

# *A Political Scientist's View of the Income Maintenance Experiments*

---

*Richard F. Elmore\**

The negative income tax experiments were designed to influence political debate on income support policy in at least two ways. One was methodological — to focus the policy debate on a few key empirical questions and produce more definitive evidence than would have been available through nonexperimental methods. Another was political — to legitimize the idea of a universal cash transfer program, scaled to the recipient's income, as an alternative to the patchwork collection of in-kind and categorical assistance programs that had grown up since the New Deal.

## *Methodological Issues*

Social experiments are thought to hold certain advantages over nonexperimental policy research. They frame policy questions in more precise terms. They permit more precise statistical tests of effects. And they introduce a dimension of greater empirical discipline to discussions of policy issues.

These advantages are, however, purchased at some cost. The very rigor of social experiments limits the utility of their results. First, experimental treatments are, of necessity, packages of discrete elements — in the negative income tax experiments, various combinations of guarantee levels and marginal tax rates. Inferring the effects of treatments not represented in the experiment from experimental results

---

\*Professor of Education, Michigan State University and Senior Research Fellow, Center for Policy Research in Education.

requires extrapolation between or beyond treatments. This extrapolation is no more rigorous than most nonexperimental research. Second, treatments are, of necessity, implemented in specific sites, which means that experimental results have to be combined with nonexperimental survey data to estimate the effects for the population as a whole and to estimate the cost of generalizing the treatment. Combining experimental and nonexperimental results reintroduces most of the methodological tangles that experiments were designed to eliminate. Finally, experiments take place at particular times, against a background of particular policies and particular economic, political, and social conditions that influence the responses of subjects. These background factors cannot be controlled experimentally. Hence, it is always problematical whether experimental results gained under conditions at Time 1 will generalize to a different set of conditions at Time 2.

These methodological problems are well-known, at least to those familiar with social experiments. They have been addressed with impressive ingenuity and technical virtuosity in the various analyses and reanalyses of the negative income tax experiments, including those reported in this volume. On balance, though, the kind of precision the experiments have supplied may be of doubtful utility to policymakers. As the experiments have matured, a kind of infinite regress into methodological and theoretical complexity distances the results from the concerns of policymakers.

Imagine a member of Congress innocently asking, "How will an income guarantee, based on the principles of the negative income tax, affect the likelihood that its recipient will work, the amount of work that a recipient will do, or the amount of income that a recipient will earn?" An answer to this question that is consistent with the experimental evidence would have to be qualified in at least the following ways: (1) the nature of benefit packages available to experimental subjects was limited; (2) benefit packages available to control groups at the time of the experiment varied unsystematically from one setting to another; (3) the effects of an income guarantee on labor supply were different for different population groups within experimental settings; (4) different experiments produced quite different estimates of effects; (5) misreporting of income and work by experimental subjects influenced the results in ways that cannot be fully explained; (6) benefit packages available to poor people have changed since the experiments, making it difficult to estimate the effect and cost of changing the existing system; (7) little correspondence exists between the benefit packages tested in the experiment and the range of benefit packages that could be made available under a feasible reform proposal; and (8) the population of potential beneficiaries of an income guarantee would be significantly different from the population tested in the experiment.

To be sure, the answer could be simplified considerably by trimming off these methodological uncertainties, and by averaging results across experiments. To do so, however, sidesteps the methodological rationale for doing experiments in the first place. Being precise, it seems, also means being complex and equivocal. Methodological precision is apparently not the same as clarity or policy relevance.

Social science researchers are accustomed to these inward-turning spirals of ever-increasing complexity. This is the stuff of which social science reputations are made. The main product of research is, after all, proposals for more research. Elected officials may be forgiven, however, if they do not share this enthusiasm, since they operate under different time constraints and different incentives.

One of the major lessons of the negative income tax experiments, then, is that methodological precision is not positively related to the clarity or policy relevance of results. This is hardly a novel finding,<sup>1</sup> but it bears repeating before the promises for the next round of social experiments are made.

### *Political Issues*

In addition to policy research, the negative income tax experiments were also political advocacy. The New Jersey Income Maintenance Experiment was conceived, in 1966-67, by a group of policy analysts in and around the Office of Economic Opportunity as a way of forcing the negative income tax into the political debate on income support.<sup>2</sup> To be sure, more politically astute ways exist of legitimizing a novel policy idea than to run a costly, long-term experiment with indeterminate results. But the advocates of the negative income tax were social scientists, not politicians, and they did what social scientists do best — they wrote a research proposal.

Over time, the experiments expanded and were accompanied by related research on poverty, spawning a significant analytic group within the Department of Health, Education, and Welfare and an enormous contract research industry outside the government. This analytic subgovernment became the main source of continuity in advocacy of the negative income tax during the welfare reform debates that followed.

Variants on the negative income tax found their way into the presidential or congressional arena on at least five occasions:

- In 1965 and 1966, the Office of Economic Opportunity presented President Johnson with proposals for a universal negative income tax. The proposals fell to budget pressures and the opposition of Undersecretary (later Secretary) Wilbur Cohen of the Department of Health, Education, and Welfare.<sup>3</sup>

- In 1969, President Nixon proposed the Family Assistance Plan, which combined an income guarantee with food stamps. The Plan passed the House in 1970, languished in the Senate, was reintroduced in 1971, failed again, and was abandoned by Nixon in 1972.<sup>4</sup>
- In 1974, Congresswoman Martha Griffiths, Democrat from Michigan, proposed an income maintenance plan consolidating several existing cash and in-kind programs. The proposal gained the endorsement of Griffiths' Subcommittee on Fiscal Policy of the Joint Economic Committee, and was the subject of hearings in both houses in 1976, but did not progress further.<sup>5</sup>
- At the same time as the Griffiths proposal, Health, Education, and Welfare Secretary Caspar Weinberger proposed an income-guarantee reform to President Gerald Ford, who rejected it in favor of more modest incremental changes.<sup>6</sup>
- In 1977, the Carter administration, after a period of bruising intramural combat between the Departments of Labor and Health, Education, and Welfare, introduced the Program for Better Jobs and Income which combined income supplements, earned income tax credits, and public jobs, in a supposedly zero-cost, comprehensive reform. The proposal never reached the floor of either house of Congress. The reasons offered for this failure included opposition from organized interests, conflicting Administration priorities, and budget constraints.<sup>7</sup>

The preparation of each of these proposals entailed considerable analytic staff work, which benefited, no doubt, from the existence of a well-staffed analytic subgovernment. In this sense, the federal government's investment in income support policy research paid off.

The published record shows, however, that the results of the negative income tax experiments entered the policy debate explicitly only twice. The first occasion was in early 1970, when the Office of Economic Opportunity, under pressure from Congress, released preliminary results from the New Jersey experiment less than two years after its commencement, showing what it claimed was a negligible impact of a negative income tax on labor supply. This report was immediately rebutted by a General Accounting Office study that labeled the results "premature."<sup>8</sup> The second occasion was in late 1978, when Senator Daniel Patrick Moynihan, Democrat from New York, announced in a speech on the Senate floor that evidence of high rates of family dissolution among recipients in the Seattle-Denver experiment had caused him to question his earlier advocacy of a negative income tax.<sup>9</sup> Neither of these occasions captures what policy researchers have in mind when they think about the use of research in policymaking. Both, however, demonstrate the political uses of evidence and the limited

value of methodological sophistication in clarifying policy choices.

Another striking feature of political debate on income support is the relative infrequency with which empirical estimates of effects, of the type represented by the negative income tax experiments, figure in decisionmaking. Most of the debate on the Family Assistance Plan and on the Program for Better Jobs and Income, both within the executive and between the executive and Congress, focused on what might be called the architecture of proposals — the way various existing programs and their disparate benefit levels could be melded into a defensive reform package — and on the winners and losers created by various alternatives to the existing system. On occasion, the debate required estimates of population parameters and total costs. If these could be supplied by the analytic subgovernment, so much the better. If they could not, that was OK too — any reasonable estimate would suffice. In other words, policymakers were mainly concerned about the incremental effect of changes, on the architecture of benefit programs and on winners and losers. They were not overly constrained by the lack of solid causal evidence on the issues regarded as important to designers of the negative income tax — notably labor supply and family structure.

Judged in terms of their direct effect on policy, then, the influence of the experiments has been modest-to-negligible. But direct effects are deceptive. If we have learned nothing else from two decades of systematic research in the service of policymaking, it is that — to use Carol Weiss's terminology — the "decision-drive" model of policy research is less accurate descriptively than the "climate" model.<sup>10</sup> That is, research influences policy not by marshalling specific evidence in support of specific decisions, but rather by shaping policymakers' perceptions of the relevant problems and the feasible range of solutions.

This view tracks with John Kingdon's research on political decisionmaking, which concludes that "Academics, researchers, and consultants affect the alternatives more than the agenda, and affect long-term directions more than short-term outcomes."<sup>11</sup> In Kingdon's view, the formulation of a working agenda of policy problems, the formulation of alternative solutions to those problems, and the politics of decisionmaking normally operate as three separate, more or less autonomous processes. Only on rare occasions do these three streams converge into what Kingdon calls "policy windows," or opportunities for major changes in policy.<sup>12</sup>

Viewed from this perspective, the negative income tax experiments, and related policy research on poverty and income support, have had about the effect one would expect. The analytic subgovernment that grew up around the negative income tax is one of a few select locations for stockpiling policy options — or "option depots" — in the event that the problem and decisionmaking streams converge. This subgovern-

ment has recently been joined by other, less social science oriented, more ideological option depots. When the problem-identifying and decisionmaking streams occasionally converge, as they have on four or five occasions over the past 20 years, these option depots supply some of the raw material for policy debate. This raw material gets unpacked and repackaged with other raw material in a variety of ways during the policy debate. The results are, for the most part, horrifying and depressing to the personnel of option depots, who see their careful work being defiled by amateurs and dilettantes.

This view also explains why the negative income tax experiments have had such a modest effect on income support policy and why the clear-cut vision of reform offered by their advocates is constantly messed up by policymakers. First is the issue of timing. The experiments ran on what might be called "social science time," or S-time. It has taken close to 20 years to get the experiments up and running, to accumulate and analyze the results, and to subject those results to the kind of critical scrutiny and secondary analysis required to tease out their strengths and limits. Even then, researchers have argued that the treatments were not in place long enough to give useful information about their long-term effects on labor supply. The policy process, on the other hand, runs on "political time," or P-time. The key determinants of P-time, for purposes of income support reform, are presidential elections and annual budget cycles. The calendar on P-time is shorter and more compressed than on S-time. On the few occasions when income support policy reform has surfaced in P-time, the results of the income support experiments have not been sufficiently mature in S-time to provide useful guidance to policymakers.

Second, the process of problem identification and political agenda-building contains much intelligence of importance to policymakers that is considered to be random noise by policy researchers. Advocates of the Family Assistance Plan and the Program for Better Jobs and Income, for example, consistently ignored signals from individual members of Congress and political advisers within the executive branch that "guaranteed income" and "negative income tax" were terms that carried very problematical overtones for certain key members and their constituencies. After the Family Assistance Plan proposal, the negative income tax became anathema to Al Ullman, Chairman of the House Ways and Means Committee, because of its extreme complexity, its potentially enormous cost implications, and its association with an earlier political debacle. Policy researchers found this aversion to be irrational and nonsensical, because after all, one could demonstrate analytically how existing benefit programs contained perverted and inferior versions of the same basic elements as the negative income tax. In other words, for policy analysts and researchers, welfare reform was the

deliberate design of an income support system around certain policy variables with certain outcomes specified; for elected officials, welfare reform was the repackaging of programs so as to create a winning coalition. In the debate about the Program for Better Jobs and Income, analysts from the Department of Health, Education, and Welfare continued to serve up negative income tax-like options in the face of mounting evidence that key members couldn't stand them. The analysts were laboring in their own option depot, defending their product, disdainful of the difficulties their work was causing in the political arena.

Third, the process of political decisionmaking usually works in ways that are upside down, backwards, or perpendicular to the proposals of policy researchers. In the debates, for example, both President Nixon and President Carter stipulated certain budget and policy constraints early in the decisionmaking process that forced executive staff to make serious compromises in the architecture of the plans. The Family Assistance Plan contained a budget constraint and a stipulation that food stamps had to be treated separately and not cashed out. The Program for Better Jobs and Income contained a zero-cost constraint and the forced marriage of the Health, Education, and Welfare and Labor Departments through the amalgamation of cash assistance and jobs. In both these exercises, one could have produced a "better" plan — in both political and analytic terms — without the constraints imposed by presidential leadership, but the Presidents insisted on them.

Finally, policy research tends to focus, as it should, on the effects of options on recipients. In political decisionmaking, though, the relevant units of analysis also include organized interest groups and state and local governments. In both the Family Assistance Plan and the Program for Better Jobs and Income, the architecture of the benefit packages could be defended for recipients overall, but the plans collapsed under scrutiny of their effects on recipients on the margins of existing programs and their effects in high benefit states and localities. The intended beneficiaries of income support policy are not, for the most part, the interests that have the most influence in the formulation of that policy.

These conditions of political decisionmaking mean that large-scale policy research efforts of the type embodied in the negative income tax experiments will inevitably have a very limited influence on policy. The notion that social experiments could be used to leverage income support policy in the direction of a negative income tax has been proved by experience to be patently absurd. The problems of timing, or marshalling the right political intelligence for the right moment, and of building political constituencies around specific proposals far outrun the complexity of designing and running an experiment. The notion of deliberate investment in option depots and of the careful stockpiling of evidence on the complex and equivocal effects of reform policies is not

absurd, however, if it is accompanied by low expectations that these options and evidence will have a direct effect on policy.

---

<sup>1</sup>Cohen, David and Michael Garet, "Reforming Educational Policy with Applied Research," *Harvard Educational Review*, 45 (1975) pp. 17-43; Lindblom, Charles and David Cohen, *Usable Knowledge*, New Haven, CT: Yale University Press, 1979.

<sup>2</sup>Williams, Walter, *The Struggle for a Negative Income Tax*, Seattle, Washington: University of Washington, Institute of Government Research, 1972, pp. 2-11.

<sup>3</sup>*Ibid.*, pp. 4-5.

<sup>4</sup>Moynihan, Daniel Patrick, *The Politics of a Guaranteed Income*, New York: Random House, 1973; Lynn, Laurence and David Whitman, *The President as Policymaker: Jimmy Carter and Welfare Reform*, Philadelphia, PA: Temple University Press, 1981, pp. 18-34.

<sup>5</sup>Lynn and Whitman, pp. 42-43.

<sup>6</sup>*Ibid.*

<sup>7</sup>*Ibid.*, passim.

<sup>8</sup>Williams, op. cit., pp. 23ff.; Moynihan, op. cit., pp. 191ff.

<sup>9</sup>Lynn and Whitman, op. cit., pp. 247ff.

<sup>10</sup>Weiss, Carol, "Research for Policy's Sake: The Enlightenment Function of Social Research," *Policy Analysis*, 3 (1979), pp. 531-545; "Improving the Linkage Between Social Research and Public Policy," in Laurence Lynn, ed., *Knowledge and Policy: The Uncertain Connection*, Washington, DC: National Academy of Sciences, 1978, pp. 23-81.

<sup>11</sup>Kingdon, John, *Agendas, Alternatives, and Public Policies*, Boston: Little, Brown, 1984, p. 71.

<sup>12</sup>*Ibid.*, pp. 92-93 and 173ff.

## *Discussion*

---

*Robert D. Reischauer\**

Welfare reform has always presented policymakers with two related, but distinct, problems. The first has been the technical problem of designing a reform package that could achieve the desired results within the budget constraint. The second has been the political problem of creating a constituency that could get the reform package enacted.

This second problem exists because, unlike farm, housing or education policy, welfare policy has been neither an area popular with the public nor an area that provided much political payoff for politicians. The potential beneficiaries of welfare reform have never been politically active and, therefore, have had little clout in Washington; and the public has held a generally hostile view of welfare programs. When politicians have supported welfare reforms, they have not done so because they anticipated PAC contributions from welfare rights organizations or long lines of appreciative recipients at the voting booths. Rather they have supported such policies because they thought they were the right thing to do and because they concluded that the reforms would not generate a backlash from the voters.

Given the hostile political environment that exists for welfare reform, an interesting policy question is whether the income maintenance experiments could have helped to build a constituency for reform. Could the experiments have reduced the public's distaste for radical reform or mobilized the low-income population into an effective interest group in behalf of change?

The simple answer to these questions is "no." In fact, there were good reasons to expect that the experiments would, if anything, strengthen the hand of those opposed to reform. Part of the reason for

---

\*Senior Fellow, The Brookings Institution.

this negative conclusion lies in the nature of the negative income tax and part lies in the particular focus of the income maintenance experiments.

Richard Elmore has argued that the political objective of the experiments was to legitimate the negative income tax concept — to give it standing with policymakers. An essential first step in this process was for the Office of Economic Opportunity technocrats to convince the White House and congressional policymakers that a negative income tax was a technically feasible policy option. The income maintenance experiments played a modest role in doing this. While the results of the experiments may have entered the public record of the policy debate only twice, their contribution was more significant than Elmore suggests. The findings of the experiments were discussed at dozens of planning and strategy meetings between congressional advocates of welfare reform and policy formulators in the executive branch. They influenced the design of the Carter welfare reform plan in numerous ways.

However, the major political hurdle was not to convince the policymakers that a negative income tax was technically feasible but rather to convince the American public that this radical approach to the poverty problem was acceptable. And the experiments were not capable of doing this, primarily because the negative income tax was designed to address the deficiencies that the policy elite saw in the existing welfare system, not the shortcomings that most concerned the general public. The public felt that welfare costs were too high and that caseloads were expanding too rapidly; they imagined that undeserving freeloaders who were capable of work were weaseling their way onto the rolls in increasing numbers; they feared that the system was encouraging marital instability; they suspected that cash assistance was being squandered on booze, color TVs and other unnecessary expenditures; and they felt that welfare was creating a permanent dependent class. The policy elite, in contrast, was more concerned with the inequities created by interstate differences in payment levels and eligibility requirements, the failure to provide assistance to two-parent families, the general inadequacy of benefits, the stigma associated with receipt of welfare, and the political contentiousness that surrounded discussions of income redistribution policy.

On many dimensions, a negative income tax was bound to exacerbate the very aspects of the existing welfare system that most concerned the public. For example, a negative income tax would increase costs and welfare rolls, provide assistance to more able-bodied adults, reduce popular in-kind benefits in favor of more cash assistance, and allow families to remain on welfare for life. In addition, it should be noted that a true negative income tax would cut away what little there was in the way of a mainstream political constituency for welfare programs, the providers of in-kind benefits: the farmers who benefited from the food

stamp program; the real estate interests who gained from the various housing subsidy programs; the professionals who provided the various social services; and so on. Overall, a negative income tax clearly went against the popular grain and, therefore, would be difficult to legitimate.

Given this situation, the political task for the experiments was to provide information to convince the public that they would like the results of the new system even if they did not like the manner in which a negative income tax would provide assistance. In other words, the experiments would have to highlight those aspects or responses to a negative income tax that might attract the support of the nonrecipient public. This strategy would have suggested emphasizing its impact on family stability, crime and delinquency, health status, nutrition, school achievement, and other effects that would indicate that the recipients of a negative income tax ultimately could attain self-sufficiency.

While the income maintenance experiments did examine such responses, these dimensions were of secondary interest. The experiments focused on the measurement of the labor supply responses to a negative income tax. Inevitably, this focus made the political problem worse because any negative labor response would constitute a political liability. While the policy experts might have been overjoyed to find that the reduction in labor supply was small, a skeptical public and congressional critics would not differentiate between a 5 percent and a 20 percent reduction in labor supply. In both cases, more indolence was being rewarded at the taxpayer's expense.

The experiments were not only incapable of allaying the public's apprehensions concerning radical welfare reform, they were also not capable of building a low-income constituency for reform. Because the experiments involved only a tiny fraction of the low-income population, most of the potential recipients of a negative income tax knew nothing about the options that were being tested. Moreover, to the extent that politicians used the experimental results to generate support for a negative income tax among low-income groups, they ran the risk of alienating the general public. A reformed system that reduced stigma, intruded less into the lives of the poor, provided more generous benefits, and offered assistance to families that previously had been ineligible might rally the low-income population, but would lose the middle class.

The general conclusion that arises from examining the experience of the income maintenance experiments is that, under most circumstances, social experiments have very limited political utility. Policy analysts may find them a useful way to convince politicians that a certain policy is technically feasible. But in doing so they are likely to focus attention on the politically unpopular behavioral responses to the proposed policy.

Imagine what might have happened if a social experiment had been

mounted in the 1920s to test the feasibility of a government-subsidized old-age pension system. The results from this experiment would have revealed that a large proportion of the elderly would drop out of the labor force under the proposed pension program. While some of these retirees would have health problems that made it difficult for them to meet the physical demands of their jobs, many would be quite capable of continued labor force participation. The experiment's results would also have shown that the proposed program would induce some workers to drop out of the labor force even before they were eligible to receive a pension. Would this be rewarding indolence?

Careful analysis of the experiment's effects would also have revealed that the new program would lead to a reduction in private retirement saving and a slight tendency for employers to cut back on their pension programs. In other words, only a portion of the federal pension payment would represent a net increase in the living standards of the retirees.

The sociologists and psychologists analyzing the results of the experiment would have discovered that the proposed program would threaten accepted family patterns. Many of the affected elderly would choose to move out of their children's homes, some to small squalid apartments. A few of the elderly would even move to distant places like Florida where their children could not care for them. By all measures, the proposed program would lead to a reduction in contacts between parents and their children and a decrease in the sense of responsibility that children would feel for their aged parents.

These results, while not surprising, would have represented a political liability for advocates of social security because they would have clashed with prevailing values and behavioral norms. But major social reforms, like social security or a negative income tax, inevitably will change prevailing values, behavioral norms and the political environment. In prospect, many of these changes may appear threatening and therefore undesirable. After some years, however, they will become not only accepted but also desirable attributes of modern living. For this reason, advocates of major social policy changes should think twice about the desirability of experimentation. What is gained in the way of an understanding of the micro effects of a proposed policy may be lost in the political realm.