculated. By taking the appropriate weighted averages of these means across normal-income categories, an unbiased nonparametric estimate of the expected average payment can be made for each of the 11 NIT programs.

Specifically, let  $i = 1, \dots, 6$  index the normal-income categories, and let  $j = 1, \dots, 11$  index the NIT programs. Define  $D_{ij}$  to be the average subsidy received by experimental families in normal-income category  $i$  in NIT program  $j$ .<sup>4</sup> Likewise, define  $D_{ij}^0$  to be the average subsidy that would

<sup>4</sup> That is,  $D_{ij} = G_j - \tau_j Y_i$ , where  $Y_i$  is mean earnings for experimental families in normal-income category  $i$ .

have been received by control families in normal-income category  $i$  in NIT program  $j$ . Variable  $D_{ij}^0$  is our estimate of the subsidy that families in normal-income category  $i$  in NIT program  $j$  would have received if they had not changed their labor-supply behavior. Now let  $w_i$  be the number of families in normal-income category  $i$  in the general population (or the population for which we wish to determine the labor-supply effects of the NIT program). The mean NIT payment that would be paid to the experimental group if it were distributed across normal-income categories as is the general population may be estimated as

$$D_j = \sum_i \frac{w_i}{w} D_{ij}, \quad (2)$$

where  $w = \sum_i w_i$  is the total population. Likewise, the mean predicted NIT payment that would be paid if (a) there were no labor-supply effects and (b) the control group were also distributed across normal income categories as is the general population may be estimated as

$$D_j^0 = \sum_i \frac{w_i}{w} D_{ij}^0. \quad (3)$$

It follows that the differences  $\Delta_j = D_j - D_j^0$  provide estimates of the labor-supply responses to the presence of NIT program  $j$ . The estimates  $D_j$  and  $D_j^0$  are both scaled so as to be representative of the general population, and so the labor-supply responses are also scaled to represent what we should expect from the general population.

The estimates  $D_{ij}$  and  $D_{ij}^0$  are specific to individual programs and normal-income categories. They form the most basic nonparametric estimates of the incentive effects of the negative income tax that the experiment has to offer. The problem with the program-by-normal-income-category specific estimates of incentive effects,  $\Delta_{ij} = D_{ij} - D_{ij}^0$ , is that, in the SIME/DIME, they are very imprecisely estimated. After tabulating these estimates, we realized that there were so many, and that they were so large compared to their estimated sampling errors, that they were nearly impossible to interpret. The problem is that the number of normal-income categories (6) and programs (11) is so large that the 66 estimated treatment effects are often based on samples of no more than 20–30 families.

In order to reduce this imprecision, it is necessary to weight up the estimates  $D_{ij}$  and  $D_{ij}^0$  to produce estimates of  $D_j$  and  $D_j^0$ , as in equations (2) and (3). The question naturally arises, What weights  $w_i$  should be used for this purpose? One natural procedure would be to use weights that reflect the distribution of normal-income categories in the aggregate population. So far as we are aware, however, the SIME/DIME experimenters

have never attempted to produce such weights. In short, such weights do not exist. As a result, it is not known precisely how the aggregate population is distributed using the methods for determining “normal incomes” that the experimenters used.

Given the data that we have, there are two obvious candidates to use for the weights  $w_i$ . On the one hand, we could use as weights the  $w_i$  in the control group. This has the advantage that we would have one set of weights that was commonly used in all contrasts between the experimental and control groups. As a result, even if there are interactions, so that treatment effects vary systematically by normal-income category, estimated effects for different treatments would be comparable. The disadvantage of using the  $w_i$  from the control group as weights is that the precision of the estimates we obtain may be very low. After all, the optimal weighting would attach more weight to the effects that are more precisely estimated, that is, to those estimated from the largest sample sizes. This suggests that more precise estimates may well be obtained by using the weights  $w_{ij}$  for treatment group  $j$  in preparing the contrast of  $D_j$  with  $D_j^0$ . In fact, this is the procedure for which we have opted throughout. Thus, the estimated treatment effect for program  $j$  is

$$\begin{aligned}\hat{\Delta}_j &= \sum_i w_{ij} D_{ij} - \sum_i w_{ij} D_{ij}^0 \\ &= \sum_i w_{ij} (D_{ij} - D_{ij}^0),\end{aligned}\tag{4}$$

where the  $w_{ij}$  are taken from table 2. The disadvantage of (4) is that it may produce estimated treatment effects by program that are not comparable if treatment effects vary systematically by normal-income category.

There are two final practical problems with which we must deal. First, the actual payment made to an experimental family could, in principle, be computed directly from the payment history provided with the SIME/DIME data. When we did this, the calculated payments were very small, and they did not seem to be correlated with earned income. Why this should be so is a mystery to us, but it suggests serious problems may have existed in the compilation of the payment data from the experiment. As an alternative, we have calculated what the payment to each family should have been given the income the family reported to the experimenters on a semiannual basis.

Second, another difficulty in computing these simple average effects arises from the fact that over the course of the experiment some families are lost from the experiment and no longer report their incomes. This attrition may be a serious problem if it is not random. If the presence of the experimental treatment changes the likelihood of the families leaving, then the control and experimental samples have been compromised, and



**Table 1**  
**Parameters of the 11 Negative Income Tax Programs**

Program Number	$G$ (\$)	$\tau$	Declining Tax Rate	Break-even Income (\$)
1	3,800	.5	No	7,600
2	3,800	.7	No	5,429
3	3,800	.7	Yes	7,367
4	3,800	.8	Yes	5,802
5	4,800	.5	No	9,600
6	4,800	.7	No	6,857
7	4,800	.7	Yes	12,000
8	4,800	.8	Yes	8,000
9	5,600	.5	No	11,200
10	5,600	.7	No	8,000
11	5,600	.8	Yes	10,360

NOTE.—Terms are explained in text.

the randomization has broken down. We examine this attrition effect in detail in the empirical results reported below.

**IV. The Design of the SIME/DIME NIT Programs**

A brief overview of the experiment performed in Seattle and Denver is necessary before presenting our empirical results. In table 1 we present the parameters characterizing the eleven different NIT treatments to which families were exposed. The guarantee level is the amount of money a family of four would receive in the absence of any income.<sup>5</sup> The tax rate is the amount by which that grant decreases for each dollar a family earns. For five of the programs, the tax rate declines at the rate of  $r = 0.000025$  per dollar earned.<sup>6</sup> The break-even point is the amount of income at which the subsidy is zero and is a rough measure of overall program generosity.

Income information is available for families on a semiannual basis beginning a year before the experiment was implemented and continuing at least 1 year after implementation of the experiment. The pre- and post-experimental data give us an opportunity to “check” the random assignment in the program and to gauge intertemporal substitution effects. If, in fact, the control and experimental families are randomly assigned (within normal-income category), then the predicted mean payment difference before the experiment should not be significantly different from zero. The post-

<sup>5</sup> The grant level of the program offered was adjusted for family size. Standardizing all incomes to a family size of four did not change the empirical results reported below.

<sup>6</sup> In a nondeclining tax-rate program, the formula for the subsidy is  $D = G - \tau Y$ , where  $Y$  is dollar income. In the declining rate programs, it is  $D = G - (\tau - rY)Y$ . The break-even income is  $G/\tau$  in the nondeclining tax-rate programs, and it is the solution to the quadratic equation  $G - \tau Y - rY^2 = 0$  for the declining tax-rate programs.

experimental difference may be positive if there are intertemporal income effects. Experimental families are exposed to the treatment for either 3 or 5 years. Thus we have between 5 and 7 years of data on each family that does not leave the experiment. For control families, the full 7 years of data are usually available. We use data from the dual-headed household file.

As mentioned earlier, a nonrandom assignment method was used based on a categorization of “normal income” into one of six groups. Table 2 shows the initial distribution of normal income for each of the 11 NIT programs for dual-headed households. These distributions change slightly over time due to attrition, but the basic design of assigning low-income families to the low break-even program persists throughout the experiment.

**Table 2**  
**Percentage Distribution of NIT Program Assignment**  
**by Normal-Income Category**

Experiment Duration by Program (Years)	Normal-Income Category (%)*						Number Initially in Program
	1	2	3	4	5	6	
1:							
3	8.7	9.6	21.7	55.7	3.5	.9	115
5	2.1	8.5	25.5	59.6	2.1	2.1	47
2:							
3	3.7	9.3	32.0	50.0	.0	.0	54
5	6.5	22.6	9.7	58.1	.0	3.2	31
3:							
3	9.6	17.8	42.5	30.1	.0	.0	73
5	2.8	25.0	38.9	33.3	.0	.0	36
4:							
3	3.7	11.1	50.0	33.3	.0	1.9	54
5	.0	16.7	60.0	23.3	.0	.0	30
5:							
3	4.3	9.7	10.8	7.5	37.6	30.1	93
5	.0	5.4	10.7	5.4	51.8	26.8	56
6:							
3	.0	2.0	17.2	43.4	33.3	4.0	99
5	.0	.0	16.7	35.4	41.7	6.3	48
7:							
3	2.5	7.5	10.0	21.7	33.3	25.0	120
5	.0	.0	15.6	15.6	57.9	10.5	19
8:							
3	2.4	2.4	24.4	27.6	23.2	.0	82
5	.0	4.2	20.8	50.0	25.0	.0	48
9:							
3	.0	10.0	12.0	8.0	26.0	44.0	50
5	.0	.0	12.0	4.0	20.0	64.0	25
10:							
3	.0	3.4	9.0	31.5	50.6	5.6	89
5	.0	3.0	15.2	33.3	48.5	.0	33
11:							
3	.9	.0	8.1	20.7	32.4	37.8	111
5	.0	.0	5.1	20.3	32.2	42.4	59

\* The normal income categories are: 1 = \$1–\$1,000, 2 = \$1,000–\$3,000, 3 = \$3,000–\$5,000, 4 = \$5,000–\$7,000, 5 = \$7,000–\$9,000, and 6 = \$9,000–\$11,000. Families with over \$11,000 in normal income were excluded from the sample altogether.

Table 3  
Experimental Payment minus Predicted Control Payment for 3-Year  
Dual-headed Experimental Families, Attrition Families Excluded  
(Standard Errors in Parentheses)

G (\$)	$\tau$	Declining Tax Rate	Preexperimental Payment (\$)	Payments for Year of Experiment (\$)			Postexperimental Payment (\$)
				1	2	3	
3,800	.5	No	193.78 (143.45)	248.46 (149.58)	368.95* (170.75)	389.24* (182.99)	138.56 (188.20)
3,800	.7	No	124.96 (223.77)	185.18 (237.91)	317.28 (252.99)	218.37 (325.57)	-47.85 (314.66)
3,800	.7	Yes	-33.37 (178.05)	68.94 (176.07)	158.44 (213.59)	324.84 (230.50)	29.28 (222.42)
3,800	.8	Yes	75.40 (229.44)	336.06 (237.18)	221.54 (245.92)	160.83 (264.53)	91.52 (261.84)
4,800	.5	No	52.02 (192.31)	85.17 (184.85)	294.55 (201.73)	337.23 (221.73)	70.22 (219.58)
4,800	.7	No	220.76 (160.04)	288.33 (169.04)	496.85* (197.88)	543.25* (204.50)	178.32 (194.03)
4,800	.7	Yes	136.99 (127.36)	281.98* (137.19)	423.30* (157.51)	348.03* (162.38)	23.96 (140.58)
4,800	.8	Yes	-16.87 (175.54)	305.09 (209.24)	417.90 (234.32)	317.39 (274.11)	121.47 (239.59)
5,600	.5	No	-163.12 (252.05)	200.75 (258.13)	664.41* (283.28)	717.15* (280.65)	124.93 (287.04)
5,600	.7	No	-59.97 (164.95)	23.34 (156.41)	386.12 (200.59)	744.94* (263.80)	267.69 (259.45)
5,600	.8	Yes	-27.64 (121.47)	-51.03 (126.67)	117.85 (138.52)	273.44 (157.96)	121.53 (169.26)

NOTE.—Terms are explained in text.  
\* Denotes mean is more than twice its standard error.

In table 3 we present the basic results for families participating in the 3-year program. The first three columns give the various program parameters. The column label “Preexperimental” gives the difference between the mean payments that would have been made to experimentals and controls based on preexperimental income. Since no family is exposed to an NIT at that point, this difference should be zero. In fact, none of the preexperimental differences are significantly different from zero, although some are rather large, particularly the (3,800, .5, No) (4,800, .7, No) and the (5,600, .5, No) programs. About one-half the preexperimental differences are positive.

The results for the 3 experimental years are given in the next three columns. All but one of the 33 differences in mean payments are positive, and 10 are significantly different from zero, thus showing evidence of some experimental effect on labor supply. The largest effects are in the (5,600, .5, No) program, which has the largest break-even level (\$11,200). The experimental-control differences tend to become larger over time.

It is difficult to tell from casual examination of these numbers if the

increased payments are due to income or substitution effects. Income and substitution effects work in the same direction for any NIT program since wage rates are lowered and lump-sum income is increased. However, the extent to which the decreased earnings are a result of income or substitution effects makes some difference for the design of welfare programs. If there is no substitution effect, then holding the break-even income constant, local changes in the implicit tax rate will have no local earnings effect. The design of the experiment allows us to do several comparisons in an attempt to isolate income and substitution effects. We could, in theory, compare programs with identical tax rates and different guarantee levels to get an idea of the income effects involved. However, since the assignment to programs is nonrandom, across-program comparisons may be misleading because, in the weighting system we use, the distribution of incomes underlying the averages presented differs.

We note also that the postexperimental predicted-payment differences tend to be smaller than any experimental differences and not statistically different from zero. Nevertheless, most of the postexperimental treatment/control differences are positive. This indicates that the *experiment* in fact was changing earnings during the course of the experiment. Thus the observed differences do not seem to be artifacts of the peculiar sample design or some other “unseen” circumstances affecting the sample makeup. On the other hand, the persistence of the positive treatment effects into the postexperimental period suggests there may have been some effects of the experimentation on permanent income.

Results for families exposed to the NIT for 5 years are presented in table 4. It is worth noting that the sample sizes for the 5-year experimental programs are substantially smaller than those for the 3-year programs, and thus the results show more variation over time and across program. The small samples are also reflected in the large standard errors. For the most part, the payment differences seem to be larger for the 5-year families, and there is some mixed evidence of growth in the payment differences over time. Attrition becomes a serious problem in the fourth and fifth years of the experiment, and thus the small sample problem becomes more acute.

These two tables present the basic nonparametric results of the experiment. These seem to us to be the logical starting point for any further analysis of the results of this experiment since any additional structure will add additional exogenous and ad hoc information onto the basic data. All the information available on the effect of these programs on predicted payments is summarized in these tables. Absent any parametric assumptions, these are the results of the experiment.

## V. Attrition

There is one important problem confronting the legitimacy of these results: attrition. The results reported thus far are dependent on the as-

**Table 4**  
**Experimental Payment minus Predicted Control Payment for 5-Year Dual-headed Experimental Families,**  
**Attrition Families Excluded (Standard Errors in Parentheses)**

G (\$)	$\tau$	Declining Tax Rate	Preexperimental Payment (\$)	Payment for Year of Experiment (\$)					Postexperimental Payment (\$)
				1	2	3	4	5	
3,800	.5	No	102.24 (185.55)	345.68 (221.42)	526.02 (241.53)	110.30 (265.28)	390.07 (307.01)	169.82 (286.76)	229.70 (309.06)
3,800	.7	No	81.16 (309.85)	23.30 (316.06)	-99.33 (330.14)	98.20 (383.52)	-16.42 (388.07)	-122.01 (352.95)	-406.46 (314.40)
3,800	.7	Yes	6.99 (234.01)	490.00 (288.13)	176.14 (272.87)	23.22 (300.28)	324.70 (386.93)	-59.79 (331.68)	-598.09* (102.72)
3,800	.8	Yes	-130.30 (271.23)	349.73 (286.56)	189.80 (280.63)	329.94 (365.58)	1207.82* (463.10)	1108.49* (487.83)	307.38 (453.29)
4,800	.5	No	-23.66 (183.73)	30.15 (208.90)	160.40 (199.26)	399.28 (236.33)	419.73 (247.25)	434.30 (254.52)	251.09 (242.45)
4,800	.7	No	-129.98 (185.46)	25.71 (208.14)	-4.47 (211.44)	569.10 (314.73)	493.42 (357.32)	219.74 (340.60)	-38.46 (228.01)
4,800	.7	Yes	75.66 (234.21)	224.96 (280.43)	387.66 (367.56)	340.71 (404.05)	-130.10 (308.90)	34.61 (445.67)	189.49 (491.52)
4,800	.8	Yes	467.89 (252.40)	325.17 (276.31)	599.43* (724.39)	398.62 (280.50)	537.21 (365.56)	506.95 (351.98)	346.28 (337.43)
5,600	.5	No	-224.97 (286.39)	560.51 (298.21)	723.08* (306.90)	782.53* (327.39)	592.40 (366.88)	313.82 (387.31)	-53.07 (325.66)
5,600	.7	No	-158.74 (239.17)	500.18 (311.24)	1194.68* (416.25)	890.38* (391.61)	825.39 (467.76)	435.01 (609.49)	588.91 (510.52)
5,600	.8	Yes	-6.48 (175.15)	193.54 (199.51)	617.29* (255.89)	906.13* (315.98)	888.72 (337.38)	877.71 (398.38)	75.21 (216.12)

NOTE:—Terms are explained in text.

\* Denotes mean is more than twice its standard error.

sumption of random assignment (within normal-income category) throughout the experiment. In such data-gathering efforts, attrition is always a problem. If it is not random, then the reliance on the assumption of random assignment is unjustified, and, in fact, nonparametric estimation is impossible. When there is attrition, some assumption must be made regarding the behavior of the families who left the experiment and how it differs from those remaining. If one is interested in inferring from the experiment the results of a potential national welfare reform, then experimental attrition must be accounted for in results such as those presented in tables 3 and 4.

In fact, tables 3 and 4 make an implicit assumption about attrition that it is in fact random and that, therefore, the comparability between controls and experimentals does not change over time. This assumption implies that the result of the attrition is only a decrease in sample size and a decrease in the precision of estimation of the treatment effects.

Attrition is a measurable phenomenon, and comparisons between attrition from the experimental sample and the control sample can be made. The key question is whether the experimental treatments are changing the rate of attrition of experimentals. Within each NIT program, is the rate of experimental attrition equal to what would be expected given the rates of control family attrition? If the assignment of experimentals were truly random, a simple comparison of rates would be possible. But since assignment is based on normal-income level, a weighted measure must be computed. Specifically, let  $w_{ij}$  be the proportion of experimental families eligible for treatment  $j$  who are from the  $i$ th normal income category. Let  $a_i$  be the attrition rate for control families in the  $i$ th normal-income category. Then the predicted rate of attrition for experimental group  $j$  is

$$a_j = \sum_i w_{ij} a_i.$$

Note that  $a_j$  differs across programs only because the weights,  $w_{ij}$ , differ across programs. The assumption that attrition is random thus implies that attrition rates across programs are the same and that they are equal to the control-group attrition rate.

In the sample of dual-headed households, there are two types of attrition that occur. The first results from families that simply stopped reporting income. The second results because one of the heads of household left the family unit, thus changing the family's earning structure. Although payments continued to the "new" family, the structure of the data set makes such families very difficult to follow. In theory, either type of attrition could be affected by the experimental treatment. For example, families whose income was substantially above the break-even level for the particular treatment had little incentive to continue to report income since no transfer

payment would ever be made.<sup>7</sup> For families who expected to receive a transfer payment, there was a clear incentive to report income regularly. One would expect to see less attrition in programs that were more generous (had higher break-evens points), holding constant the income distribution. After all, families who were in the generous treatment groups had a substantial incentive not to disappear. The literature on household breakups caused by the experiment is large,<sup>8</sup> and we make no attempt to summarize it here. Here we simply observe that such attrition could result in a non-random sample of experimental families.

Table 5 presents predicted and actual rates of attrition for experimental families who simply left the program. It excludes attrition due to family head separation. The top panel of table 5 shows the results ordered by the usual program parameters presented before. The bottom panel is ordered by break-even level of the program. The result is striking: *actual attrition among experimentals is always less than predicted*. And as the break-even income of the program increases, actual attrition decreases dramatically relative to predicted attrition. In short, families who were in programs where their incomes were unlikely to fall below the cutoff where they would receive a payment were far more likely to leave the program than were other families. (The correlation coefficient between the break-even income and the difference between actual and predicted attrition is .84.) These results lead us to believe that families who left the program did so because it was not in their economic interest to stay. That is, the expected payment to an experimental subject, given attrition, may have been zero. If this is the case, the computations in tables 3 and 4 are biased since experimentals are underrepresented in them and zero-payment experimentals are most underrepresented. Since we know the number of families leaving the experiment, we can calculate an "actual mean experimental payment" under the assumption that the expected payment to families that leave is zero to correct for this potential bias. The results of that modification to tables 3 and 4 are presented in tables 6 and 7. Averaging in zero payments to attrition families lowers the average experimental payment and also changes the distribution of families within the experiment across normal-income categories. For the most part, differences between predicted and actual payments decrease substantially and, in several instances, become statistically indistinguishable from zero. Nevertheless, most of the experimental/treatment contrasts are positive, indicating the presence of a labor-supply effect.

In table 8, we present the predicted and actual attrition rates for those families for whom data become unavailable due to changes in family struc-

<sup>7</sup> There was a \$20 payment made for reporting income every month. Thus there was some mild incentive to do so.

<sup>8</sup> For a summary, see SRI International (1983).

**Table 5**  
**Attrition Rates for Families Leaving Program**  
**A. Attrition by Program**

G (\$)	$\tau$	Declining Tax Rate	Number of Families	Actual % Attrition	Predicted % Attrition	Difference
3,800	.5	No	162	11.1	22.2	-11.1*
3,800	.7	No	85	21.2	21.9	-.7
3,800	.7	Yes	109	16.5	21.1	-4.6
3,800	.8	Yes	84	14.3	20.3	-6.0
4,800	.5	No	150	10.0	20.8	-10.8*
4,800	.7	No	147	14.3	20.8	-6.5
4,800	.7	Yes	140	8.6	21.1	-12.5*
4,800	.8	Yes	130	14.6	21.1	-6.5
5,600	.5	No	75	5.3	21.5	-16.2*
5,600	.7	No	123	13.8	20.4	-6.6
5,600	.8	Yes	170	10.6	20.6	-10.0*

**B. Attrition by Break-even Income**

Break-even Income (\$)	Actual % Attrition	Predicted % Attrition	Difference
5,429	21.2	21.9	-.7
5,802	14.3	20.3	-6.0
6,857	14.3	20.8	-6.5
7,367	16.5	21.1	-4.6
7,600	11.1	22.2	-11.1*
8,000	13.8	20.4	-6.6
8,000	14.6	21.1	-6.5
9,600	10.0	20.8	-10.8*
10,360	10.6	20.6	-10.0*
11,200	5.3	21.5	-16.2*
12,000	8.6	21.1	-12.5*
			$\rho = -.84^*$

NOTE.—Terms are explained in text.

\* Denotes that the actual rate is significantly different from the predicted rate assuming the predicted rate is non-random.

ture. Again we present the results in the usual manner in the upper panel and sorted by break-even income in the lower panel. There does not seem to be any particular relationship between program generosity and attrition. From these results alone, there is no obvious assumption to be made regarding what payment would have been made if we had continued to observe these families.

The fact of attrition necessitates some sort of “parametric” assumption about the sample. Either we assume the departing families are randomly chosen from the distribution of incomes and their expected excess payment is the same as those families remaining (tables 3 and 4), or we assume they in fact should be assigned zero payment (tables 6 and 7). We could do something in between, but either way an arbitrary assumption is necessary.<sup>9</sup>

<sup>9</sup> The problem of attrition in experimental data suggests that, rather than using survey data, administrative data should be used. For example, in a paper by Wood-



Table 6  
Experimental Payment minus Predicted Control Payment for 3-Year  
Dual-headed Experimental Families, Attrition Families Included and  
Single-headed Families Excluded (Standard Errors in Parentheses)

G (\$)	$\tau$	Declining Tax Rate	Preexperimental Payment (\$)	Payment for Year of Experiment (\$)			Postexperimental Payment (\$)
				1	2	3	
3,800	.5	No	193.78	225.83	335.34*	308.59	138.56
			(143.45)	(148.12)	(167.16)	(173.32)	(188.20)
3,800	.7	No	124.96	185.18	267.18	112.67	-47.85
			(223.77)	(237.91)	(238.02)	(273.31)	(314.66)
3,800	.7	Yes	-33.37	9.02	60.00	166.80	29.28
			(178.05)	(174.05)	(205.08)	(221.16)	(222.42)
3,800	.8	Yes	75.40	314.12	184.40	100.15	91.52
			(229.44)	(234.11)	(233.37)	(238.01)	(261.84)
4,800	.5	No	52.02	99.92	264.63	334.01	87.91
			(192.31)	(186.19)	(196.30)	(213.00)	(222.46)
4,800	.7	No	220.76	264.26	451.21*	465.12*	178.32
			(160.04)	(164.90)	(188.21)	(189.28)	(194.03)
4,800	.7	Yes	136.99	273.49*	399.37*	299.34	23.96
			(127.36)	(136.73)	(155.81)	(158.22)	(140.58)
4,800	.8	Yes	-16.87	279.95	302.62	134.21	121.47
			(175.54)	(205.49)	(216.43)	(233.79)	(239.59)
5,600	.5	No	-163.12	200.75	635.11*	686.28*	124.93
			(252.05)	(258.14)	(280.70)	(278.23)	(287.04)
5,600	.7	No	-59.97	-9.20	308.18	590.95*	267.69
			(164.95)	(152.60)	(188.81)	(239.35)	(259.45)
5,600	.8	Yes	-27.64	-62.75	91.95	228.00	121.53
			(121.47)	(125.40)	(134.82)	(152.62)	(169.26)

NOTE.—Terms are explained in text.  
\* Denotes mean is more than twice its standard error.

VI. Discussion of the Results

There are two clear conclusions that we may draw from the results in tables 3–7. First, differences of actual from predicted payments depend strongly on the probability that a family received a payment. As table 5 indicates, the family was more likely to receive a payment the greater the break-even income in the NIT program they were offered was. This relationship is not primarily behavioral but, instead, reflects the fact that more families will naturally be eligible for payments if the income threshold

bury and Speigelman (1987), an experiment on the effects of various incentives to end unemployment was administered using the record-keeping system already established by the state of Illinois for unemployed persons. These records did not rely on self-reporting but on reports by employers regularly obtained by the state. Thus there was no possibility for attrition on the part of participants. It is unclear how this might be done with welfare programs, although Social Security records are one avenue of approach. See Greenberg and Halsey (1983). Another advantage of Social Security records is that they mimic the kind of records that would probably exist if a negative income tax were actually adopted.

**Table 7**  
**Experimental Payment minus Predicted Control Payment for 5-Year Experimentals, Attrition Families Included and Single-headed Families Excluded (Standard Errors in Parentheses)**

G (\$)	$\tau$	Declining Tax Rate	Preexperimental Payment (\$)	Payment for Year of Experiment (\$)					Postexperimental
				1	2	3	4	5	
3,800	.5	No	102.24 (185.55)	347.68 (221.42)	501.64* (237.61)	79.55 (252.07)	302.80 (286.09)	16.93 (251.87)	229.70 (309.06)
3,800	.7	No	81.16 (309.85)	24.31 (306.23)	-127.77 (276.63)	-17.73 (319.75)	-122.39 (318.68)	-234.97 (272.62)	-406.46 (314.40)
3,800	.7	Yes	6.99 (234.01)	451.72 (289.00)	19.53 (261.83)	-198.65 (267.67)	11.81 (338.90)	-359.08 (261.22)	-598.09* (102.72)
3,800	.8	Yes	-130.30 (271.23)	366.43 (286.04)	207.49 (270.73)	179.04 (316.47)	818.42* (401.60)	652.98 (402.68)	307.38 (453.29)
4,800	.5	No	-23.66 (183.73)	36.34 (206.02)	140.35 (196.16)	374.48 (232.51)	358.79 (243.46)	347.62 (246.59)	251.09 (242.45)
4,800	.7	No	-129.98 (185.46)	8.72 (204.18)	-76.22 (196.61)	434.62 (292.76)	277.88 (329.27)	-13.02 (292.73)	-38.46 (228.01)
4,800	.7	Yes	75.66 (234.21)	161.92 (275.52)	184.27 (339.43)	39.33 (345.03)	-355.17 (262.18)	-222.78 (368.51)	189.49 (491.52)
4,800	.8	No	467.89 (252.40)	326.53 (276.29)	571.15* (270.06)	387.91 (275.22)	405.38 (333.56)	322.45 (313.93)	346.28 (373.43)
5,600	.5	No	-224.97 (286.39)	560.51 (298.21)	695.72* (306.52)	759.18* (326.79)	544.91 (366.66)	216.44 (386.88)	-53.07 (325.66)
5,600	.7	No	-158.74 (239.17)	514.92 (311.19)	1207.34* (416.15)	951.05* (391.12)	847.96 (453.51)	254.61 (503.70)	588.91 (510.52)
5,600	.8	Yes	-6.48 (175.15)	166.13 (197.13)	614.26* (252.46)	824.43* (306.92)	721.03* (316.72)	615.81 (354.91)	75.21 (216.12)

NOTE.—Terms are explained in text.  
 \* Denotes mean is more than twice its standard error.

Table 8  
Attrition Rates due to Change in Family Structure  
A. Attrition by Program

G (\$)	$\tau$	Declining Tax Rate	Number of Families	Actual % Attrition	Predicted % Attrition	Difference
3,800	.5	No	162	36.4	36.2	.2
3,800	.7	No	85	40.0	37.2	2.8
3,800	.7	Yes	109	37.6	39.9	-2.3
3,800	.8	Yes	84	29.8	39.0	-9.2
4,800	.5	No	150	27.3	32.5	-5.2
4,800	.7	No	147	35.4	32.1	3.3
4,800	.7	Yes	140	32.1	32.5	-.4
4,800	.8	Yes	130	35.4	33.6	1.8
5,600	.5	No	75	24.0	32.3	-8.3
5,600	.7	No	123	30.1	31.4	-1.3
5,600	.8	Yes	170	24.7	34.5	-9.8*

B. Attrition by Break-even Income

Break-even Income (\$)	Actual % Attrition	Predicted % Attrition	Difference
5,429	40.0	37.2	2.8
5,802	29.8	39.0	-9.2
6,857	35.4	32.1	3.3
7,367	37.6	39.9	-2.3
7,600	36.4	36.2	.2
8,000	30.1	31.4	-1.3
8,000	35.4	33.6	1.8
9,600	27.3	32.5	-5.2
10,360	24.7	34.5	-9.8*
11,200	24.0	32.3	-8.3
12,000	32.1	32.5	-.4
			$\rho = -.37$

NOTE.—Terms are explained in text.  
\* Denotes that the actual rate is significantly different from the predicted rate assuming the predicted rate is non-random.

they must pass in order to lose their payments is high. (A detailed analysis is contained in Ashenfelter [1983] and Plant [1984].) As a result, any model that is designed to provide a simplified representation of the nonparametric results in tables 3, 4, 6, and 7 will almost certainly have to be composed of two parts: one part allowing for the interaction of the income distribution and the break-even income in the determination of program participation, and the other allowing for the effect of the program on labor supply, given that a family is eligible to receive payments.

Second, it is quite clear from the results in tables 3, 4, 6, and 7 that the actual experimental payments are larger than what would be expected if no labor-supply response to the incentives set up by an NIT were present. This implies that there were incentive effects working to reduce labor supply as a result of the experimental treatment. Precisely how to untangle

the effects of the participation and the magnitude of income and substitution effects remains a problem for further research.

## VII. Conclusion

This article contains the first tabulation of the nonparametric estimates of labor-supply responses to a negative income tax program available from the Seattle and Denver Income Maintenance Experiments. The results do indicate that the incentive effects on labor supply induced by an NIT would have a real effect on the transfer costs of such a program. Unfortunately, precisely how these results should be used to estimate aggregate program effects has still to be worked out. Moreover, the estimated size of these nonparametric estimates of incentive effects depends significantly on how nonrandom attrition from the treatment groups is handled.

The major lessons we have learned from our analyses are methodological. First, we think that the experimental method offers enormous advantages in the estimation of labor-supply effects because the analyst no longer has to depend on a whole host of untestable econometric assumptions. In practice, however, most of the analyses of the SIME/DIME data have taken little or no advantage of the experimental nature of the treatments. It seems to us that too many opportunities for setting out the basic experimental results were missed. As we have shown, the fact that the treatment assignment design was random within, but not across, normal-income categories does not prohibit the computation of simple nonparametric program treatment effects.

On the other hand, the reliance on a massive survey research effort to collect data on the labor-supply responses of members of the treatment and control groups was terribly misguided. Not only was this data collection effort enormously costly (the majority of the experiment's costs went for data collection and analysis, not program payments), it also damaged the integrity of the experiment by inducing a correlation between realized family incomes and attrition from the data-collection system. The result is that even nonparametric estimates of NIT program effects must rely on assumptions that the experimental design was supposed to make unnecessary.

## References

- Ashenfelter, Orley. "Determining Participation in Income-tested Social Programs." *Journal of the American Statistical Association* 78 (September 1983): 517–25.
- Greenberg, David, and Halsey, Harlan. "Systematic Misreporting and Effects of Income Maintenance Experiments on Work Effort: Evidence from the Seattle-Denver Experiment." *Journal of Labor Economics* 1 (October 1983): 380–407.

- Orcutt, Guy H., and Orcutt, Alice G. "Incentive and Disincentive Experimentation for Income Maintenance Purposes." *American Economic Review* 58 (September 1968): 754-72.
- Plant, Mark W. "An Empirical Analysis of Welfare Dependence." *American Economic Review* 74 (September 1984): 673-84.
- Rees, Albert. "An Overview of the Labor Supply Results." *Journal of Human Resources* (Spring 1974): 158-80.
- SRI International. *Final Report of the Seattle-Denver Income Maintenance Experiment*, vol. 1. Menlo Park, Calif.: SRI International, 1983.
- Woodbury, Stephen A., and Spiegelman, Robert G. "Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois." *American Economic Review* 77 (September 1987): 513-30.