CHAPTER 6

BIG Experiments of the 1970s and the Public Reaction to Them

Abstract This chapter discusses the five Negative Income Tax experiments conducted in the 1970s in the United States and Canada, summarizes their findings, and shows how badly they were misunderstood at the time. It argues that although the experiments succeeded in the scientific goal of obtaining useful data, they badly failed in the goal of improving public understanding of the issue. This experience provides extremely important lessons for the current round of basic income experiments.

Keywords SIME/DIME • Income maintenance experiments • Mincome • Basic income experiments • Negative Income Tax experiments • Social science experiments • Basic income • Universal Basic Income • Inequality • Poverty

Between 1968 and 1980, the US and Canadian governments conducted five NIT experiments. They got started when what I've called the second wave of the UBI movement was at its height. The United States had declared "War On Poverty." Civil rights activists were turning their attention to poverty and inequality. The United States was rethinking its welfare system with an eye to expanding and improving it. All of this created a strong interest in BIG, especially in the form of the NIT, but UBI (under various names) was also in the public discussion in the era. The last of

K. Widerquist, *A Critical Analysis of Basic Income Experiments for Researchers, Policymakers, and Citizens*, Exploring the Basic Income Guarantee, https://doi.org/10.1007/978-3-030-03849-6_6

these experiments wound down and their results came out at a time when expanding and improving the welfare system was much less popular.[1] This political context probably had a significant effect on the experiments and the reception of their results. Lessons from these experiments affect the argument throughout this book.

1 Labor Market Effects of the NIT Experiments of the 1970s

Unfortunately, most of the attention of the 1970s experiments was directed not at the effects of the policy (how it affects the welfare of net beneficiaries), but to one potential side effect (how it affects the labor time of test subjects). And so that issue takes up most of the discussion here. This section draws heavily on an earlier work, entitled, "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments."[2]

Table 6.1 summarizes the basic facts of the five NIT experiments. The first four columns show the name of the experiment, where it was conducted, the years it ran, and its sample size, usually showing how much it decreased due to dropouts. The specifications of each experiment varied considerably and so the last three columns summarize information about the makeup of the people being studied, the grant level, and the marginal tax rate.

The largest NIT experiment was the SIME/DIME. The main study was conducted from 1970 to 1976 for most participants, but a small subsample (discussed in Chap. 4) continued to receive the grant until 1980.

The Canadian government initiated the Manitoba Basic Annual Income Experiment (Mincome) in 1975 when the US experiments were winding down. It was the only experiment to include a saturation study (along with an RCT). At the time of writing, Mincome remains the only BIG saturation study conducted in a higher-income nation. Disappointingly, by the time data collection was completed in 1978, interest in the guaranteed income was seriously on the wane and the Canadian government cancelled the project in 1980 before the data was fully analyzed. It would be decades before researchers would go back to it.

Scholarly and popular media articles on the NIT experiments focused, more than anything else, on the NIT's "work-" or "labor-effort response"—the comparison of how much the experimental group worked relative to the control group. Table 6.2 summarizes the findings of several

[1] Widerquist, "Three Waves of Basic Income Support."

[2] "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments?"

Table 6.1 Summary of the Negative Income Tax experiments in the United States and Canada

Name	*Location(s)*	*Data collection*	*Sample size: initial (final)*	*Sample characteristics*	*Grant level*[a]	*Marginal tax rate*[b]
The New Jersey Graduated Work Incentive Experiment (NJ)	New Jersey and Pennsylvania	1968–1972	1216 (983)	Black, white, and Latino, two-parent families in urban areas with a male head aged 18–58 and income below 150% of the poverty line	0.5 0.75 1.00 1.25	0.3 0.5 0.7
The Rural Income-Maintenance Experiment (RIME)	Iowa and North Carolina	1970–1972	809 (729)	Both two-parent families and female-headed households in rural areas with income below 150% of poverty line	0.5 0.75 1.00	0.3 0.5 0.7
The Seattle/Denver Income Maintenance Experiments (SIME/DIME)	Seattle and Denver	1970–1976, (some to 1980)	4800	Black, white, and Latino families with at least one dependent and incomes below $11,000 for single parents, $13,000 for two-parent families	0.75, 1.26, 1.48	0.5 0.7, 0.7–0.025y, 0.8–0.025y[c]
The Gary, Indiana Experiment (Gary)	Gary, Indiana	1971–1974	1799 (967)	Black households, primarily female-headed, head 18–58, income below 240% of poverty line	0.75 1.0	0.4 0.6
The Manitoba Basic Annual Income Experiment (Mincome)	Winnipeg and Dauphin, Manitoba	1975–1978	1300	Families with, head younger than 58 and income below $13,000 for a family of four	C$3800 C$4800 C$5800	0.35 0.5 0.75

[a]The "grant level" or "guarantee level" is the maximum NIT level for a person or family with no income other than the NIT. This is the equivalent of the UBI level. US grant levels were reported as a percentage of the poverty line. Canadian grant levels were reported in Canadian dollars (C$)

[b]The "marginal tax rate" or the "take-back rate" is the rate at which the NIT is reduced as income rises. This is equivalent to the rate at which income is taxed in an income-tax-financed equivalent UBI

[c]"y" stands for family income

Source: Reproduced from Widerquist (2005)

Table 6.2 Summary of findings of labor-reduction effect

Study	*Data source*	*Labor reduction*[a] *(in hours per year*[b] *and percentage)*			*Comments and caveats*
		Husbands	*Wives*	*SFH*	
Robins (1985)	4 US[c]	−89 −5%	−117 −21.1%	−123 −13.2%	Study of studies that does not assess the methodology of the studies but simply combines their estimates; finds large consistency throughout, and "in no case is there evidence of a massive withdrawal from the labor force"; no assessment of whether the work response is large or small or its effect on cost; estimates apply to a poverty-line guarantee rate with a marginal tax rate of 50%
Burtless (1986)	4 US	−119 −7%	−93 −17%	−79 −7%	Average of results of the four US experiments weighted by sample size, except for the SFH estimates, which are a weighted average of the SIME/DIME and Gary results only
Keeley (1981)	4 US	−7.9%			A simple average of the estimates of 16 studies of the four US experiments
Robins and West (1980)	SIME/ DIME	−128.9 −7%	−165.9 −25%	−147.1 −15%	Estimates "labor supply effects"; it goes without saying that this is different from "labor-market effects"
Robins and West (1980)	SIME/ DIME	−9%	−20%	−25%	Recipients take 2.4 years to fully adjust their behavior to the new program
Cain et al. (1974)	NJ	–	−50 −20%	–	Includes caveats about the limited duration of the test and the representativeness of the sample; notes that the evidence shows a smaller effect than nonexperimental studies
Watts et al. (1974)	NJ	−1.4% to −6.6%	–	–	Depending on size of G and t

Rees and Watts (1976)	NJ	−1.5 hpw[b] −0.5%	−0.61%	–	Found anomalous positive effect on hours and earnings of blacks
Ashenfelter (1978)	RIME	−8%	−27%	–	"There must be serious doubt about the implications of the experimental results for the adoption of any permanent negative income tax program"
Moffitt (1979)	Gary	−3% to −6%	0%	−26% to −30%	No caveat about missing demand, but careful not to imply the results mean more than they do
Hum and Simpson (1993)	Mincome	−17 −1%	−15 −3%	−133 −17%	Smaller response to the Canadian experiment was not surprising because of the make-up of the sample and the treatments offered

[a]The negative signs indicate that the change in labor effort is a reduction

[b]Hours per year except where indicated "hpw," hours per week

[c]US refers to four US studies of the era; it excludes Canada and the studies conducted in the twenty-first century

NJ New Jersey graduated work incentive experiment; *SIME/DIME* Seattle/Denver income maintenance experiment; *Gary* Gary income maintenance experiment; *RIME* rural income maintenance experiment; *Mincome* Manitoba income maintenance experiment; *SFH* single female "head of household"

Source: Reproduced from Widerquist (2005)

of the studies on the labor-effort response to the NIT experiments, showing the difference in hours (the "labor reduction") by the experimental group relative to the control group in foregone hours per year and in percentage terms. Results are reported for three categories of laborers, husbands, wives, and "single female heads" (SFH), which meant single mothers. The relative labor reduction varied substantially across the five experiments from 0.5% to 9.0% for husbands, which means that the experimental group worked less than the control group by about 0.5 hour to 4 hours per week, 20–130 hours per year, or 1–4 fulltime weeks per year. Three studies averaged the results from the four US experiments and found relative labor-reduction effects in the range of 5–7.9%.[3] One study using computer simulations estimated that the labor reduction in response to a national program would be only about one-third of the reduction in the Gary experiment (1.6% rather than 4.5%) because the sample was drawn from a relatively small portion of the population (people living near or below the poverty line).[4]

The response of wives and single mothers was somewhat larger in terms of hours and substantially larger in percentage terms because they tended to work fewer hours, to begin with. Wives reduced their labor effort by 0–27% and single mothers reduced their labor effort by 15–30%. These percentages correspond to reductions of about 0–166 hours per year. The labor-market response of wives had a much larger range than the other two groups, but this was usually attributed to the peculiarities of the labor markets in Gary and Winnipeg, where particularly small responses were found.

Studies that I reviewed did not place great stress on how reliable estimates were considered to be of the possible national response. Most of the data I have below represents point estimates of the difference between the control and experimental groups rather than confidence intervals or estimates of the national response.

[3] G. Burtless, "The Work Response to a Guaranteed Income. A Survey of Experimental Evidence," in *Lessons from the Income Maintenance Experiments*, ed. A. H. Munnell (Boston: Federal Reserve Bank of Boston, 1986). M.C. Keeley, *Labor Supply and Public Policy: A Critical Review* (New York: Academic Press, 1981). P.K. Robins, "A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments," *Journal of Human Resources* 20, no. 4 (1985).

[4] R.A. Moffitt, "The Labor Supply Response in the Gary Experiment," ibid. 14 (1979).

All or most of the figures reported above are raw comparisons between the control and experimental groups: they are not predictions of how labor-market participation is likely to change in response to a national NIT or UBI. Consider four of the many reasons why.

First, participants tended to be drawn from a small segment of the population: people with incomes near the poverty line. This part of the income distribution is about where one would expect the largest negative labor-effort effect because the potential grant is high relative to their earned income. Thus, the response of the group studied is likely to be much larger than the response of the entire labor force to a national program. As mentioned above, one study using computer simulations estimated that the labor reduction in the Gary experiment (4.5%) would translate into a 1.6% labor-effort reduction in a national program.[5] I wonder whether numbers like 1.6%—more easily perceived as negligible—would have had a different effect on the discussion of the results at the time.

Second, the figures do not include any demand response, which economic theory predicts would lead to higher wages and a partial reversal of the labor reduction (see this chapter). One study using simulation techniques to estimate the demand response found it to be small.[6] Another found that, "[r]eduction in labor supply produced by these programs does tend to raise low-skill wages, and this improves transfer efficiency."[7] That is, it increases the benefit to recipients from each dollar of public spending.

Third, the figures were reported in average hours per week and very often misinterpreted to imply that 5–7.9% of primary breadwinners dropped out of the labor force. In fact, few, if any workers simply dropped out of the labor force for the duration of the study, as knee-jerk reactions to guaranteed income proposals often assume.[8] Primary breadwinners in both the experiment and control groups left their jobs (whether voluntarily or by getting fired or laid off) at about the same rate. The observed

[5] Ibid.

[6] D.H. Greenberg, "Some Labor Market Effects of Labor Supply Responses to Transfer Programs," *Social-Economic Planning Sciences* 17, no. 4 (1983).

[7] J.H. Bishop, "The General Equilibrium Impact of Alternative Antipoverty Strategies," *Industrial and Labor Relations Review* 32, no. 2 (1979).

[8] Robert Levine et al., "A Retrospective on the Negative Income Tax Experiments: Looking Back at the Most Innovative Field Studies in Social Policy," in *The Ethics and Economics of the Basic Income Guarantee*, ed. Karl Widerquist, Michael A. Lewis, and Steven Pressman (Aldershot: Ashgate, 2005).

labor-effort reduction was mainly caused by workers in the experimental group taking longer to find their next job if and when they became nonemployed.

Fourth, the experimental group's labor "reduction" was only a relative reduction in comparison to the control group. Although this language is standard for experimental studies, it is often wrongly taken to imply that receiving the NIT was the major determinant of labor hours. In fact, in some studies, labor hours increased for both groups, and in all studies, the labor hours of both groups tended to rise and fall together along with the macroeconomic health of the economy—implying that when good jobs were plentiful, both groups took them, but when they were less plentiful, the control group searched harder or accepted less attractive jobs.[9]

A bigger problem than misinterpretations of the size of the labor-effort reduction was that most laypeople writing about the NIT experiments assumed any labor reduction, no matter how small, was an extremely negative side effect. But it is not obviously desirable to put unemployed workers in the position where they are desperate to start their next job as soon as possible. It's obviously bad for workers and families to be in that position. It's not only difficult to go through, but also it reduces their ability to command desirable wages and working conditions. Increased periods of nonemployment might have a social benefit if they lead to better matches between workers and firms.

Another problem with the focus on labor effort was that it distracted attention from the question of how well the NIT achieved its main goals of reducing poverty and increasing the well-being of low-income people. Assessing these issues requires looking at nonlabor-market effects.

2 Nonlabor-Market Effects of the NIT Experiments

The experimental results for various quality-of-life indicators were substantial and encouraging. Some studies found significant positive influences in elementary school attendance rates, teacher ratings, and test scores. Some studies found that children in the experimental group stayed in school significantly longer than children in the control group. Some found an increase in adults going on to continuing education. Some of the experiments found

[9] Widerquist, "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments?"

desirable effects on many important quality-of-life indicators, including reduced incidents of low-birth-weight babies, decreased household indebtedness, increased food consumption, improvements in medical treatment, and increased nutritional content of the diet, especially among children. Some even found reduced domestic abuse and reduced psychiatric emergencies.[10]

Much of the attention to nonlabor-market effects focused not on the presumed goals of the policy but on another side effect: a controversial finding that the experimental group in SIME/DIME had a higher divorce rate than the control group. Researchers argued forcefully on both sides into the early 1990s, with no conclusive resolution in the literature. The finding was not replicated by the Manitoba experiment, which found a lower divorce rate in the experimental group. The higher divorce rate in some studies examining SIME/DIME was widely presented as a negative effect, even though the only explanation researchers had for it was that the NIT must have relieved women from financial dependence on husbands.[11] It is at the very least questionable to label one spouse staying with another solely because of financial dependence as a "good" thing.

3 An Overall Assessment?

Most of the researchers involved considered the results extremely promising overall. Comparisons of the control and experimental groups indicated that the NIT was capable of significantly reducing the material effects of poverty, and the relative reductions in labor effort were probably within the affordable range and almost certainly within the sustainable range.

But experiments of this type were not capable of producing a bottom line. Nonspecialists examining the results were left asking: what *was* the cost exactly? How much were the material effects of poverty reduced? What is the verdict from an overall comparison of costs and benefits?

As this book argues throughout, experiments cannot answer these questions, although they can contribute towards attempts to address these questions. Simply reporting experimental comparison without explaining what they contribute to these larger issues leads to misunderstanding—as the following section illustrates.

[10] Levine et al.

[11] Ibid.; Widerquist, "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments?"

4 Public Reaction to the Release of NIT Experimental Findings in the 1970s

As promising as the results were to the researchers involved, the NIT experiments were seriously misunderstood in the public discussion at the time. The discussion in Congress and in the popular media displayed little understanding of the complexity of experimental results or difficulties of extrapolating them into answers to any bottom-line question. The results were spun or misunderstood and used in simplistic arguments to reject any form of guaranteed income offhand.

The experiments were of most interest to Congress during the period from 1970 to 1972, when President Nixon's Family Assistance Plan (FAP), which had elements of an NIT, was under debate in Congress. None of the experiments were ready to release final reports at the time. Congress insisted researchers produce some kind of preliminary report, which was criticized by members of Congress for being "premature," just as researchers had warned.[12]

Results of the fourth and largest experiment, SIME/DIME, were released while Congress was debating a policy proposed by President Carter, which had already moved quite away from the NIT model. Dozens of technical reports with large amounts of data were simplified down to two statements: NIT decreased labor effort and supposedly increased divorce. The smallness of the labor disincentive effect hardly drew any attention. Although from the start, researchers expected some labor-reduction effect and were pleased to find it was small enough to make the program affordable, many members of Congress and popular media commentators acted as if the mere existence of a labor-reduction effect was enough to disqualify the program.

The public discussion displayed little, if any, understanding that the 5–7.9% difference between the control and experimental groups is not a prediction of the national response. In an earlier work, I reviewed nonacademic articles on the experiments and found that they had little or no understanding that the labor-effort response would be much smaller as a percentage of the entire population, that it could potentially be counteracted by the availability of good jobs, or that it could be the first step necessary for workers to command higher wages and better working conditions, which could partly counteract the labor-reduction effect.[13]

[12] "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments?"

[13] Ibid.

The United Press International simply got the facts wrong, saying the SIME/DIME study showed, "adults might abandon efforts to find work," as did *the Rocky Mountain News*, which claimed that the NIT "saps the recipients' desire to work." The *Seattle Times* presented a relatively well-rounded understanding of the results, but despite this, it simply concluded that the existence of any decline in labor effort—regardless of size—was enough to "cast doubt" on the plan.

Others went even farther, saying that the existence of a work-disincentive effect was enough to declare the experiments a failure. Headlines such as "Income Plan Linked to Less Work" and "Guaranteed Income Against Work Ethic" appeared in newspapers following the hearings. Only a few exceptions such as Carl Rowan for the *Washington Star* considered that it might be acceptable for people working in bad jobs to work less, but he could not figure out why the government would spend so much money to find out whether people work less when you pay them to stay home.[14]

Senator Daniel Patrick Moynihan, who was one of the few social scientists in the Senate, also failed to understand the experimental findings. He wrote, "But were we wrong about a guaranteed income! Seemingly it is calamitous. It increases family dissolution …, decreases work, etc. Such is now the state of the science, and it seems to me we are honor bound to abide by it for the moment." Senator Bill Armstrong, mentioning *only the existence* of a labor-disincentive effect, declared the NIT "An acknowledged failure," writing, "Let's admit it, learn from it, and move on."[15]

Robert Spiegelman, one of the directors of SIME/DIME, defended the experiments in an op-ed piece, in which he argued that the experiments provided much-needed cost estimates that demonstrated the feasibility of the NIT. He said that the decline in labor effort was not dramatic and could not understand why so many commentators drew such different conclusions than the experimenters. Gary Burtless remarked, "Policymakers and policy analysts … seem far more impressed by our certainty that the efficiency price of redistribution is positive than they are by the equally persuasive evidence that the price is small."[16]

The experiments produced a great deal of useful evidence, but failed to communicate those results either to Congress or to the public. The literature review reveals neither supporters nor opponents who appeared to

[14] Ibid.
[15] Ibid.
[16] Burtless.

have a better understanding of the likely effects of the NIT or any income guarantee in the discussions following the release of the results of the experiments in the 1970s.[17]

The late-1970s reaction to experimental results reflected the times, as politicians like Ronald Reagan were attracting support to the idea of cutting the welfare system rather than expanding and improving it, often by vilifying almost anyone who was eligible for redistributive programs. Many of the commentaries on the SIME/DIME results reflected such a perspective, but I would caution against reading too much into the timing. The complaint about giving too much support to the sturdy beggar has been a perennial demagogic talking point in English-speaking countries since the Elizabethan era. And it remains a tempting talking point for opponents of redistribution almost anywhere. Thus, while keeping the context in mind, I ask readers to consider the potential that this experience might contain more widely applicable lessons.

Whatever the causes of it, an environment with a low understanding of complexity is highly vulnerable to spin with simplistic or even vacuous interpretation. All sides spin, but in the late-1970s NIT debate, only one side showed up. The guaranteed income movement that had been so active in the United States at the beginning of the decade had declined to the point that it was able to provide little or no counterspin to the enormously negative discussion of the experimental results in the popular media.

Whether the low-information content of the discussion in the media resulted more from spin, sensationalism, or honest misunderstanding is hard to determine. But whatever the reasons, the low-information discussion of the experimental results put the NIT (and, in hindsight, UBI by proxy) in an extremely unfavorable light, when the scientific results were mixed to favorable.

Researchers working on the experiments were blind-sided by the level of spin. They had not been asked to make special efforts to explain their results to laypeople in a way that would head off possible spin. If they had been asked, they would have had no particular expertise in doing so. And even if they or some science communication specialist had tried, it would have been extremely difficult, if not impossible, to communicate the complexities to most nonspecialists in the time a reasonable person typically devotes to the issue.

[17] Widerquist, "A Failure to Communicate: What (If Anything) Can We Learn from the Negative Income Tax Experiments?"

Thus, it would be wrong simply to blame researchers for failing to communicate their results clearly. The problem came from the inherent difficulty of communicating complex and tentative scientific findings to a lay audience looking for definitive answers on questions that are only partly related to those findings. Everyone involved has a responsibility not to be blind-sided by spin and misunderstanding next time. The political context will be different, but the warning needs to be considered. The rest of this book is an effort to help reduce similar misunderstandings in future experiments.