THE RELATIVE ASSIGNMENT VARIABLE APPROACH TO SELECTION BIAS IN PRETEST-POSTTEST GROUP DESIGNS[1]

William M. K. Trochim, Cornell University
Clifford H. Spiegelman, Consultant[2]

The purpose of this paper is to explore a potential analytic solution, termed here the "relative assignment variable" approach, to the problem of selection bias in pretest-posttest group designs. The presentation includes a description of the most common pretest-posttest group designs and an explanation of why some of these are susceptible to selection bias problems. In addition there is a discussion of the relative assignment variable approach, and the illustration of this approach through simulations and the analysis of real data.

Pretest-Posttest Group Designs and Selection Bias. Pretest-posttest group designs have in common the existence of an observed pretest, x_i, and posttest, y_i, and can be distinguished by the manner in which persons are assigned to condition (e.g., program and comparison groups, program 1 and program 2, etc.) where the assignment is usually represented by the variable z_i (i.e., z_i = 1 if the research participant is in the program, 0 otherwise).

Four such designs are defined here in terms of their assignment strategies. The first of these is the true randomized experiment which is characterized by random assignment to condition. As a result, the groups which are created can, on the average, be assumed equivalent on the pretest (i.e., the expected value, $E(x_i)$, is equal for both groups). Although the true experiment does not require a pretest, the version of the design which is of interest in this paper includes one as a covariate to increase the statistical precision of the program effect estimate. The second of these, the regression-discontinuity design, is characterized by assignment on the basis of a cutoff score on the pretest. That is, all persons scoring on one side of a selected cutoff value are assigned to one condition with the remaining participants assigned to the other. The third of these, the "fuzzy" regression-discontinuity design, occurs when there is misassignment relative to this cutoff value. Here, some persons who should have been assigned by the cutoff to one condition are incorrectly included in the other. Finally, the non-equivalent group design allows for non-equivalence between the groups on the pretest and can occur when the random assignment in a true experiment is not maintained or when individuals or intact groups (e.g., classrooms, agencies, governmental units, etc.) are assigned to condition in a nonrandom manner. In fact, the distinction between the fuzzy regression-discontinuity and non-equivalent group designs is somewhat arbitrary (Reichardt, 1979). As will be shown later, data representative of both designs can be generated for simulation purposes by means of the same general models.

Pretest-posttest group designs are commonly chosen techniques for the evaluation of social programs. For information on the use of these methods one can consult Boruch et al (1978) who describe a number of randomized experiments, Cook and Campbell (1979) and Reichardt (1979) who speak of the utility of non-equivalent group designs, and Campbell (1969) and Trochim (1980) who discuss applications of the regression-discontinuity and fuzzy regression-discontinuity designs.

Despite an apparent conceptual similarity beween pretest-posttest group designs in terms of observed x_i, y_i and z_i, no general small sample analytic strategy is known to the authors. While the regression of y_i on x_i and z_i (and perhaps polynomials in x_i and interactions of x_i and z_i) will yield unbiased estimates for some of these designs, it cannot be used for all of them, at least in part because of the selection bias problem. The problem of selection bias has been viewed as specification error or omitted-variable bias by Barnow, Cain and Goldberger (1978) who state:

> "Selectivity bias addresses the question of whether there is some characteristic of the treatment (or control) group that is both associated with receipt of the treatment and associated with the outcome so as to lead to a false attribution of causality regarding treatment and outcome. So stated, selectivity bias is a version of omitted-variable bias, which is commonly analyzed under the rubric of specification error in econometric models." (p. 4)

Selection bias may affect program estimates when a variable related to z_i and y_i is not included in the analytic model. The development of a general analytic scheme which can compliment the conceptual similarity of pretest-posttest group designs is seen here then as dependent, at least in part, on an analytic solution to the selection bias problem.

The four designs outlined above can be examined for their potential for selection bias. If correctly implemented, the true randomized experiment is free of selection bias because assignment is random and independent of all pre-program measures. Within the context discussed in this paper, Goldberger (1972) has shown that the regression-discontinuity design is free of selection bias as long as assignment is adhered to and the underlying model is linear. This is because assignment is entirely based on a cutoff value on the pretest and any other measures related to assignment must, by definition, be related to the pretest (and will be accounted for in the regression of y_i on x_i and z_i by the presence of the pretest). In the fuzzy regression-discontinuity design assignment is not based entirely on the pretest alone. In fact, "fuzziness" can be defined as misassignment relative to what the pretest cutoff value would have dictated. As a result, the pretest does not perfectly account for assignment (as in regression-discontinuity) and the potential for selection bias exists. This is also the case for the non-equivalent group design. Here, assignment is not based on the pretest but rather on a judgment or determination of the pre-program "equivalence" of the groups. The extent to which the factors which determine the assignment to group affect z_i and y_i and are unaccounted for in the analytic model determines the potential for bias.

The need for an analytic solution to the selection bias problem in the fuzzy regressiondisctoninuity and non-equivalent group designs is especially apparent when one considers the frequency with which these two designs occur in practice. While both can occur in their own right, they also represent the degraded versions of the true regression-discontinuity and randomized experiment (i.e., versions where the assignment strategies for either are incorrectly implemented). Connors (1977) for example, points out the difficulties of adhering to random assignment in practice while Trochim (1980) describes the almost universal occurrence of misassignment relative to the cutoff value in implementations of the regression-discontinuity design within the context of compensatory education evaluation.

To summarize, in the true experiment and regression-discontinuity designs the assignment procedure is known and is perfectly accounted for by the inclusion of the pretest, x_i and the assignment variable, z_i in the analytic model. With the fuzzy regression-discontinuity and nonequivalent group designs assignment is not perfectly accounted for by regressing y_i on x_i and z_i and an analytic model based on these is likely to exhibit selection bias. The development of an analytic solution to the selection bias problem is seen here as a step towards unifying analytically this set of conceptually similar designs.

The Relative Assignment Variable Approach. Suppose as in Spiegelman (1976, 1977 and 1979) that x_i^*, v_i and q_i are unobserved variables where x_i^* denotes true ability, v_i denotes pretest random measurement error and q_i denotes posttest random measurement error. The data analyst and program evaluator observe x_i, y_i, and z_i which are related to the unobservables (for simplicity of exposition) by the equations

$$x_i = x_i^* + v_i$$

and

$$y_i = b_0 z_i + b_1 + b_2 x_i^* + q_i$$

where $z_i = 1$ if the research participant has received the program and 0 otherwise. In general terms, the approach to selection bias recommended here relies on an estimate of $E(z_i|x_i)$, which is termed the relative assignment variable, $\tilde{z}_i$, in place of z_i in the analytic model. Spiegelman (1976, 1977 and 1979) has shown that an appropriate estimate of b_0 based on an estimate of $E(z_i|x_i)$ is asymptotically unbiased under rather general conditions. Specifically, it is argued here that the regression of y_i on x_i and $\tilde{z}_i$ (instead of z_i) will yield unbiased estimates for common selection bias situations. The estimate, $\tilde{z}_i$ is not assumed to be related in any way to x_i or x_i^* except that it may not be perfectly colinear with x_i (i.e., $\tilde{z}_i \neq a_1 + a_2 x_i$).

It is useful to picture what $\tilde{z}_i$ is estimating. First, consider assignment in the true experiment. Here, $E(z_i|x_i) = .5$ for any given x_i, which is to say that for any given pretest value one expects on average about half the cases will be assigned to the program and half to the comparison group. In this case, the relative assignment variable can be described in relation to x_i by a horizontal straight line at $\tilde{z}_i = .5$ as shown in Figure 1. In these graphs, $\tilde{z}_i$ is on the vertical axis and can take values from 0 to 1 (i.e., none or all in the program group). The pretest values, x_i, are shown on the horizontal axis. Second, consider the regression-discontinuity design when assignment is "sharp" relative to a pretest cutoff value. Here, it might be that $E(z_i|x_i) = 1$ if x_i is less than or equal to the selected cutoff and 0 if it is greater. This step-function is shown in Figure 2. Finally, for fuzzy regression-discontinuity or the non-equivalent group design the relative assignment can be described by a function which ranges between the horizontal line of the true experiment and the step-function of the sharp regression-discontinuity design. Several functions of this type are sketched in Figure 3. It is clear that $\tilde{z}_i$ can be viewed as the estimated probability of assignment or as an estimate of the proportion of cases assigned to the program for any given pretest value.

Two methods for estimating relative assignment are offered here. The simplest and most straightforward can be termed the "assignment percentage" method. It can be calculated in two ways. With the first procedure, cases are ordered by their pretest values and divided into equal size pretest intervals. In the second procedure, cases are similarly ordered by the pretest but are divided into intervals having an equal number of cases. For both procedures, the percent of cases assigned to the program is calculated within each defined interval and then divided by 100 to yield values which range from 0 to 1. These values are then assigned to the individual cases within the intervals. Spiegelman (1976) has shown that for extremely large n estimates from both procedures will on average be equal.

The second method for estimating the relative assignment variable comes from the work of Spiegelman (1976, 1977 and 1979) and can be termed the nearest neighbor moving average method. Three steps are involved:

(1) The set of observations (x_i, y_i and z_i) are put in ascending order according to the pretest, x_i

(2) values of A and B are computed as the greatest integer part of:

$$A = n^{7/10}/2$$

$$B = n^{4/5}/2$$

(3) The relative assignment variable, $\tilde{z}_i$ (i.e., $E(z_i|x_i)$), $\tilde{y}_i$ (i.e., $E(y_i|x_i)$) and $\tilde{y}_i^2$ (i.e., $E(y_i^2|x_i)$) are estimated:

$$\tilde{z}_i = \sum_{i-A}^{i+A} z_i/2A$$

$$= \tilde{z}_A \text{ if } i < A$$

$$= \tilde{z}_{n-A} \text{ if } i > n-A$$

$$\tilde{y}_i = \sum_{i-B}^{i+B} y_i/2B$$

$$= \tilde{y}_B \text{ if } i < B$$

$$= \tilde{y}_{n-B} \text{ if } i > B$$

$$\tilde{y}_i^2 = \sum_{i-B}^{i+B} y_i^2/2B$$

$$= \tilde{y}_B^2 \text{ if } i < B$$

$$= \tilde{y}_{n-B}^2 \text{ if } i > B$$

$$\text{Then } \hat{\sigma}_y^2(x) = E(y_i^2|x_i) - (E(y_i|x_i))^2$$

$$= \tilde{y}_i^2 - (\tilde{y}_i)^2$$

Essentially, the procedure involves computing the moving average of the z_i's for cases ordered by x_i. The window for the moving average is of width 2A. Conditions are specified such that the A-1 values of $\tilde{z}_i$ at either end of the series are assigned the values of the first and last estimates having 2A observations. The estimate $\hat{\sigma}_y^2$ in the procedure is a weighting factor for the regressions. In this paper the assignment percentage and moving average procedures will be shown with and without this weighting.

A third procedure which might be useful for estimating relative assignment is suggested in the work of Maddala and Lee (1976) and Barnow, Cain and Goldberger (1978). Essentially it uses the maximum likelihood probit analysis of z_i on x_i to estimate relative assignment. There are three reasons for not including the probit analysis approach here. First, it is based on the assumption that the procedure on which assignment was based is known. This will often not be the case. Second, probit analysis is only appropriate here if the relative assignment variable follows the cumulative normal distribution. Finally, the probit approach tends to be more complex computationally than the other procedures.

To summarize, the relative assignment variable, $\tilde{z}_i$, is an estimate of $E(z_i|x_i)$, that is, an estimate of the probability of assignment to the program for any given x_i value. Two methods are suggested for estimating $\tilde{z}_i$, the assignment percentage and moving average approaches.

Illustrative Simulations. The relative assignment variable approach is illustrated here on simulated data. This requires constructing a pretest, x_i, a posttest, y_i, and an assignment variable, z_i. For the true or "sharp" regression-discontinuity design in the case of compensatory education (where the most "needy" student receives the program) the assignment might be represented as

$$z_i = 1 \text{ iff } x_i \leq x_0$$
$$= 0 \text{ otherwise}$$

where x_i is the pretest value for a given student and x_0 is the pretest cutoff value for assignment to the program. To generate data for the fuzzy regression-discontinuity or non-equivalent group designs one must assign using a variable which is not perfectly related to the pretest. The differences between these two designs is one of degree not of kind. To generate fuzzy regression-discontinuity data one can begin with true regression-discontinuity data and introduce slight misassignment in terms of the cutoff value, x_0. To generate non-equivalent group design data one allows greater amounts of misassignment thus leading to groups which are more nearly equivalent on the pretest. For convenience, the discussion presented here is phrased in terms of fuzzy regression-discontinuity rather than the non-equivalent group design.

Five models of misassignment are used to generate data for the simulations and are indicated by the symbols $z1_i$ to $z5_i$. To begin with, we generate a true score, x_i^*, such that $x_i^* \sim N(\mu, \sigma_{x_i^*}^2)$. In all runs, $\mu = 0$ and $\sigma_{x_i^*}^2 = 9$. In addition, we generate three error terms, v_i, q_i, and w_i such that each is normally distributed with variances equal to 1 or 4 units depending on the simulation. Here, w_i can be considered assignment error and v_i and q_i are pretest and posttest error, respectively. We can now construct a pretest, x_i, such that

$$x_i = x_i^* + v_i$$

Once we generate z_i using one of the five models described below we can construct a posttest, y_i, such that

$$y_i = b_0 z_i + x_i^* + q_i$$

where b_0, the program effect, is either 0 or 3 (i.e. the null case or a gain of three units). The five models used to generate z_i are

(1) Assignment by pretest plus independent assignment error:

$$z1_i = 1 \text{ iff } (x_i^* + v_i + w_i) \leq 0$$
$$= 0 \text{ otherwise}$$

(2) Assignment by true score:

$$z2_i = 1 \text{ iff } x_i^* \leq 0$$
$$= 0 \text{ otherwise}$$

(3) Assignment by true score plus independent assignment error:

$$z3_i = 1 \text{ iff } (x_i^* + w_i) \leq 0$$
$$= 0 \text{ otherwise}$$

(4) Assignment by true score and pretest

$$z4_i = i \text{ iff } x_i^* \leq 0 \text{ and } x_i \leq 0$$
$$= 0 \text{ otherwise}$$

(5) Assignment by true score intervals:

$$z5_i = 1 \text{ iff } x_i^* \leq -1.0 \text{ or } (.5 < x_i^* \leq 0)$$
$$= 0 \text{ otherwise}$$

For each of the five models of misassignment we use relatively low or high error variances (i.e., equal to 1 or 4) and a gain, b_0, of either 0 or 3 units. Thus we have 5(assignment models) X 2(gain) X 2(error variance) = 20 separate conditions. For each condition twenty independent simulations were carried out yielding a total of 20 X 20 = 400 runs, each based on 1000 individual cases (i.e., n=1000).

For each run the following general linear regression model was used to estimate the effect:

$$y_i = b_0 z_i' + b_1 + b_2 x_i + e_i$$

where

y_i = posttest for individual i

x_i = pretest for individual i

b_0 = parameter for program effect estimate

b_1 = parameter for intercept

b_2 = parameter for linear slope

e_i = residual $\sim N(0, \sigma_e^2)$

z_i' = real assignment (i.e., $z1_i$ ... $z5_i$) or estimate of $\tilde{z}_i$ as described below

For each run five analyses were conducted:

(1) Analysis using real assignment (i.e., $z1_i$ ... $z5_i$, depending on the simulation) in place of z_i'.
(2) Analysis using moving average estimate of $\tilde{z}_i$.
(3) Analysis using assignment percentage estimate of $\tilde{z}_i$.
(4) Weighted analysis using moving average estimate.
(5) Weighted analysis using assignment percentage estimate.

With the analysis based on real assignment we expect treatment estimates to be biased for all assignment models except for $z1_i$, assignment by pretest plus independent assignment error. In this case misassignment occurs randomly with respect to the pretest and will be reflected equally on the average in both groups. If the relative assignment variable approach successfully removes selection bias, the four analyses based on $\tilde{z}_i$ should yield unbiased estimates for all five assignment models.

The results are presented in Table 1 (b_0 = 0; low error variances) Table 2 (b_0 = 0; high error variances), Table 3 (b_0 = 3; low error variances) and Table 4 (b_0 = 3; high error variances). Each table presents, for all five assignment models and all five analyses, the average gain, the standard error of the average gain and the minimum and maximum obtained gain for twenty runs. Results will be considered biased if the true gain, b_0, lies outside the interval $\bar{b}_0 \pm 3SE(\bar{b}_0)$.

Several conclusions can be drawn from the tables. First, as expected, estimates from the analyses based on real assignment are biased except when misassignment is random. Second, the moving average estimates of relative assignment appear to yield unbiased estimates of gain for most of the models and conditions which were studied. Even for the three (out of twenty) sets of conditions where bias is detected two of these had average estimates which were not greatly biased, especially when considered relative to estimates from the analyses by real assignment. Third, it appears that estimates from the moving average analyses are in general less biased than the ones from the assignment percentage ones. This may be in part because the assignment percentage functions in these simulations are bases on only 50 intervals of only 20 z_i values each. Thus, the estimate of $\tilde{z}_i$ can only take on twenty values between 0 and 1 (i.e., 0, .05, .10 ... .95, 1.0) whereas the moving average estimate is more finely differentiated. Finally, the estimates yielded by relative assignment variable analyses appear to be less biased when error variances are low. It may be that with large sample sizes (i.e., larger than n=1000) and correspondingly greater statistical power, estimates would in general be unbiased. In fact, Spiegelman (1976, 1977 and 1979) has been careful to point out that the method is efficient only for large sample sizes.

Illustrative Real Data Analyses. Two sets of fuzzy regression-discontinuity data were constructed from the Third Grade Reading scores for a Title I compensatory education reading program in Providence, Rhode Island (Trochim, 1980). It is useful to apply the relative assignment variable approach to such data to see how the assignment functions differ from the simulations and to detect any unforseen difficulties in application. The linear model used in the simulations is applied here because visual inspection of the data indicates that a linear model may be appropriate and because there are relatively few program participant cases available for estimating changes in slope or function. Only the weighted and unweighted moving average analyses were carried out (in addition to analysis by real assignment) because the illustrative simulations indicate that they were less likely to exhibit bias than the assignment percentage estimates. In a previous analysis of data from this program where sharp regression-discontinuity data were used, the estimate of gain for the same linear model was $\hat{b}_0$ = 29.73 with a standard error of 6.12 (Trochim, 1980).

The first set of fuzzy data results from the use of the vocabulary subscale of the reading pretest rather than the total score. Assignment was sharp relative to the total score but is fuzzy relative to the subscale. The bivariate plot of the data is shown in Figure 4. Here, the analysis by real assignment, z_i, showed no significant gain ($\hat{b}_0$ = 11.13, SE($\hat{b}_0$) = 6.22) whereas the relative assignment variable analyses showed gains similar to the one found in the sharp regression-discontinuity case ($\hat{b}_0$ = 30.05, SE($\hat{b}_0$) = 10.13 for the unweighted moving average analysis and $\hat{b}_0$ = 29.43, SE($\hat{b}_0$) = .56 for the weighted moving average analysis).

The second set of fuzzy data is from the same program and results from the inclusion of the scores of children who come from schools in the district which were ineligible for service. Some of these students qualify for the program on the basis of their pretest score. The total reading score is used for the pretest and posttest and the bivariate distribution is shown in Figure 5. Here, all estimates of program effect are significant at the .05 level although the estimate from the analysis by real assignment appears smaller than the relative assignment estimates ($\hat{b}_0$ = 23.32, SE($\hat{b}_0$) = 5.60 for real assignment analysis, $\hat{b}_0$ = 47.82, SE($\hat{b}_0$) = 7.37 for unweighted moving average analysis and $\hat{b}_0$ = 48.77, SE($\hat{b}_0$) = .41 for the weighted moving average analysis).

Clearly, the results of analyses based on real assignment tend to differ from those based on relative assignment. Given that the former are likely to be biased and the latter are not (at least under the conditions specified here), one might place greater faith in the relative assignment analyses and conclude that this reading program had a positive effect.

Conclusions. While the relative assignment variable approach, especially using a weighted moving average analysis, appears in general to yield unbiased estimates in several models where selection bias is expected, there are still important unanswered questions. For example, it is not clear whether unbiased estimates will be obtained under more realistic or complex assignment models. Specifically, it is important to determine by simulations whether estimates are biased when the pretest-posttest relationship is nonlinear, when a wider variety of sample sizes are tested, and when misassignment occurs nearer the extremes of the pretest distribution. In addition, it is not yet clear whether the assignment percentage procedure yields biased results in general or whether the biases obtained here are related to sample size, interval size or other conditions chosen for these simulations. More definitive simulations than these illustrative ones require a greater number of runs for a wider variety of conditions.

It is reasonable to conclude that appropriate estimates of the relative assignment variable can be used to produce realistic estimates of program effect under many conditions where selection bias is expected. On this basis we might tentatively advance the outline of a more general analytic approach for pretest-posttest group designs. First, if the true randomized experiment or regression-discontinuity design are used and assignment has been implemented correctly the analysis may be based on the regression of y_i on x_i, z_i polynomials in x_i, interactions of x_i and z_i, and other appropriate covariates. Second, if the fuzzy regression-discontinuity or non-equivalent group designs are used or if the assignment procedures of a true regression-discontinuity or randomized design are not correctly implemented, an estimate of $\tilde{z}_i$, the relative assignment variable, can be used in place of z_i in the analytic model, at least as one part of a multiple analysis scheme (as described in Trochim, 1980) for estimating program effect.

Footnotes

[1]This work was supported in part by NSF Grant DAR 7820374 (Robert Boruch, Principal Investigator), NIE Grant G-79-0128 (Robert Boruch and David Cordray, Principal Investigators) and NSF Grant BNS 792-5577 (Donald T. Campbell, Principal Investigator).

[2]Currently with the National Bureau of Standards.

References

Barnow, B.S., Cain, G.C. and Goldberger, A.S. Issues in the analysis of selection bias. Unpublished manuscript, preliminary draft, August, 1978.

Boruch, R.F., McSweeny, A.J. and Soderstrom, E.J. Randomized experiments for program development and evaluation: An illustrative bibliography. Research Report NIE-01BMS, Methodology and Evaluation Division, Psychology Department, Northwestern University, Evanston, Illinois, 1977.

Campbell, D.T. Reforms as experiments. American Psychologist, 1969, 24, 409-429.

Conner, R.F. Selecting a control group: An analysis of the randomization process in twelve social reform programs. Evaluation Quarterly, 1977, 1, 195-244.

Cook, T.D. and Campbell, D.T. Quasi-experimentation: Design and analysis issues for field settings. Rand McNally: Chicago, 1979.

Goldberger, A.S. Selection bias in evaluating treatment effects: Some formal illustrations. Discussion Papers, #123-72. Madison: Institute for Research on Poverty, University of Wisconsin, 1972.

Maddala, G.S. and Lee, L. Recursive models with qualitative endogenous variables. Annals of Economic and Social Measurement, 1976, 5, 4, 525-545.

Reichardt, C.S. The statistical analysis of data from non-equivalent groups. In T.D. Cook and D.T. Campbell (Eds.), Quasi-experimentation: Design and Analysis Issues for Field Settings. Rand-McNally: Chicago, 1979, 147-206.

Spiegelman, C.H. Ph.D. Thesis, Northwestern University, 1976.

Spiegelman, C.H. A technique for analyzing a pretest-posttest nonrandomized field experiment. Statistics report M435, Florida State University, 1977.

Spiegelman, C.H. Estimating the effect of a large scale pretest posttest social program. Proceedings of the Social Statistics Section, American Statistical Association, 1979, 370-373.

Trochim, W. The regression-discontinuity design in Title I evaluation: Implementation, analysis and variations. Ph.D. Dissertation, Northwestern University, August, 1980.

Table 1

$b_0 = 0$, error variances = 1

Model	Analysis*	$\hat{b}_0$	SE($\hat{b}_0$)	min($\hat{b}_0$)	max($\hat{b}_0$)
$z1_i$	Real	.007	.027	-.293	.210
	MA	.023	.050	-.375	.545
	AP	.037	.043	-.306	.482
	MA(w)	.040	.055	-.394	.553
	AP(w)	.050	.051	-.342	.514
$z2_i$	Real	-1.171	.026	-1.356	-.918
	MA	-.051	.056	-.540	.283
	AP	-.185	.058	-.745	.251
	MA(w)	-.086	.058	-.552	.290
	AP(w)	-.167	.057	-.613	.288
$z3_i$	Real	-.922	.022	-1.105	-.728
	MA	-.178	.054	-.701	.316
	AP	-.350	.047	-.686	.067
	MA(w)	-.143	.058	-.685	.328
	AP(w)	-.265	.049	-.744	.129
$z4_i$	Real	-.623	.033	-.843	-.361
	MA	-.003	.033	-.248	.423
	AP	-.054	.032	-.283	.341
	MA(w)	.009	.038	-.239	.491
	AP(w)	-.032	.036	-.272	.431
$z5_i$	Real	-.979	.024	-1.177	-.771
	MA	.040	.059	-.316	.641
	AP	-.129	.055	-.437	.385
	MA(w)	.078	.065	-.405	.648
	AP(w)	-.054	.057	-.409	.429

Table 2

$b_0 = 0$, error variances = 4

Model	Analysis*	$\hat{b}_0$	SE($\hat{b}_0$)	min($\hat{b}_0$)	max($\hat{b}_0$)
$z1_i$	Real	.001	.054	-.422	.606
	MA	-.019	.187	-1.253	1.794
	AP	-.060	.148	-.969	1.664
	MA(w)	-.049	.184	-1.246	1.931
	AP(w)	-.074	.150	-.914	1.695
$z2_i$	Real	-2.715	.034	-2.979	-2.418
	MA	-.398	.183	-1.518	.946
	AP	-1.113	.142	-2.154	.131
	MA(w)	-.371	.183	-1.586	.866
	AP(w)	-1.031	.148	-2.143	.191
$z3_i$	Real	-1.685	.038	-1.985	-1.350
	MA	-.231	.266	-2.486	1.823
	AP	-1.020	.207	-2.654	.685
	MA(w)	-.163	.246	-2.563	1.751
	AP(w)	-.922	.207	-2.703	.620
$z4_i$	Real	-1.706	.045	-2.268	-1.423
	MA	.039	.080	-.722	.791
	AP	-.141	.081	-.954	.530
	MA(w)	.045	.083	-.631	.935
	AP(w)	-.137	.083	-.864	.677
$z5_i$	Real	-2.490	.054	-2.939	-2.157
	MA	-.143	.173	-1.335	.929
	AP	-1.020	.153	-2.331	.321
	MA(w)	-.110	.167	-1.356	.892
	AP(w)	-.945	.148	-2.220	.163

*Real=real assignment; MA=moving average; AP=assignment percentage; MA(w)=weighted moving average; AP(w)=weighted assignment percentage

Table 3

b_0 = 3, error variances = 1

Model	Analysis*	$\hat{b}_0$	SE($\hat{b}_0$)	min($\hat{b}_0$)	max($\hat{b}_0$)
$z1_i$	Real	2.992	.025	2.808	3.135
	MA	3.141	.048	2.747	3.535
	AP	3.025	.043	2.679	3.442
	MA(w)	3.126	.053	2.646	3.510
	AP(w)	3.017	.048	2.593	3.425
$z2_i$	Real	1.860	.021	1.686	2.079
	MA	2.995	.047	2.659	3.380
	AP	2.804	.049	2.393	3.209
	MA(w)	2.995	.049	2.653	3.355
	AP(w)	2.830	.048	2.435	3.199
$z3_i$	Real	2.086	.027	1.785	2.225
	MA	2.976	.066	2.316	3.485
	AP	2.758	.053	2.257	3.160
	MA(w)	2.976	.062	2.365	3.532
	AP(w)	2.783	.048	2.277	3.055
$z4_i$	Real	2.425	.034	2.012	2.705
	MA	3.184	.041	2.812	3.564
	AP	2.969	.039	2.616	3.278
	MA(w)	3.158	.040	2.790	3.572
	AP(w)	2.975	.038	2.626	3.323
$z5_i$	Real	1.992	.023	1.862	2.231
	MA	3.098	.058	2.599	3.431
	AP	2.908	.049	2.479	3.199
	MA(w)	3.117	.063	2.588	3.648
	AP(w)	2.960	.056	2.553	3.431

Table 4

b_0 = 3, error variances = 4

Model	Analysis*	$\hat{b}_0$	SE($\hat{b}_0$)	min($\hat{b}_0$)	max($\hat{b}_0$)
$z1_i$	Real	2.908	.056	2.203	3.356
	MA	2.944	.148	1.844	4.146
	AP	2.874	.122	1.867	3.855
	MA(w)	2.951	.157	1.851	4.273
	AP(w)	2.876	.127	1.889	3.936
$z2_i$	Real	.401	.047	-.134	.785
	MA	2.969	.130	1.930	4.131
	AP	2.007	.102	1.304	2.956
	MA(w)	3.017	.125	2.047	4.261
	AP(w)	2.134	.098	1.437	3.047
$z3_i$	Real	1.277	.043	.894	1.508
	MA	2.001	.271	.039	4.764
	AP	1.553	.168	.363	3.016
	MA(w)	2.038	.266	.047	4.697
	AP(w)	1.596	.176	.351	3.335
$z4_i$	Real	1.343	.051	.869	1.821
	MA	3.139	.081	2.248	3.869
	AP	2.841	.081	1.814	3.627
	MA(w)	3.164	.084	2.242	3.936
	AP(w)	2.892	.083	1.847	3.713
$z5_i$	Real	.524	.038	.235	.900
	MA	2.568	.165	1.001	3.895
	AP	1.901	.126	.908	2.806
	MA(w)	2.650	.183	1.094	4.229
	AP(w)	2.019	.141	1.032	3.161

*Real=real assignment; MA=moving average; AP=assignment percentage; MA(w)=weighted moving average; AP(w)=weighted assignment percentage

Figure 1
Relative Assignment Variable Function for the True Experiment

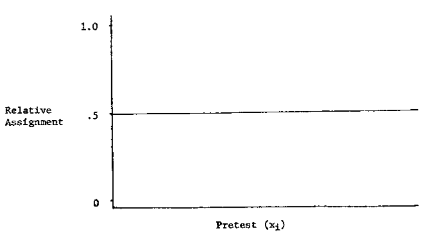

Figure 2
Relative Assignment Variable Function for Regression-Discontinuity

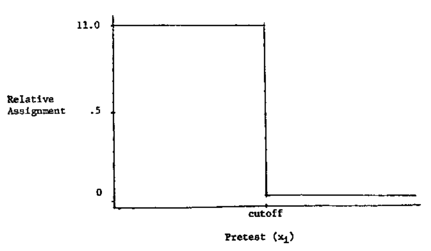

Figure 3

Relative Assignment Variable Functions for Fuzzy Regression-Discontinuity and the Non-equivalent Group Design

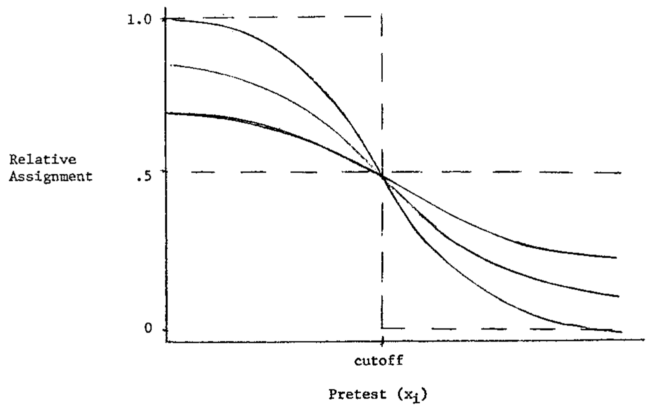

Figure 4

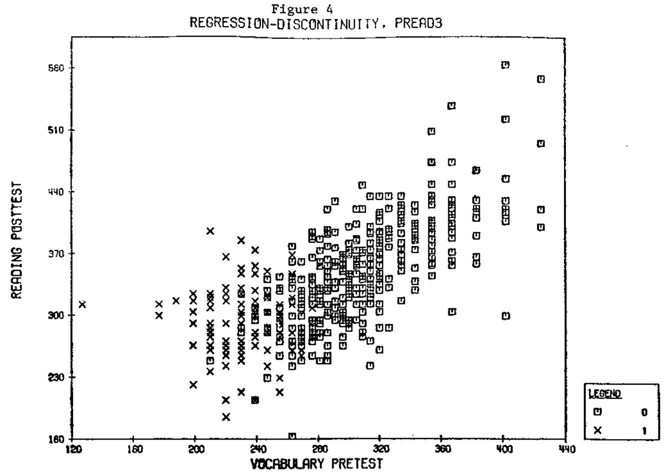

Figure 5