215-74

# INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

## PARAMETERIZATIONS OF THE "TREATMENT" IN NEGATIVE INCOME TAX EXPERIMENTS

Dale J. Poirier

UNIVERSITY OF WISCONSIN ~ MADISON

# PARAMETERIZATIONS OF THE "TREATMENT" IN NEGATIVE INCOME TAX EXPERIMENTS

Dale J. Poirier

August 1974

The research reported here was supported in part by funds provided to the Institute for Research on Poverty at the University of Wisconsin-Madison by the Office of Economic Opportunity pursuant to the Economic Opportunity Act of 1964. The opinions expressed are those of the author.

## ABSTRACT

In analyzing *ex post* the *Final Report of the New Jersey Graduated Work Incentive Experiment*, one is confronted with various approaches to characterizing or parameterizing the experimental treatment. In addition, the recent origin of controlled social experimentation implies that most researchers are likely to have little, if any, previous experience with these approaches. The purpose of this discussion is to summarize some of these approaches in the hope of facilitating parameterization choices by future researchers. While the discussion will deal only with the New Jersey-Pennsylvania (Urban) Experiment, most of the parameterizations are applicable to other negative income tax experiments as well.

PARAMETERIZATIONS OF THE "TREATMENT" IN
NEGATIVE INCOME TAX EXPERIMENTS

Dale J. Poirier*

July 1974

In analyzing *ex post* the *Final Report of the New Jersey Graduated Work Incentive Experiment*, one is confronted with various approaches to characterizing or parameterizing the experimental treatment. In addition, the recent origin of controlled social experimentation implies that most researchers are likely to have little, if any, previous experience with these approaches. The purpose of this discussion is to summarize some of these approaches in the hope of facilitating parameterization choices by future researchers. While the discussion will deal only with the New Jersey-Pennsylvania (Urban) Experiment, most of the parameterizations are applicable to other negative income tax experiments as well.[1]

Briefly, the experimental treatment is defined by the guarantee rate g, expressed as a percentage (divided by 100) of the poverty level $P_n$ for a family of size n, and the tax rate t.[2] The tax rate gives the rate at which the transfer payment is reduced as the family's income increases, and hence is often called the "rate of reduction." Letting $\pi$ denote the dollar amount of the transfer payment, then

$$\pi = \begin{cases} G - tY, & Y < G/t \\ 0, & Y \geq G/t \end{cases} \quad , \tag{1}$$

where $G = gP_n$ is the dollar amount of the guarantee and Y is the family's income.[3] The point at which the transfer payment ceases, i.e., G/t, is called the *break-even point*.

Since most of the analysis in the Final Report has been performed in a regression context, that will be the only framework considered here.

Specifically, models of the form $y = f(Z) + g(X,Z) + u$ will be considered, where $y$ is the dependent variable of interest, $f(Z)$ is a control function of various socioeconomic-demographic (nonexperimental) variables $Z$, $g(X,Z)$ is an experimental response function that depends on the treatment variables $X$ and possibly $Z$, and $u$ is a random disturbance term whose properties depend on the particular context. It should be emphasized that the control function $f(Z)$ must not contain any induced experimental effects such as would arise from including a contemporaneous variable such as income. Hence, $f(Z)$ often involves pre-enrollment values and estimated "normal" variables such as found in Poirier and Watts [14] and Watts [17]. This condition must be met if the experimental response function $g(X,Z)$ is to isolate the effects of the treatment. In what immediately follows, attention will be focused on the parameterization of $X$ alone; later, mention will be made of the many possibilities for interactive responses with $Z$.

Undoubtedly, the simplest parameterization involves using an experimental dummy, which equals one for an experimental observation and zero for a control observation. This approach implies that the effect of the treatment only manifests itself through different intercepts for the control and experimental groups. The regression coefficients of all other variables in the control function $f(Z)$ are assumed to be identical for the two groups. The statistical significance of the treatment is determined by simply examining the t-ratio of the experimental dummy's coefficient.

Clearly this approach is highly restrictive. If the effects of various independent variables are believed to be different for the experimental and control groups, then the experimental dummy can be interacted with these independent variables providing for different effects between the two groups. Conventional joint F-tests can then be used to determine the significance of the treatment response. In the extreme case in which the coefficients of all independent variables differ between the two groups, the model can be estimated separately for experimentals and for controls.[4] Testing for the equality of coefficients between the two groups can again be analyzed by a simple F-test.[5]

However, the preceding uses of the experimental dummy ignore differences in the various levels of treatment that were administered. One simple approach to overcome this is to replace the single experimental dummy by eight dummies--one for each of the plans. (The "omitted group" is the control group.) Unfortunately, experience with this approach indicates that more often than not, such a representation is "too flexible," i.e., there seldom appears any systematic variation across plans, and few statistically significant differences between plans are found. A slight modification, which has been mildly successful, has been used by Watts et al. [20]; it assigns each plan to one of the three groups that measure "degrees of generosity." The "low" guarantee-tax plans (in percentages) are the 50-50 and 75-70 plans, both of which were dominated by New Jersey welfare during most of the experiment and both of which were subject to very high attrition rates.[6] The "medium" plans include the 50-30, 75-50, and 100-70 plans, and the "high" plans include the 75-30, 100-50, and 125-50 plans.

In direct contrast to dummy variable approaches is the approach that focuses on a continuous response in the tax-guarantee dimension. Indeed, as Skidmore [16, p. 43] points out, even in the early planning stages of the experiment it was felt that the various plans should not be treated as totally distinct, but rather their similarities (common tax or guarantee rates) should be exploited. As a result, the approach outlined in Watts [18] is often an attractive alternative. Briefly, this approach involves using the experimental dummy, hereafter denoted by $s_1$, together with other variables involving explicit tax and guarantee rates.[7] Translating the origin of the tax rate guarantee rate dimension to the 75-70 plan,[8] the following two variables are added: $s_2 = s_1(g - .75)$ and $s_3 = s_1(t - .5)$. Hence, the coefficient of $s_1$, say $\beta_1$, measures the planar experimental-control differential evaluated at the central 75-50 plan. The coeffieients $\beta_2$ and $\beta_3$ of $s_2$ and $s_3$ are simply partial derivations with respect to the guarantee and tax, respectively. Nonlinear but additive effects of the tax and/or guarantee rates can be allowed for by adding all or some of the following linear spline variables: $s_4 = \max(g - .75,0)$, $s_5 = \max(t - .5,0)$, and $s_6 = \max(g-1.0,0)$. Introducing $s_4$ permits a change in the partial derivative with respect to the guarantee rate at $g = .75$, i.e., the partial equals $\beta_2$ for $g < .75$ and $\beta_2 + \beta_4$ for $g > .75$. Similarly, if $s_6$ is included, then the partial derivative changes again at $g = 1.0$, i.e., for $g > 1.0$ the partial equals $\beta_2 + \beta_4 + \beta_6$. A similar interpretation holds for $s_5$, which permits a "kink" at $t = .5$. One attractive feature of this representation is that the t-ratio correspondence to $\beta_4$, $\beta_5$, and $\beta_6$ lead to direct tests as to whether these changes in marginal effects are significant.[9]

Nonadditive effects (i.e., interaction effects) can be permitted by including the variables $s_7 = s_2s_3 - s_4s_5$ and $s_8 = s_4s_5$. The coefficient $\beta_7$ measures the amount (times 20) by which the 50-30 plan deviates from the extrapolation of a plane passing through responses at the 75-30, 75-50, and 50-50 plans. Likewise, the coefficient $\beta_8$ measures the deviation of the 100-70 plan from the plane determined by the 75-50, 100-50, 75-70 plans. The inclusion of $s_7$ and $s_8$ implies that the representation of tax and guarantee rates is no longer simply a sum of two linear splines, but rather is a bilinear spline.[10] Bilinear splines have yielded interesting results when applied to age and education dimensions,[11] however, in most studies, the inclusion of $s_7$ and $s_8$ has contributed little in representing the tax- and guarantee-rates dimensions.[12]

In direct contrast to all of the preceding approaches, it may be argued that the treatment should be represented explicity in terms of dollar amounts. Some support for this view is offered by Knudson et al. [7] who found, based on the 13th Quarterly (Follow-up) Interview, that experimentals had little perception of their actual tax and guarantee rates. However, it must be remembered that this interview can be accused of being heavily dependent on the test-taking ability of the interviewees.

In any case, if the dollar value of the guarantee is believed to be more relevant than its ratio to the family's poverty level, $P_n$, then the transformed guarantee variables $s_2$, $s_4$, and $s_6$ for each observation can be multiplied by the $P_n$. Similarly, it may be argued that it is not the tax rate that is relevant, but rather the (dollar) amount by which hourly earnings are reduced. In this case $s_3$ and $s_5$ can be multiplied by an hourly earnings rate for the observation. Likely candiates for such hourly earnings rates are pre-enrollment wage rates and, better yet, the normal wage rate constructed by Poirier and Watts [10]. Use of the pre-enrollment wage suffers from not being available for those not working at pre-enrollment and from a likely large transitory component.[13] Since the Poirier-Watts normal wage rate is based on panel data, it is influenced less by transitory components. Futhermore, it was imputed for every head and spouse in the continuous 693 sample (whether or not they worked).[14] The most important aspect of both of these hourly wage measures is their freedom from any induced experimental effects. The same cannot be said for current wage rates.

Approaches that merely convert the treatment into a dollar amount are not equivalent to a "payments representation." The most obvious payments representation involves simply including the dollar value of the experimental payments as a regressor. However, if a labor supply variable, such as hours worked, is the dependent variable being considered, including actual payments on the right-hand side results in a simultaneity problem because actual payments depend in turn on hours worked.[15] This simultaneity problem can be avoided by using instead payments formula (1) with a family income estimate purged of experimental effects, e.g., pre-enrollment income, Watts's [17] estimated normal income, or Hollister's [3] estimated normal income.[16]

One advantage of such payments representations is that they identify experimentals over the break-even level who receive no payments--something that is camouflaged by the simple tax-guarantee rates representations However, the story isn't quite so simple. An over-break-even experimental receiving no payments is not quite the same as a control since such an experimental is eligible to receive payments if its family income drops sufficiently.[17] Often it is useful to differentiate between experimentals receiving payments, experimentals over break-even, and controls by constructing an over-break-even dummy.[18] Interacting such a dummy with other experimental variables permits different effects for the over-break-even group. For example, the guarantee effect for those over the break-even level is largely one of "security," while for those experimentals under the break-even, their payments actually depend upon the guarantee.

Rather than simply say that over-break-even experimentals may respond to the treatment, it may be desirable to say that their response depends upon how far above their break-even point they are. For example, in their analysis of male labor supply Watts et al. [20] constructed a variable $\Theta$ that embodied the assumption that a male head who is more than 20-hours-worth of work per week above his family's break-even point is "immune" from the effects of the experimental treatment and can be regarded as equivalent to a control observation. In other words $\Theta = 0$ for controls and those experimentals who could forego 20 or more hours of work at their normal wage without falling below their break-even point. Otherwise it is defined as

$$\Theta_{ij} = \frac{G_{ij}/t + 20\hat{W}_{ij} - \hat{Y}_{ij}}{10\hat{W}_{ij}} ,$$

where $G_{ij}/t = gP_{in_{ij}}/t$ is the dollar break-even level for the ith household (which depends of course on the family size $n_{ij}$) in period j, $\hat{W}_{ij}$ is the Poirier-Watts normal wage for the ith male head at time j, and $\hat{Y}_{ij}$ is the Watts' normal family income for the ith family at time j. The variable $\Theta$ will equal 2 for an observation with $\hat{Y}_{ij}$

precisely at the break-even level, and takes on higher positive values for cases that are below the break-even level.[19] By scaling $\Theta$ in terms of (tens of) hours it can be argued that it has greater comparability on an interpersonal basis than would the simple dollar amount of the gap between the family's income and break-even point.

Having defined $\Theta$, the treatment response function used by Watts et al. [20] is given by

$$X = (\alpha_{11} + \alpha_{12}s_2 + \alpha_{13}s_3)\Theta + (\alpha_{21} + \alpha_{22}s_2 + \alpha_{23}s_3)\Theta^2, \qquad (2)$$

where, as before, $s_2 = s_1(g - .75)$ and $s_3 = s_1(t - .5)$. As stated, the response function is homogeneous in $\Theta$, but of course this restriction can be removed by adding a constant term. The coefficients $\alpha_{11}$ and $\alpha_{21}$ give directly the coefficients for $\Theta$ and $\Theta^2$ when $g = .75$ and $t = .5$. Experience has shown that the quadratic term is important whenever there is a substantial response. Alternatively, replacing $\Theta^2$ $\max(\Theta - 2,0)$ a linear spline analogy to (2) can be formulated which has a "knot" at the break-even point $\Theta = 2$.

By taking the appropriate partial derivatives of (2) many different effects can be considered. The partial $dX/d\Theta$ is a "gap" effect; $(dX/dG)_{dt=0}$ is an income effect; $(dX/dG)_{d(G/t)=0}$ is a "pivot" effect; $(dX/dt)_{dG=0}$ is a price effect; and $(dX/dt)_{d\Theta=0}$ is a substitution effect. In the "pivot" effect the break-even point held constant. It shows the consequence of increasing the guarantee without changing the level of income at which benefits begin. The income-compensated substitution effect is derived under the constraint that benefits (B) remain constant under a change in the tax rate. Unfortunately, the estimated signs of these partial derivatives seldom conformed to theoretical expectations in the empirical work done by Watts et al. [14].[20]

While ignored up to this point in the discussion, still another treatment variable is "experimental time," i.e., the time that has passed since the beginning of the experiment at the particular site in question. Since special "start-up" and "termination" effects may be expected, often only the middle two years are considered by investigators.[21] On the other hand, the "spurious wage hypothesis" entertained by Poirier and Watts [14] and Watts and Mamer [19] was investigated in part by including experimental time explicitly in the model through the use of a natural cubic experimental time spline with interior knots at quarters two and six.[22] The rational for such a piecewise representation was once again special start-up effects. Entering experimental time explicitly into the model can be justified not only as a proxy for omitted variables, but also on a "learning-by-doing" basis.[23]

As if all the different types of parameterization discussed up to this point were not enough, their interaction with various socio-economic-demographic variables increases their number many-fold. In the past popular candiates for interactions have been age, earning capacity, education, ethnicity, health, home ownership, job characteristics, sex, site, and variance in normal income. Often it seems the question is not what to interact with the treatment, but what <u>not</u> to interact with the treatment. It seems reasonable to expect that future researchers will expand rather than narrow down the list.

FOOTNOTES

* The author is a Visiting Assistant Professor of Economics at the University of Wisconsin, on leave from the Department of Economics of the University of Illinois at Urbana - Champaign.

[1] Other such experiments, at various stages of completion, are the Rural Experiment (Iowa and North Carolina), the Gary Experiment, and the Seattle-Denver experiment.

[2] The actual eight tax-guarantee plans used are given on page 3. For the distribution across of experimentals from the husband-wife continuous sample of 693 families, see Rees [15, p. 169]; the distribution based on the 1309 families initially enrolled can be found in Kershaw and Fair [5, p. 157].

The annual poverty levels at the beginning of the experiment were as follows:

| Family Size (n) | Poverty Level ($P_n$) |
|---|---|
| 2 | $2000 |
| 3 | 2750 |
| 4 | 3300 |
| 5 | 3700 |
| 6 | 4050 |
| 7 | 4350 |
| 8 and over | 4600 |

These poverty levels were adjusted annually for changes in the consumer price index.

[3] The distribution of payments for the 693 sample, broken down by plan, site, ethnicity, and experimental time can be found in Rees [15, pp. 170-1].

[4] In such cases, the two groups are tied together only in the estimation of the error variance which should be based on all observations if the model is homoskedastic across the two groups. See Kmenta [6, p. 421] for more discussion.

[5] See Kmenta [6, p. 373].

[6] Approximately 90% of the experimentals on these two plans were either on welfare or had incomes above the break-even point. Hence, they are often omitted from the analysis. Measuring the effect of welfare on estimated experimental response is one the major problems in analyzing the data and is far too complicated to go into here. See Avery [1] and Garfinkel [2] for discussions.

FOOTNOTES (cont.)

[7]If the control group is viewed as belonging to a 0 - 0 plan and if it is desired that the experimental surface be continuous at the origin, then the experimental dummy should be omitted. However, controls aren't really faced with zero marginal tax rates, and so the validity such an approach is questionable.

[8]The 75-50 plan is close to the mean coordinates--approximately 85-50-- among experimentals in the 693 sample.

[9]For an empirical application see Poirier and Watts [14].

[10]See Poirier [10] and [13] for details.

[11]See Poirier and Watts [14] and Watts [17].

[12]The inclusion of $s_1$ through $s_8$ is equivalent to the insertion of eight dummies, one for each distinct treatment, in the sense that the estimated values at each plan-coordinate will be the same for both representations, The $R^2$ and other statistics will also be the same for each. These results follow from the fact that any linear combination of eight independent variables will fit an eight-dimensional space as well as any other combination of the same sort. As Watts [18, pp. BI-17 - BI-18] points out, the advantage of the bilinear spline sequence is that it provides for more interpretable intermediate coefficients (using less than eight degrees of freedom) and yields directly the more relevant tests of the presence of interactions and nonlinearities than the dummy representation.

[13]The use of any pre-enrollment variable as a proxy for a "normal" variable can cause special problems in panel data models since its approximation power is likely to diminish in later time periods. See Metcalf [8, CIII-29] and Hollister [3, III-1 - III-6], for more details.

[14]The appropriateness of such wage measures for those who work very little is of course not clear. In cases where the observational unit is the family, some sort of a weighted (possibly by "normal" hours worked) average of the head's and spouse's normal wage rates could be used.

[15]This problem does not arise when the dependent variable is, say, consumption.

[16]Still different approaches toward using dollar amounts could be formulated by using explicity the estimated experimental components from the Poirier - Watts [14] normal wage rate, or the normal income variable of Watts [17]. As of now neither of these approaches has been tried.

FOOTNOTES (cont.)

[17] See Metcalf [8, CIII-29 - CIII-30] for a further discussion.

[18] Depending on the context, there may exist many different candidates for the family income variable needed in constructing the over-break-even dummy e.g., current income, pre-enrollment income, and normal income. If it is desired that over-break-even status should not be experimentally induced, then current family income should not be used.

[19] A mean of five and a standard deviation of three are good round numbers to describe the distribution of $\Theta$.

[20] Somewhat similar to the $\Theta$ variable, Horner [4] introduces a treatment variable Q defined as follows:

$$Q = \frac{G}{2000\ \bar{\bar{W}}\ (1-t)} \quad ,$$

where G is the family's dollar guarantee, 2000 is the "normal" full-time, full-year work effort, $\bar{\bar{W}}$ is a "normal" wage variable developed by Horner [4], and t is the tax rate. Since $\bar{\bar{W}}(1-t)$ is the effective hourly wage, Q can be interpreted as the ratio of the basic hourly subsidy to the effective wage rate. This parameterization was derived from a rather restrictive utility maximization framework, and hence should be used with great caution.

[21] See for example Watts et al. [20].

[22] See Poirier [9] and [11] for details.

[23] The role of experimental time in the Seattle-Denver experiment is even more crucial since experimentals were placed on plans of varying time lengths. As a result, the piecewise spline representations should prove to be especially valuable tools for representing experimental time in anaylsis of these data.

REFERENCES


1. Avery, Robert. "The Effects of the Welfare 'Bias' on Family Earnings. Response to the Experiment." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, BVb-1 to BVb-36.

2. Garfinkel, Irwin. "The Effects of Welfare on Experimental Response." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, CII-1 to CII-70.

3. Hollister, Robinson. "The Labor-Supply Response of the Family." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. BVa-1 to IX-9.

4. Horner, David. "The Impact of Negative Taxes on the Labor Supply of Low-Income Male Family Heads." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, BIIb-1 to BIIb-38.

5. Kershaw, David N., and Jerilyn Fair, eds. Operations, Surveys, and Administration. Final Report of the New Jersey Graduated Work Incentive Experiment, Vol. IV. Madison: Institute for Research on Poverty, 1974.

6. Kmenta, Jan. Elements of Econometrics. New York: Macmillian, 1971.

7. Knudsen, Jon Helge; John Mamer; Robert A. Scott; and Arnold R. Shore. "Information Levels and Labor Response." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, BX-1 to BX-38.

8. Metcalf, Charles E. "Predicting the Effects of Permanent Programs from a Limited Duration Experiment." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, CIII-1 to CIII-64.

9. Poirier, Dale J. "Piecewise Regression Using Cubic Splines." Journal of the American Statistical Association, 68 (September 1973): 515-524.

10. ______________. "Technical Note on Bilinear Splines." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harlod W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, EI-1 to EI-7.

11. ______________. "Technical Note on Cubic Splines." Final Report of the New Jersey Graduated Work Incentive Experiment. Edited by Harold W. Watts and Albert Rees. Madison: Institute for Research on Poverty, 1974, EII-1 to EII-5.