The Income Maintenance Experiments and the Issues of Marital Stability and Family Composition

*Glen G. Cain**

Between 1968 and 1978 four negative income tax experiments were conducted; they were designed to measure labor supply and earnings. The experiments were not designed to measure the effects of government programs on such demographic behavior as marital dissolution, fertility, family composition, or the decision to marry or remarry. Nevertheless, the data from the experiments have been used to analyze all these family issues, and they are the subject of this paper.

The essential reform examined in the negative income tax experiments was the extension of a guaranteed minimum income to poor families with an able-bodied, non-aged husband or father as the potential provider. The income plans tested in the experiments were expected to lead to reductions in the labor supply and earnings of the participating married-couple families. By a twist of fate, however, the most influential research finding of the experiments turned out to be not about labor supply but about marital stability, a family issue. The findings on labor supply showed reductions neither large enough nor small enough to permit a definitive verdict about the negative income tax. In contrast, the findings about marital stability appeared decisive.

The most important research on marital stability was conducted by Groeneveld, Hannan, and Tuma, based on the Seattle-Denver income maintenance experiment.[1] They concluded that the negative income tax increased marital dissolutions, even though it had been designed to

*Professor of Economics, University of Wisconsin. For help in getting the data the author is grateful to Katherine Dickinson, Mario Lopez-Gomez, Philip Robins, Daniel Weinberg, and Richard West. James Albrecht, Edward M. Gramlich, and Nancy B. Tuma provided helpful comments. Douglas Wissoker provided excellent counsel as well as programming and statistical help.

cover and assist families headed by married couples as well as families headed by women. Indeed, their finding applied to a negative income tax plan of the same level of generosity as the prevailing aid to families with dependent children (AFDC) plan. This conclusion was unambiguously unfavorable to advocates of a negative income tax that would cover married couples, for two important reasons. First, increased marital breakups among the poor would increase the numbers on welfare and the amount of transfer payments, principally because the separated wife and children would receive higher transfer payments. Second, marital dissolutions and the usual accompanying absence of fathers from households with children are generally considered unfavorable outcomes regardless of whether or not the welfare rolls increase.

Besides appearing decisive, the experimental findings about marital stability were dramatic: the reported increase in marital splits was large; it was counter to the outcome hoped for and expected by advocates of the negative income tax reform; and it was counter to the predictions of social scientists, and in particular economists. The dramatic findings received considerable attention by the press and intense scrutiny by scholars, who were skeptical but eventually accepted the findings.

This review of the negative income tax experiments offers two main messages. The first is that the evidence about the issue of marital stability is not decisive, or even persuasive. A second message is that family issues such as marital stability are not well suited to experimental research. The costs of a properly designed experiment seem too high.

Social Experimentation and Family Issues: The Case of Marital Stability

The belief that marital stability among low-income families has been adversely affected by our current welfare system seems firmly entrenched, even though empirical evidence in support of this belief has been difficult to marshal.[2] The general upward trend in divorce, separation, and female headship of families throughout this century[3] applies to all income strata; welfare programs are likely to be a factor in marital stability only among the lower half of the family income distribution and only during the last 25 years or so. In recent decades the generosity of welfare programs systematically, although not steadily, increased in ways that tended to lower the financial cost of marital dissolution to a married couple with children. For a mother, the income from welfare, which may include such in-kind payments as food stamps, Medicaid, and housing subsidies, as well as the cash payments from AFDC, provides an alternative to her husband's income. Welfare is likely in poor families to exceed the income the wife receives from the husband.

Moreover, mothers in poor families are less likely to be capable of earning enough to be self-supporting at a level of income that exceeds welfare. Finally, the availability of welfare essentially required the husband-father to leave the marriage. For a departing father, the welfare programs provided a de facto if not a legal alternative to alimony and child support.

The increase in marital dissolutions is, however, only one of the trends contributing to the increase in the proportion of families headed by a woman. Increases in the number of unwed mothers, in the length of time that mothers remain without husbands (regardless of whether they are divorced or never married), and in the proportion of mothers who establish separate households, are all sources of increased female headship and of welfare recipiency.

In summary, the spread and increased generosity of welfare programs have reduced the price (or cost) of marital dissolution. The AFDC program has been decidedly nonneutral regarding marital dissolutions in its dispensing of transfer payments and other benefits.[4] One advantage widely claimed for a negative income tax was that it would move the income maintenance system toward neutrality in marital decisions.

An Economic Framework for Analyzing Marital Stability and the Negative Income Tax Experiments

To determine how marital stability among the population of already married couples will be affected by a negative income tax, two regimes must be compared:

- The current system of welfare programs, referred to as AFDC, which provides a net subsidy to a dissolved marriage, given the presence of dependent children and assuming that the divorced woman meets the income criteria for eligibility.
- A negative income tax regime, in which the current system is amended to add welfare assistance for married couples who meet the income criteria for eligibility.

We may assume for the negative income tax, as we did for AFDC, that the only way it affects marital stability is by the income changes that it brings about. Income changes in turn induce an ''income effect'' in a regime where the income receipts are neutral with respect to marital status, and a ''price effect'' that refers to the nonneutrality of the income change with respect to marital status. Even with the simplifying assumption that income changes are all that matter, three differences between the negative income tax and AFDC are critical.

First, a negative income tax provides transfer payments for married couples whose incomes are low. Economists and sociologists appear to agree that modest increases in the incomes of married couples ought to have a positive effect on marital stability. They find both an empirical negative relation between family income and the probability of a marital dissolution and a theoretical argument that poverty puts a strain on a marriage, creating tensions and dissatisfactions that contribute to a subsequent dissolution.[5]

Against this apparent consensus are one empirical finding and three theoretical arguments. First, the time trend is not supportive of the idea that rises in income are associated with increased marital stability. Second, consider the plausible hypothesis that in many instances income has a positive effect on divorce. After all, more income can make desired divorces and separate living arrangements affordable. Third, income is partly determined by personal traits that are themselves related to marital stability, so the empirical positive relation between income and marital stability does not imply a causal relation. Fourth, income from welfare may not be "ordinary income," because it may carry a stigma that is destabilizing.[6] Although no direct evidence indicates such a stigma effect is operative and destabilizing, that issue will be discussed later. To summarize, a negative income tax carries a direct income effect to the recipient married couple that is commonly viewed as promoting marital stability. We may consider marital stability to be, on average, a "normal good," but the evidence and theories appear only weakly supportive.

A second major difference is that AFDC requires the presence of dependent children for the receipt of transfer payments, whereas some proposed income maintenance plans, including those adopted in the Seattle-Denver experiment, do not. Because the presence of dependent children in a marriage has a negative effect on marital dissolution, this is an important difference in the two regimes. Restricting our attention to families with children will solve two problems: it will provide the proper comparison with existing AFDC programs, and also provide information about the only type of negative income tax legislation that is likely to be considered by Congress.

A third potential difference between the existing AFDC regime and the proposed negative income tax plan involves the level of payments received by a recipient family. To simplify matters, let us assume that the payment depends on three parameters: (a) the income guarantee for a family of a given size with no other income, (b) the benefit-reduction rate, and (c) the differential in income guarantees for families of different size and demographic composition. Regarding (c), let us assume that the negative income tax guarantee amounts are structured so that approximate neutrality in the economic well-being of different-sized families is achieved, if the guarantee were the only income received by the family. Given such a structure, the higher the income guarantee and the lower the benefit-reduction rate, the more generous is the plan, because more families are eligible and because recipient families will receive larger transfer payments.[7] In the following three comparisons of predicted effects on marital stability, let us assume that the mother maintains custody of the children.

1. *If the negative income tax and AFDC plans are equally generous* to a mother who is without her husband, then a net decrease in marital breakups is predicted. The main reason is that the price subsidy to the unmarried state is reduced. The mother would gain the same amount as before (under AFDC) if she is divorced or separated, and she would receive more than before if she remained married. The same statement applies to the husband-father, assuming that the negative income tax plan, like AFDC, provides no income support to the separated husband. If the husband were to become the provider of a new family and had a sufficiently low income, then he could gain under a negative income tax regime relative to AFDC. Let us assume that this possibility is sufficiently remote that the potential gain is negligible.

The negative income tax experiments, including Seattle-Denver, did differ from AFDC by providing an income guarantee and potential transfer payments to the departing husband, even if he remained single. This could be a major difference in the incentives to marital dissolution. However, the break-even levels of income for single persons were so low—around $2000 a year in 1970 dollars—that I will assume, unless otherwise noted, that the actual benefits to the husband from this provision of the negative income tax are negligible and can be ignored.

Returning to the payments received by the divorced mother, the term "independence effect" was used by Groeneveld, Hannan, and Tuma in their studies of marital dissolution to refer to the woman's opportunity to use these payments to support herself and her children. I prefer to speak of the price subsidy to being divorced and to focus on the neutrality or nonneutrality of the subsidy: the payments received when the woman is separated from her husband relative to the payments received if she stays married. (As noted above, the price subsidy also affects the husband by reducing his subjective obligation to pay child support or alimony.)

A second reason for fewer marital dissolutions under the negative income tax regimen relative to AFDC is that the married couple may receive income transfers. Here the focus is on the receipt of income, per se. Although the income effect does not have an unambiguous sign, prevailing opinion suggests that it should be weakly pro-stability.

In summary, if the negative income tax is as generous as AFDC, its price and income effects should lead to a reduction in marital splits. However, only a negligible decrease would be predicted by those who: (i) minimize the income effect; (ii) emphasize the fact that there has been no change in the independence effect; and (iii) believe that there are actually few cases under our current welfare system in which a father leaves his wife and children to permit them to qualify for transfer payments. A negative income tax plan of equal generosity obviates taking this drastic step, but apparently there has never been any concrete evidence that this behavior occurs.[8] Of course, the lack of empirical evidence does not mean that the events have not occurred.

2. *If the negative income tax is less generous than AFDC, and the two plans coexist,* then again a net decrease in marital breakups should occur, but the decrease should be even smaller than in the case of an equally generous negative income tax. In brief, the negative income tax again reduces the price subsidy to being divorced relative to being married, although the ratio of payments received with and without a divorce is smaller. Also, the presumed pro-stability income effect is smaller than in the previous case, because the transfer payments to the married couple are lower. Again, the potential transfer payments to the separated husband under a negative income tax plan probably play no role in his decision.

3. *If the negative income tax is more generous than AFDC,* then the two theoretical effects we have considered have opposite signs with respect to marital stability. The higher payments of the negative income tax would dominate AFDC, and the latter would disappear for lack of customers. Under the more generous negative income tax the payment to the divorced woman is increased, but so is the payment to the woman (and her husband) if she remains married. It is likely that the higher level of payments to the woman if she divorces dominates the comparisons in her decision to remain married or become divorced. In this sense, the "independence effect" of the income maintenance system has been increased, implying that the effective price subsidy to being divorced is increased, leading in turn to an increase in marital dissolutions. A more generous plan increases the payments to married couples, and the pure income effect, which is the second theoretical effect, may be assumed to promote marital stability. Thus, a more generous negative income tax induces a net price change that promotes marital breakups and an income change that promotes marital stability. For reasons discussed above, the price effect appears, a priori, to be the stronger.

A Broader Research Agenda for Marital Stability

The economic framework presented in the previous section was overly simplified in several respects. A more realistic setting and objectives will be useful for the reanalysis of the Seattle-Denver study and will reinforce the message of the difficulty, perhaps intractability, of social experiments on family issues.

A fundamental question that has been only partially addressed by past research is the precise purpose of analyzing marital dissolutions and other demographic outcomes of the experiments. Given our attention to welfare reform, one important purpose of an experiment is to measure the fiscal costs of the reform. The costs will tend to rise or fall depending on whether the reform increases or reduces marital dissolutions. Unlike reductions in labor supply and earnings, however, the change in dissolutions may have no effect on national income. For example, a married couple that is not receiving welfare benefits may split, after which the mother and children begin to receive welfare payments. The divorced husband may continue working as much as he did before, and the divorced wife may work in the market no less, and perhaps even more, than she did before. The change in marital stability has no clear effect on national income, even though its relation to program costs is useful to measure.

We would also like to know how the well-being of low-income married couples is affected by a change in marital dissolutions. With the introduction of a negative income tax, married couples have an expanded set of options regarding their living arrangements, and if they choose to change them, we may presume that they are better off. This is, of course, an application of the economist's conventional assumption of consumer sovereignty and rational behavior. The important point here is that those who use the experimental results to design welfare programs are not going to be able to answer these questions about well-being.

Finally, what is the effect on the well-being of the children in the low-income families that experience a change in marital status as a result of the negative income tax? This difficult question may require many years to elapse before it can be answered. However, the social concern about deleterious effects of marital breakups on the well-being of children is sufficiently widespread that we may agree on the importance of using the experiment to measure marital dissolutions in families with children.

Given the two purposes of measuring program costs and marital dissolutions in families with children, we are in a firmer position to assess the strengths and weaknesses of different experimental designs. A well-known weakness of the experiments, in terms of measuring the demographic consequences of a nationally legislated negative income tax, is their short duration—three to five years. Short-duration experiments do not simulate the incentives of a permanent negative income tax pertaining to demographic behavior. The apparent bias in a short-duration experiment would be to understate the program's effect on the lifetime or steady-state incidence of births, marriages, and divorces, since the present value of the subsidies from a short-duration program are lower. However, the timing of these outcomes may be so affected that a short-duration experiment could overstate, rather than understate, the impact of a permanent program.

To illustrate these biases, let us use the example of births. With a permanent negative income tax, a married couple might decide to have three instead of two children in response to the incentive of the extra transfer payments they will receive during the 16 to 20 or so years that the child is their dependent. Another incentive is that any reduction in market earnings by the mother during the first 10 or so years of the additional child's life may also be partially offset by an additional increase in transfer payments. Clearly, the total value of these extra transfers under a permanent negative income tax is considerably larger than the payments received under a short-duration experiment. Thus, the lifetime incidence of the births of additional children may be substantially understated in the experiment. But to illustrate an opposite bias whereby fertility is overstated, consider all couples who plan to have an extra child and who would do so whether or not a negative income tax program exists. They might respond to the subsidy for births in the short-duration experiment by bearing that child "now" rather than "later." The short-duration experiment will, in these cases, overstate the fertility effect of a negative income tax.

It is partly a matter of judgment and partly a matter of ingenuity in analyzing the experimental data to determine which behavioral outcomes are affected by these duration biases and to measure the bias. Regarding marital dissolutions, we might suspect that teenage cohabiting couples who did not have a legal marriage and had no children would be more likely to alter the timing (as well as the incidence) of dissolutions in response to financial incentives than would legally married couples in their thirties with children present.

Another shortcoming of the negative income tax experiments regarding the measurement of marital stability is the reliance on already married couples. There are two problems here. One is that the relevant population, given our interest in children in families without a father present, includes women who have never married and some who are divorced or separated and are not now married. A second and related problem is that the measure of marital breakups with a sample of already married couples may be biased even as a measure of dissolutions among married couples. Both problems arise because a permanent negative income tax may be expected to affect the decision to marry and to remarry, as well as the decision to dissolve an existing marriage.

In considering the problem of the relevant populations in connection with the well-being of children, an example may illustrate some important issues. Let us define family stability in terms of the proportion of time that children spend growing up with their mother and father present. So defined, an increase in marital dissolutions could be consistent with an *increase* in family stability according to the following scenario. Assume that a negative income tax that covers married couples increases the proportion of young unwed mothers who marry the fathers of their children. Assume further that these unwed mothers would not have married in the absence of the program. These two assumed outcomes are realistic because the current AFDC program, which provides transfer payments to the mother only if she does not marry, is assumed to be superseded by a negative income tax that provides transfer payments to the mother if she does marry (and is income eligible). Finally, assume that the proportion of these marriages that ends in divorce is higher than the proportion of divorces in the rest of the married-couple population. (This is also a realistic assumption.) The end result is that the overall proportion of marital dissolutions is increased. Nevertheless, the assumed marriages of unwed mothers who would not have married otherwise increases family stability as defined by the presence of a father and mother during the time of the upbringing of their children. The same apparent paradox—an increase in family stability accompanied by an increase in marital dissolutions—will result from a similar scenario applied to remarriages of divorced mothers with children.

The above example illustrates the point that female headship is probably more influenced by the current welfare system than is marital stability. AFDC may not create female-headed families by providing a monetary incentive for a father to leave his family nearly as often as it does by discouraging the marriage of young unwed mothers and of currently divorced and separated women who are receiving AFDC and other welfare assistance.[9] Thus, the proportion of dissolved marriages among women who marry could be *reduced* by the current system because the system discourages certain marriages from occurring in the first place.

Now consider a second bias in measuring marital dissolutions that arises when examining only the existing stock of married couples.[10] Two examples will illustrate the problem. First, assume that unmarried women tend to have preferences in favor of singleness and against marriage. If a negative income tax offsets the current incentive to singleness and encourages more marriages among these women, we might expect that the proportion of divorces in these marriages would be above average. As a consequence, the long-run impact of a negative income tax on divorces is understated by observing only the already married population.

An opposite bias is also possible. Assume that the population of unmarried women is composed of two groups: one that is committed to singleness and will not marry and another that is planning on marrying but is taking more time to search more carefully. If a negative income tax encourages more, or even less delayed, marriages among this second group, we might expect fewer divorces among these marriages than the average. If so, then the full impact on divorces is overstated by using only the already married population in an experiment.

The data from the negative income tax experiments do not realistically allow analyses to estimate the effect on (a) first marriages of never-married women; (b) marriages of unwed mothers; (c) remarriages of divorced women who are not currently married; and (d) subsequent marital decisions. One problem is the small sample sizes of some of the populations. A second is the short duration of the experiment relative to the time horizons for these decisions. Again, the downward bias of a short-duration plan, stemming from the lower present value of the plan's transfer payments, is competing with the upward bias from intertemporal substitution.

The case of new marriages by women and men without children is worth special attention to reveal some of the complexities in using the experiments to analyze marital behavior. The Seattle-Denver experiment created unusual bonuses for new marriages. Imagine an 18- to 20-year old unmarried son or daughter in an experimental family who is considering marrying or cohabiting. The first bonus to marriage is eligibility for cash transfers for the couple, including an additional transfer allowance that the experimental plan will assign to the new partner. These cash transfers are not available to *either* partner under the existing welfare system. It is also unlikely that a couple without children would be eligible for cash transfers in a nationally legislated plan. The son or daughter in a family eligible to receive experimental negative income tax payments received a dowry to a new union unavailable to other people in that "marriage market," and operative only for the duration of the experiment. A second bonus to the new marriage or cohabitation was that eligibility for experimental payments was extended to the new partners even if they later dissolved the union and formed a second union with a different partner.[11]

My point is not to dwell on the difficulties in using existing experimental data to test for effects of a negative income tax on new marriages and subsequent dissolutions. Rather, it is that the already married couples are the only feasible group to use to examine marital stability and that there are inherent limitations in relying on this group.

Summary Points on the Experiments and the Issue of Marital Stability

Several lessons can be drawn about the use of experiments to study marital stability. First, the population of interest should include the major groups that will be affected, and in particular, young persons who are not married. Second, long-duration experiments seem necessary both because the decisions about marital status and other family issues involve long-duration plans and consequences and because the biases from using short-duration experiments are not clear in direction. Both lessons would apply to the issues of fertility and marriage as well as to marital dissolution. A third lesson, which is derived from our interest in welfare reform, is that families with children should receive priority in the design of the experiment, which should include a scheme in which families without children are not eligible for payments. Finally, a simplified economic framework for analyzing how a negative income tax influences marital dissolution suggests two predictions:

- A negative income tax plan that is as generous as an AFDC plan or less generous should promote marital stability.
- The predicted effect on marital stability of a plan that is more generous than the existing AFDC plan is ambiguous. If the negative income tax led to fewer marital splits, we could infer that the gain to a married couple that stays together dominates the extra gain that the mother would receive from the new plan if the marriage dissolves. If the negative income tax led to more marital splits, then that latter gain would appear to dominate the decision.

The Experiments, with Special Reference to the Seattle-Denver Experiment

The findings on marital stability that received by far the most attention are those from the Seattle-Denver experiment, which was the largest of the four negative income tax experiments. Seattle-Denver had the advantage of a five-year duration for a subsample of about 25 percent of the married couples; most of the other 75 percent were in the experiment for three years, as were the subjects in the other three experiments. In fact, a small subsample of 169 families was assigned to a third, 20-year duration, category. This assignment occurred after the experiment had started and after these families had been originally assigned to three-year treatment and control groups. This 20-year group was not analyzed separately in the research on marital issues, but will be referred to later.

The Seattle-Denver experiment was the last to be completed, and its research team of Groeneveld, Hannan, and Tuma reviewed and reanalyzed the findings on marital stability from the other three experiments. They concluded that these, individually or collectively, did not show any clear impact.[12] The sample size in the rural experiment was too small for conclusive evidence, particularly in view of the low proportion of families experiencing a marital dissolution—around 2 percent for both treatment and control groups. The Gary experiment showed no effect on marital splits, but the Groeneveld, Hannan, and Tuma team pointed to administrative and data flaws in this experiment that led them to discard these results. The first experiment, carried out in New Jersey and Pennsylvania, appeared to have inconclusive results regarding marital splits when originally analyzed, but the researchers reanalyzed the data and found evidence for a pro-split outcome that supported the conclusions of their Seattle-Denver research. The New Jersey study had a much smaller sample and more attrition among married-couple families than did Seattle-Denver, however.

In the Seattle-Denver experiment, marriage was defined by cohabitation, not by a marriage certificate or other legal sanction. A marital dissolution was defined by a separation, not necessarily a divorce, of at least 30 days and a statement from one of the partners that a separation had occurred.[13] Because this reanalysis is restricted to couples with children, the problem of dealing with unmarried cohabiting couples is presumed to be negligible. The Seattle-Denver definition of a separation seems to set rather loose criteria, however, and the effects of this on reports of a dissolution by experimental couples relative to control couples will be discussed below.

Several features of the Seattle-Denver experiment appear either to be obstacles to analyzing marital dissolutions or to imply reservations about the research findings.

1. The addition of a training and counseling program. A large proportion of the experimental group was given the option of a training and/or counseling program (hereinafter referred to as the training program) intended to improve the earning capacities of the adult family members. In fact, the Seattle-Denver experiment consisted of four major groups: families that were assigned only to a training treatment, families assigned only to a negative income tax plan, families assigned to both treatments, and control families. The training program complicates measuring the effect of a negative income tax on marital stability in three ways:

- No theoretical basis exists for predicting the sign of the effect of training on marital stability. The effect may differ depending on whether the husband or the wife receives the training.
- The sample size for the ''pure'' negative income tax treatment is sharply diminished.
- Training programs are difficult to administer in a way that will replicate how a nationally legislated program would be carried out and will be as unobtrusive to the experimental subjects as a nationally legislated plan. In contrast, a negative income tax plan has relatively rigid parameters that permit the experimental plan

to match closely the design of a nationally legislated plan, and its administrators play a relatively passive role.[14]

2. Small sample sizes for subgroups of interest. The Seattle-Denver experiment was directed toward three major ethnic groups: white, black, and Hispanic, the last primarily of Mexican heritage and living in Denver. There were eleven different negative income tax plans, two (or three) durations, and three training treatments. Problems of inadequate sample sizes arise when these features are extensively cross-classified.

3. The short duration of the experiment. The short duration of the experiments relative to the time horizons of such demographic behavior as marriage and divorce led Groeneveld, Hannan, and Tuma to emphasize the five-year plan, on grounds that it ''is more like a permanent program than the 3-year treatment.'' In reporting that the three-year treatment was about 75 percent as large as the five-year treatment, they commented that ''if the longer treatment more closely approximated the effects of a permanent program, a permanent program would have even larger effects than the 5-year program.''[15] These arguments imply that the intertemporal substitution bias is dominated by duration bias; that is, the five-year plan's lesser duration bias gives it an overall larger effect. However, the reanalysis of the Seattle-Denver data presented below does not support this finding.

4. Attrition. All research with longitudinal surveys has to deal with attrition bias. The attrition proportion was about 20 percent in the negative income tax experiments, including Seattle-Denver,[16] and the resulting biases may be serious. For example, the attrition bias of most concern in analyzing labor supply is that families whose earnings *declined* had an incentive *not to drop out* if they were in the experimental group, but had no such incentive if they were in the control group. Thus, the experimentals who did not drop out should overrepresent experimentals whose earnings declined, especially those who lost their jobs and had zero earnings. The controls had no such incentive not to drop out, and controls did drop out more than experimentals.

The attrition bias might be even more serious in analyzing marital stability than it is for labor supply analysis. A decline in earnings associated with decreased labor supply will consist of a continuum of small to large declines, whereas the decline in income for the mother after the departure of her husband is often a very large loss. A woman in the negative income tax experimental group whose marriage breaks up has the option of receiving the higher of the experimental payment or the AFDC payment. Women in the control group whose marriages dissolve get nothing from the negative income tax before or after the marital dissolution, so they have no economic incentive to continue responding to the interviews three times a year. Attrition rates were higher for control wives than experimental wives in the Seattle-Denver experiment, and the ratio of attrition rates, control-to-experimental, was higher for wives than for husbands. Experimental husbands were, of course, less likely to receive transfer payments after a marital dissolution than were their wives. The economic incentive was also evident within the experimental group: wives assigned to the most generous plans had lower rates of attrition than wives in the least generous plans.[17]

In summary, the sample of experimental families who remained in the study should overrepresent marital dissolutions relative to the full original sample of experimentals. The economic incentive for this attrition bias is not present in the control group, and other things equal, we might view the proportion of marital dissolutions in the remaining control group as an unbiased estimate of the proportion for the full sample of controls. However, if we believe that personal problems and traits that are associated with marital dissolution are also associated with attrition, then, with no economic incentive in operation, those controls who dropped out might well have a higher incidence of marital splits.[18]

5. Post-experimental design changes and the 20-year treatment. After the experiment began, the initial design was changed in several ways. First, shortly after the experiment began about 40 percent of the control families were assigned to a five-year control status.[19] The five-year controls were thus exempt from the substantial attrition that occurs early, and for this reason their attrition was much lower than that of the three-year controls. To the extent that the frequency rates for marital splits are affected by attrition bias, the rates for the five-year controls will differ from those for the three-year controls.

Another change in design occurred after the experiment had been in operation for about two and one-half years, when 169 families in Denver were assigned to a 20-year negative income tax plan: 112 families that were initially control families, and a second group, reassigned a few months later, of 57 families that were originally three-year experimentals. The latter maintained their originally assigned guarantee amount and benefit-reduction rate.[20]

Several complications arise from these reassignments, which are not discussed here.[21] As shown below, dealing with these complications is facilitated by the statistical techniques presented by Groeneveld, Hannan, and Tuma, which make use of multiple time periods for each couple. This allows the couple's particular and varying experimental status to be matched with the time period under investigation.

6. Fraud, reconciliations, and reporting biases. Problems associated with fraud, reconciliations, and reporting bias are all somewhat related. The issues raised can become complicated, but neither the Seattle-Denver researchers nor I view the observable evidence as indicating a serious bias from these sources. The clearest case of fraud is where an experimental couple falsely claims to be separated in order to collect

extra payments. Then the two main objectives of the experiment will suffer from a bias. Obviously, the true number of marital dissolutions will be overstated. The cost of the program is also overstated, assuming the amount of fraud in the experiment exceeds the amount that would exist in a nationally legislated plan. Reconciliations that occur during the experiment or soon after the experiment ends might be an indicator of this type of fraud. Reconciliations are also of concern because they affect the time in which children have both parents present.

Finally, the reporting for experimentals and controls differed in three ways. Controls reported their family composition once every four months, whereas the experimentals reported monthly as well as in triennial interviews. Groeneveld, Hannan, and Tuma dealt with this problem by using reported dissolutions from the same triennial interview. However, this does not equalize the incentives the experimental families have to report a dissolution or the likelihood that the multiple sources of reports by these families would spill over to their interviews.[22]

In summary, the features of the Seattle-Denver experiment that I find most likely to lead to important biases about marital dissolution in relation to a negative income tax are (a) the confounding of negative income tax treatments with training treatments; (b) the duration biases; and (c) attrition. The problem of sample size is one of reliability, not bias. The problem of the 20-year treatments is mainly that they create ambiguity in interpretation. Fraud, reconciliation, and reporting issues appear minor, but it is worth noting that the *direction* of bias is surely that of exaggerating the effect of a negative income tax on marital splits.

The Seattle-Denver Results Concerning Marital Dissolution

In their final report on marital dissolutions, Groeneveld, Hannan, and Tuma state that "the NIT plans tested in SIME/DIME dramatically increased the rates at which marriages dissolved among white and black couples . . ." They report an increase of "40 to 60 percent," and add:

> If one wishes a single set of numbers to summarize our findings one might choose the effects of the $3800 guarantee level treatments because it is closest to the current system and to likely welfare reforms. [Adjusting for attrition bias and restricting the estimates to couples with children] one obtains estimates of 58 percent increases in dissolution for blacks and 51 percent increases for whites. These are estimates of the experimental-control differences in the SIME/DIME population for the most feasible programs tested.[23]

As large as these increases are, they are smaller than those reported by these researchers in earlier published articles. In the first article the $3800 guarantee plan was estimated to increase "the annual probability of marital dissolution . . . by 63 percent for blacks, 194 percent for whites, and 83 percent for chicanas over what it would be in the control situation."[24] These results applied to the first two years of the experiment. For the first three years of the experiment the estimates of increases ranged from 57 percent to 129 percent (in a 1979 article) and from 24 percent to 114 percent (in a 1980 article).[25] In the latter article the results for Chicanos showed no increase in marital dissolutions for the treatment group.

The increase in marital splits among experimentals relative to controls was not attributable to a low proportion of splits in the control group. The proportion of white, black, and Chicano couples in the control group who experienced a marital dissolution during the first three years were 16, 24, and 20 percent, respectively. These percentages apply to the originally enrolled couples who did not drop out and they reflect the full three years of exposure to risk. The percentages are considerably higher than those reported by Sawhill et al. for poor couples in the Survey of Income Dynamics for a similar time period[26] or for comparable controls in the New Jersey negative income tax experiment.[27] Because the dissolution proportions among the controls in Seattle-Denver were high, the even higher level of dissolutions among the treatment group was noteworthy.

The Seattle-Denver results were surprising in two respects. First, previous research on the impact of AFDC on marital dissolutions had not prepared researchers to see a large effect from a negative income tax. After all, no firm evidence had been established for a large destabilizing effect of AFDC on marriages despite the fact that the system essentially provided "permanent" benefits to a wife if her marriage dissolved and no benefits to a married couple.[28] The Seattle-Denver experiment showed a large destabilizing impact from a program that did provide benefits to a couple that stayed together.

One possibility could have been that the destabilizing effect was attributable to the relative generosity of the negative income tax plans. In other words, the price effect was so large in its negative impact on stability that it dominated any positive income effect. Actually, this was not the case. In the second surprising result, the least generous negative income tax plans, which offered about the same or lower cash payments as did AFDC, induced the largest destabilizing effect, while the most generous plan had essentially no destabilizing effect. This is the opposite of the theoretically expected result discussed earlier.

The researchers rationalized the large destabilizing impact of the low-payment plan by suggesting a negligible income effect associated with the payments to the intact couple, while emphasizing a large price or "independence" effect of the payment to the divorced mother.[29] (The relative sizes of the income and independence effects are claimed to be reversed for the high-payment plan.) Although the ostensible payment to the divorced mother from the low-payment plan is no more than that available under AFDC, Groeneveld, Hannan, and Tuma suggest that the negative income tax payment is worth more because it carries less stigma, it is more certain to be received, and it involves low transaction costs because the woman does not have to file and wait for AFDC benefits.[30]

These reasons do not appear persuasive as explanations for a steady-state increase in "permanent" dissolutions, defined here as those dissolutions that prevail for many months. Instead, the explanations appear to apply to the timing of the dissolution rather than to its eventual incidence. The low transactions costs, for example, should only affect the timing. The fact that the woman may receive immediate monthly payments from the negative income tax plan increases the present value of the total payments received. However, this increase is trivial relative to the present value of AFDC payments because the latter are "permanent" and include noncash benefits, whereas the negative income tax cash payments will terminate within a year or two for most three-year plans and within three years or so for most five-year plans. Similarly, the certainty of the negative income tax payments should be important, if at all, only with respect to the timing of the marital dissolution.[31]

We do not have direct evidence for a stigma effect that discounts AFDC payments, and if we did we would need to know how a legislated negative income tax plan would enroll and monitor its participants to determine whether its administration would eliminate any stigma in receiving payments.[32] Apparently, many of the experimental families who were already receiving AFDC in Seattle-Denver were unwilling to shift to the negative income tax plans even when the latter paid larger cash transfer payments. These AFDC recipients did not want to jeopardize their Medicaid benefits or, in some cases, housing subsidies.[33] Persons already on AFDC may be inured to stigma, but their reluctance to shift to higher-paying negative income tax plans casts doubt on the strength of the stigma effect. Again, the stigma of AFDC might delay a woman's shift from receiving negative income tax payments to receiving AFDC benefits when the latter are larger, but this behavior implies that the negative income tax plan is affecting only the timing of the split, not its incidence.

Expressing these doubts about the explanation Groeneveld, Hannan and Tuma offer for their surprising results does not refute their explanation. Indeed, rather than attempt a thorough analysis of their explanations, the next section presents a reanalysis of their data.

An Empirical Reanalysis of the Seattle-Denver Data

The empirical reanalysis of the Seattle-Denver data presented in this paper will concentrate on couples with children and on marital dissolutions as the outcome of interest.[34]

The techniques of analysis follow closely the pioneering use by Groeneveld, Hannan, and Tuma of "event-history analysis," which appears preferable to any other statistical procedure for summarizing the results. These techniques focus attention on the *rate* of dissolution; that is, the sample's proportion of dissolutions per unit of time. The time-unit may be a year or as brief as a day, because the calendar date of the dissolution is recorded in the Seattle-Denver data. Remarriages, reconciliations, and subsequent dissolutions are not analyzed. The first dissolution ends the couple's record.

An important reason for using the rate instead of just measuring the incidence of a dissolution is that the treatment and control groups are exposed to the risk of dissolution for varying lengths of time. Even groups in the experiment for the same intended duration, whether three, five, or 20 years, may experience differential attrition. In particular, more attrition on the part of the control group could yield a spuriously lower incidence of dissolutions, and this bias would be all the greater if the control couples were more likely to divorce or separate after they dropped out of the experiment.

What is less obvious, however, is that the rate measure may bias (or exaggerate) marital dissolutions of treatment couples relative to control couples in the context of a short-duration experiment. As discussed earlier, the short-duration experiment provides an artificial incentive to divorce earlier rather than later. Previously, this intertemporal substitution bias was cited as a reason why the incidence of dissolution during a three-year experimental period might be higher than the incidence for the same three-year period under a permanent plan. The rate measure could increase this bias because even the same number (incidence) of dissolutions in a three-year period will produce different rates—a higher rate when the dissolutions occur early.

Table 1 illustrates these distinctions between rates and incidence, early and late dissolutions, and records with and without attrition. A hypothetical example of four couples (A,B,C, and D) and three periods is shown. Case II relative to Case I shows that later dissolutions yield a lower rate for the same incidence level. Case III relative to Case II shows how attrition will tend to understate the dissolutions if an incidence measure is used, whereas the rate will adjust for the varying exposures to risk.

Case IV relative to Case I is interesting because it reveals how the same rate may accompany different levels of incidence. Note that although Case I has the same rate as Case IV, one dissolution per four time-periods of exposure, there are only two dissolutions among the four couples in Case I and three dissolutions in Case IV. Our concern about the disruption of intact marriages and the consequences of this for the upbringing of children probably implies that Case IV is "worse." The important point is that a short-duration experiment should tend to produce Case I-type outcomes among the treatment group.

Table 1
Illustrative Examples of Differences in the Incidence and Rate of Marital Dissolutions

Legend: 1 = marital dissolution
0 = no marital dissolution
X = post-event period (no further record)
ATT = attrited (no information)

(Data beyond three years are unavailable.)

Period	Couple A	Couple B	Couple C	Couple D	
Case I					
1	1	1	0	0	Incidence: 2/4 = .50
2	X	X	0	0	
3	X	X	0	0	Rate: 2/8 = .25
Case II					
1	0	0	0	0	Incidence: 2/4 = .50
2	0	0	0	0	
3	1	1	0	0	Rate: 2/12 = .17
Case III					
1	0	0	0	0	Incidence: 2/4 = .50
2	0	0	0	ATT	
3	1	1	0	ATT	Rate: 2/10 = .20
Case IV					
1	0	0	0	0	Incidence: 3/4 = .75
2	0	0	0	0	
3	1	1	0	1	Rate: 3/12 = .25

A First Look, Using The Seattle Data

The yearly records of the Seattle experiment may be shown with three tables similar in form to Table 1. This will permit us to see the ingredients of analysis that will later be summarized in a statistical model, and the simplicity of the tables will facilitate some important observations about the data. Table 2 shows the number of couples and their record of attrition from the experiment. Tables 3 and 4 show the year-by-year record of marital dissolutions for white and black couples respectively.

Assignments to the four experimental groups were random within the stratifications of city, ethnicity, and the estimated level of the families' normal incomes.[35] As we shall see, income is not a major determinant of marital dissolution in this sample, so ignoring this variable in these tables still allows a fairly accurate picture of marital dissolutions in Seattle in response to the experimental plans. Adding the income variable would further dilute the already thin cell sizes. The five-year experimental cells are particularly small.

Seattle data are easy to interpret because the sample was stratified with only two ethnic groups and two duration groups. No family units were shifted among plans as they were in Denver. However, Seattle's outcomes are quite different from those in Denver in certain key respects. An overall assessment must wait for the statistical model for Seattle and Denver combined.

Table 2 shows Seattle's number (N) of originally married couples and the number who dropped out of the experiment, for each race for the four experimental statuses and the two duration groups. Overall, the attrition rate is 16 percent for the entire number (163 couples out of 1001), and the rates for the groups are: 20 percent for controls, 17 percent for the trainee group (TR), 15 percent for the pure negative income tax (NIT) group, and 13 percent for the group receiving the combined treatment of a negative income tax and training (NIT × TR). A striking

Table 2
Attrition in Seattle Negative Income Tax Experiment, by Treatment, Race, and Duration of Assignment

Racial Group and Duration of Experiment	Experimental Treatment											
	Control			Training			NIT			NIT x Training		
	N	ATT	R	N	ATT	R	N	ATT	R	N	ATT	R
White												
3-year	103	19	.18	96	12	.12	79	9	.11	148	18	.12
5-year	69	4	.06	33	4	.12	35	3	.09	45	5	.11
Total	172	23	.13	129	16	.12	114	12	.11	193	23	.12
Black												
3-year	47	24	.51	57	11	.19	50	8	.16	109	14	.13
5-year	44	6	.14	28	9	.32	25	9	.36	33	8	.24
Total	91	30	.33	85	20	.24	75	17	.23	142	22	.15
Total												
3-year	150	43	.29	153	23	.15	129	17	.13	257	32	.12
5-year	113	10	.09	61	13	.21	60	12	.20	78	13	.17
Total	263	53	.20	214	36	.17	189	29	.15	335	45	.13

Notes: N = number of couples at beginning of experiment. ATT = Attrition (number of couples who dropped out). R = rate of attrition (percentage dropping out). Table refers to originally married couples with children under age 21 at beginning of experiment. Cases where a spouse died during the experiment have been excluded.

finding is that the controls who were in the three-year duration group had a rate of attrition about three times that of the five-year controls (29 percent compared to 9 percent, for whites and blacks combined), despite the fact that the five-year controls had two additional years of exposure to the risk of dropping out of the experiment. The reason is surely that the five-year controls were assigned after the experiment had begun and after the considerable attrition in the beginning months of the experiment had occurred. Clearly, the designation to three-year and five-year control status was nonrandom with respect to attrition. In all subsequent analysis, all the controls in each year of their participation in the experiment are pooled to guard against the assignment being nonrandom with respect to the propensity to divorce or separate. Together they should constitute a random group.

Tables 3 and 4 show the essential information on marital dissolutions for white and black couples. For each year and for each experimen-

Table 3
Annual Rates of Marital Dissolution among Whites in Seattle Experiment[a]

Duration and Year in Experiment	Control[b] N	Control[b] D	Control[b] R	Training N	Training D	Training R	NIT N	NIT D	NIT R	NIT x Training N	NIT x Training D	NIT x Training R
3-yr, first	166	10	.060	93	7	.075	78	5	.064	143.5	10	.070
3-yr, second	149.5	9	.060	80.5	6	.075	71	7	.099	128	16	.125
3-yr, third	136	12	.088	72.5	4	.055	60.5	5	.083	107.5	6	.056
5-yr, first	included above			31.5	3	.095	36	2	.056	44	6	.136
5-yr, second	included above			27	1	.037	33.5	3	.090	37.5	3	.080
5-yr, third	included above			25.5	2	.078	29.5	1	.034	33.5	2	.053
5-yr, fourth	50	1	.020	24	1	.042	27.5	3	.109	31	0	.000
5-yr, fifth	49	1	.019	23	0	.000	24	0	.000	30.5	1	.033
Totals												
3-yr	451.5	31	.069	246	17	.069	209.5	17	.081	379	32	.034
5-yr	550.5	33	.060	131	7	.053	145.5	9	.062	176.5	12	.068
Total	550.5	33	.060	377	24	.064	355	26	.073	555.5	44	.079

	Control Dissolution Rate (Adjusted for Attrition[c])	Training Ratio[d] No Adj.	Training Ratio[d] Att. Adj.	NIT Ratio No Adj.	NIT Ratio Att. Adj.	NIT x Training Ratio No Adj.	NIT x Training Ratio Att. Adj.
3-yr	.070	1.00	.99	1.17[d]	1.13[d]	1.22	1.16
5-yr	.062	.88	.88	1.03	.90	1.13	.94
Total	.062	1.07	1.03	1.22	1.11	1.32	1.16

Notes follow Table 4. See also Notes to Tables 5 and 6

N = number of person- (couple-) years of exposure to risk. In the first year, N = the number of couples at the beginning of the experiment minus one-half of the number of couples dropping out who did not divorce or separate.

D = number of marital dissolutions.

R = rate of marital dissolution, measured as the proportion of dissolutions per years at risk.

Table 4
Annual Rates of Marital Dissolution among Blacks in Seattle Experiment[a]

Duration and Year in Experiment	Control[b] N	D	R	Training N	D	R	NIT N	D	R	NIT x Training N	D	R
3-yr, first	83.5	8	.096	54.5	4	.073	48.5	3	.062	105	23	.219
3-yr, second	65	6	.092	46.5	1	.022	43	3	.070	78.5	11	.140
3-yr, third	53	3	.057	42.5	2	.047	38.5	3	.078	65.5	4	.061
5-yr, first	included above			26	0	.000	23.5	2	.085	31.5	5	.159
5-yr, second	included above			23	4	.174	19.5	2	.103	24	3	.125
5-yr, third	included above			18.5	0	.000	16.5	0	.000	20.5	3	.140
5-yr, fourth	30.5	2	.066	17	1	.059	15.5	0	.000	17	2	.118
5-yr, fifth	28	1	.036	15	1	.125	13.5	1	.074	14	1	.071
Totals												
3-yr	201.5	17	.084	143.5	7	.049	130	9	.069	249	38	.153
5-yr	260	20	.077	99.5	6	.060	88.5	5	.056	107	14	.131
Total	260	20	.077	243	13	.053	218.5	14	.064	356	52	.146

	Control Dissolution Rate (Adjusted for Attrition[c])	Training Ratio[d] No Adj.	Att. Adj.	NIT Ratio No Adj.	Att. Adj.	NIT x Training Ratio No Adj.	Att. Adj.
3-yr	.088	.58	.56	.92	.86	1.82	1.67
5-yr	.082	.78	.73	.73	.63	1.70	1.48
Total	.082	.69	.65	.90	.74	1.90	1.70

[a]See also Notes to Tables 5 and 6.
N = number of person- (couple-) years of exposure to risk. In the first year, N = the number of couples at the beginning of the experiment minus one-half of the number of couples dropping out who did not divorce or separate.
D = number of marital dissolutions.
R = rate of marital dissolution, measured as the proportion of dissolutions per years at risk.

Notes to Tables 3 and 4

[a]Tables refer to originally married couples, with children under 21 at the beginning of the experiment; cases where spouse died during the duration of the experiment are excluded. Dropouts contribute to the total dissolution rate (or proportion) during the year that they leave the experiment by assuming they represent one-half year of exposure to risk. If they report a dissolution, the dissolution is included in the total number of dissolutions for that year.

[b]All controls are aggregated during the first three years of the experiment. Only the five-year controls are measured during the fourth and fifth years of the experiment.

[c]The attrition adjustment has two parts. First, the dissolution rate is assumed to be 25 percent greater among the control dropouts for whom no information is available. Second, the dissolution rate is assumed to be 50 percent less for experimental (treatment) dropouts who were assigned to a negative income tax plan and for whom no information is available. No attrition adjustment is made for the training-treatment group.

[d]From Table 3, the calculation of the two ratios for the NIT three-year sample of whites is demonstrated as follows:

1.17 = .081/.069 = NIT average rate for the three-year group / control average rate for the three years.

1.13 = .079/.070 = The attrition-adjusted dissolution rate for the NIT group, assuming the NIT dropouts with no information about dissolutions have a dissolution rate one-half as large as the remaining NIT couples / adjusted rate of control attrition.

All other ratios in Tables 3 and 4 are derived in the same way.

tal group, the number of dissolutions is recorded along with the number of person-years (or couple-years) at risk. A couple that drops out of the experiment in any year is assumed to provide a half-year of exposure to the risk of dissolution. A dissolution recorded for a couple that drops out in a year is counted in that year. In the following year, only the continuing and still-intact couples are at risk. The three-year and five-year treatment groups—Training (TR), Negative Income Tax (NIT), and NIT × TR—are separately recorded. The controls are pooled in the first three years, but only the five-year group is recorded for the fourth and fifth years of the experiment.

Several interesting results in Tables 3 and 4 will be shown to hold up in the final analysis when all data are used and when a number of exogenous control variables are held constant statistically.

1. Looking at the average yearly dissolution rate and the ratio of these averages to the corresponding control group, no consistent pattern emerged regarding the three-year groups versus the five-year groups or regarding three of the four experimental groups—control, Training, or the NIT group.

2. The experimental group, NIT × TR, shows a higher dissolution rate, and among blacks, the higher rate for the three-year experimentals is statistically significant at conventional levels.[36]

3. The cell sizes are too small to detect a time trend in the dissolution rates, although there is a hint of a downward trend in the experimental groups, as, for example, when the third year is compared with the first and second in the three-year groups, and the fourth and fifth years are compared with the first two years in the five-year groups. Intertemporal substitution will be examined below in more detail, although our full analysis of the time-dependence of dissolutions is not completed.

4. In the light of earlier findings by Groeneveld, Hannan, and Tuma, an unexpected result from the Seattle data is that the average annual dissolution rates among the three-year NIT and NIT × TR groups are *higher* than among their five-year counterparts. Moreover, the NIT/Control ratios of the three-year dissolution rates are higher than these ratios for the five-year dissolution rates. Again, intertemporal substitution is a possible explanation.

5. An adjustment for attrition bias can be demonstrated with these data and it turns out to be a fairly minor adjustment. Couples who dropped out and who did not report a marital split represent a certain number of subsequent unknown person-years. In sensitivity tests, the marital dissolution rate is assumed to be 25 percent higher among controls who dropped out; say, .075 per year instead of .06. When this adjustment is used, the *overall* average dissolution rate of controls is raised by .001 or .002, from, say, .06 to .061 or .062. In the next step, the

dissolution rate among dropout experimentals who were eligible for negative income tax payments is assumed to be 50 percent smaller than the rate among experimentals who stayed; say, .03 instead of .06. Applying this rate to the unknown person-years among experimentals who dropped out serves in practice to lower the overall average dissolution rate of the NIT or NIT × TR groups by .002 to .005. Thus, the attrition adjustment could change the experimental/control ratio of dissolution rates by around 5 to 10 percentage points; for example, from .06/.06 = 1 to .058/.061 = .95 or to .056/.062 = .90. These calculations merely illustrate the sensitivity of the estimates to an attrition adjustment. They will now be set aside until the concluding section of the paper.

The Full Sample and The Use of an Exponential Model

Table 5 provides a relatively complete summary of the dissolution "effects" of the various experimental statuses, using the full information for both cities, all five years of the regular experiment, and the sixth and seventh years for the small number of Denver couples assigned to the 20-year duration plans. Also, the full set of control variables used by Groeneveld, Hannan, and Tuma is included in the statistical model. The reported coefficients under the column headed b show the approximate percentage effects of the independent variables on the marital dissolution rate. The coefficients of the experimental categorical variables are related to the "multipliers" of the dissolution rate of the omitted base group of controls.[37]

In Table 5 the original numbers of couples for each group in each plan are shown in brackets, and we see the small number of families in the 20-year plans. All of these 20-year couples were originally in another group, so the total number of couples at the beginning of the experiment is given by the totals for the 3- and 5-year groups, along with the controls: 272 white controls, 182 black controls, and 93 Chicano controls. Hence, the number of observations per group may well be too small for the analysis of a relatively uncommon event like marital dissolutions. (Table 6 shows the statistical results when using fewer groups but with larger cell sizes.) The general lack of statistical significance also discourages spending much effort in investigating the effects of the still smaller subgroups of experimental treatments, such as the three training programs and the eleven (or even three) negative income tax plans. Groeneveld, Hannan, and Tuma extensively analyzed the results for high, medium, and low negative income tax plans.

The statistical model underlying the results shown in Tables 5 and 6 is the discrete-time analogue of the continuous-time model used by Groeneveld, Hannan, and Tuma; that is, their exponential rate model.

Table 5
Estimated Effects of Independent Variables on Dissolution Rates: Full Set of Interactions, Treatment x Duration x Site [a]

Independent Variable[c]	Whites[b] b	t-ratio	original n	Blacks[b] b	t-ratio	original n	Chicanos[b] b	t-ratio	original n
Constant	−2.02	(3.50)***		−1.37	(2.69)***		−1.31	(1.22)	
Normal earnings ($000's)									
0-1	−.03	(.08)		−.31	(.73)		.19	(.25)	
1-3	−.19	(.48)		−.29	(.82)		.54	(.73)	
3-5	−.42	(1.04)		−.31	(.86)		.48	(.64)	
5-7	−.67	(1.62)		−.46	(1.28)		.43	(.56)	
7-9	−.78	(1.77)*		−.10	(.27)		.64	(.81)	
9-13	−.63	(.78)		1.23	(2.13)**		−3.80	(1.12)	
Unreported	−2.83	(.85)		.37	(.35)		−3.74	(.59)	
Duration of marriage	−.08	(4.78)***		−.06	(4.09)***		−.04	(1.57)	
Wife's age	−.01	(.50)		−.03	(2.65)***		−.05	(2.12)**	
Wife's ed, 12	−.23	(1.71)*		−23	(1.68)*		−.40	(1.78)*	
Wife's ed, 12	−1.16	(2.08)**		.66	(1.41)		. . .	. . .	
Young Children	−.26	(1.40)		−.23	(1.39)		−.70	(2.36)***	
AFDC, pre	.30	(1.74)*		.04	(.20)		.70	(3.01)***	
TR x 3 x S	.21	(.73)	96	−.26	(.63)	57	. . .	. . .	. . .
TR x 5 x S	−.08	(.21)	33	−.11	(.25)	28	. . .	. . .	. . .
TR x 3 x D	.12	(.39)	83	.57	(2.19)**	80	−.12	(.36)	61
TR x 5 x D	−.41	(.81)	33	.08	(.22)	33	.20	(.50)	30
NIT x 3 x S	.19	(.69)	83	−.13	(.36)	50	. . .	. . .	. . .
NIT x 5 x S	.26	(.73)	35	−.29	(.61)	25	. . .	. . .	. . .
NIT x 3 x D	−.12	(.33)	57	.36	(1.00)	39	−.49	(1.32)	70
NIT x 5 x D	.06	(.20)	44	.88	(3.42)***	44	−.01	(.02)	30
NIT x 20 x D	.09	(.25)	35	.07	(.17)	23	−.53	(.96)	15
(NIT.TR) x 3 x S	.31	(1.36)	148	.73	(3.20)***	109	. . .	. . .	. . .
(NIT.TR) x 5 x S	.13	(.40)	45	.76	(2.43)**	33	. . .	. . .	. . .
(NIT.TR) x 3 x D	.20	(.85)	142	.74	(3.17)***	96	.07	(.26)	152
(NIT.TR) x 5 x D	−.02	(.05)	49	−.12	(.35)	39	−.26	(.83)	59

[a]See the text for a specification of the statistical model to estimate the rate of marital dissolution.

[b]b = Multiplier, approximately equal to the percentage effect of the independent variable on dissolution. See text footnote 37.

n = number of couples at beginning of experiment; DF = degrees of freedom, based on number of 6-month time periods per couple at risk (minus the number of independent variables). Whites: n = 1120, DF = 7120; Blacks: n = 815, DF = 4732; Chicanos: n = 495, DF = 2960.

[c]Independent variables are defined for their values at the beginning of the experiment:

- Duration of marriage in years.
- Wife's age, in years.
- Wife's ed (education): the category "less than 12 years of schooling" is the omitted category.
- Young children: 1 if a child under 6 years of age is present; 0 otherwise.
- AFDC, pre: 1 if wife had participated in AFDC in the year prior to enrollment; 0 otherwise.

TR x 3 x S = Training treatment only and 3-year duration and in Seattle. Other treatment statuses are defined accordingly.

(NIT.TR) = The combined treatment of an NIT plan *and* training.

*Statistically significant at the 10 percent level, two-tailed test.
**5 percent level.
***1 percent level.

Table 6
Estimated Effects of Independent Variables on Dissolution Rates: Summary Results, Combining Duration and Site Groups[a]

	Whites[b]		Blacks[c]		Chicanos[d]	
Variable	b	t	b	t	b	t
Constant	−1.92	(3.36)***	−1.30	(2.53)**	−1.30	(2.79)***
Normal earnings ($000)						
0−1	−.01	(.02)	−.37	(.88)	.12	(.16)
1−3	−.19	(.49)	−.34	(.96)	.49	(.66)
3−5	−.41	(1.02)	−.34	(.96)	.39	(.52)
5−7	−.66	(1.61)	−.48	(1.33)	.33	(.42)
7−9	−.78	(1.77)*	−.20	(.53)	.53	(.67)
9−13	−.58	(.71)	1.13	(1.98)**	−3.85	(1.14)
Unreported	−2.71	(.81)	.20	(.19)	−3.71	(.59)
Denver	−.16	(1.20)	.12	(.90)	...	...
Duration of marriage	−.08	(4.79)***	−.07	(4.29)***	−.04	(1.51)
Wife's age	−.01	(.59)	−.03	(2.79)***	−.05	(2.10)**
Wife's ed, 12	−.23	(1.72)*	−.20	(1.46)	−.41	(1.85)*
Wife's ed, > 12	−1.17	(2.10)**	.77	(1.66)	...	...
Young Children	−.25	(1.39)	−.26	(1.54)	−.69	(2.33)**
AFDC, pre	.29	(1.66)*	.00	(.03)	.66	(2.89)***
TR	.03	(.14)	.15	(.69)	.10	(.33)
NIT	.16	(.82)	.27	(1.28)	−.30	(1.00)
(NIT x TR)	.17	(.97)	.57	(3.06)***	−.01	(.03)

[a]See notes to Table 5. Denver = 1 if family lives in Denver, O if in Seattle.
[b]DF = 7129. (DF = Degrees of Freedom)
[c]DF = 4741.
[d]DF = 2964.

Define $P(t)$ as the probability that a couple experiences a dissolution at time t, conditional upon the couple being at risk at time t. The usual logit transformation of $P(t)$, related to a linear specification of explanatory variables, is:

$$\ln[P(t)/(1 - P(t))] = a + bx.$$

As the interval of time becomes smaller, the data approach continuous time. The specification of the dependent variable that provides an exact analogue to the continuous-time model is:[38]

$$\ln[-\ln(1 - P(t))].$$

Let $y = \ln[-\ln(1 - P(t))]$; T is a vector of treatment variables, and X is a vector of exogenous determinants of marital dissolution. The statistical model in Tables 5 and 6 has this double-log functional form and uses discrete data for six-month time periods:

$$y = T'a + X'\beta.$$

Estimation is by maximum likelihood logit analysis, using the GLIM statistical package.

Table 7
Estimated Effects of Independent Variables on Marital Dissolution Rates (using the same samples and variables as Groeneveld, Hannan, and Tuma)

	Black				White				Chicano			
	Cain		GHT		Cain		GHT		Cain		GHT	
	b	t	b	t	b	t	b	t	b	t	b	t
Constant	−.55	(.73)	.07	(.10)	−2.35	(3.44)***	−1.59	(2.30)	−1.76	(1.63)	−1.01	(.95)
Normal Earnings ($000's)												
0-1	.35	(.98)	.40	(1.11)	.78	(1.66)*	1.01	(2.30)**	.00	. . .	−.02	(.02)
1-3	−.13	(.36)	−.11	(.31)	.81	(2.68)***	.89	(2.87)***	−.19	(.43)	−.19	(.43)
3-5	−.17	(.76)	−.10	(.45)	.60	(2.31)**	.66	(2.44)**	.17	(.46)	.14	(.38)
5-7	−.31	(1.56)	−.28	(1.40)	.45	(1.86)*	.52	(2.00)**	.03	(.10)	.06	(.17)
7-9	−.37	(1.82)*	−.34	(1.70)*	.19	(.75)	.23	(.85)	−.05	(.14)	−.01	(.03)
9-13	−.53	(.52)	−.59	(.58)	1.28	(1.26)	1.37	(1.33)	. . .	. . .	7.52	(.11)
Denver	.28	(2.01)**	.28	(2.00)**	−.18	(1.28)	−.20	(1.43)	. . .	. . .	. . .	. . .
Dur. Marriage	−.05	(3.33)***	−.05	(5.00)***	−.09	(5.05)***	−.10	(3.33)***	−.04	(1.41)	−.03	(1.00)
Age-W	−.01	(.72)	−.01	(.50)	.01	(.42)	.01	(.50)	−.05	(1.61)	−.06	(2.00)
Ed-W	.00	(.08)	.01	(.20)	−.06	(1.89)*	−.08	(2.00)**	−.03	(.55)	−.03	(.60)
Age-H	−.03	(2.14)**	−.03	(3.00)***	−.02	(1.15)	−.02	(1.00)	.01	(.20)	.00	. . .
Ed-H	−.09	(2.61)**	−.08	(2.67)***	.01	(.31)	.02	(.67)	−.02	(.38)	.03	(.75)
Children, n	.07	(1.42)	.08	(1.60)	.04	(.75)	.05	(.83)	.12	(1.35)	.13	(1.44)
Young Children	−.24	(1.47)	−.29	(1.81)*	−.27	(1.67)*	−.29	(1.81)*	−.38	(1.35)	−.43	(1.54)
AFDC,	.04	(.22)	.05	(.26)	.45	(2.40)**	.50	(2.63)**	.61	(2.50)**	.67	(2.79)
M-1	.42	(2.00)**	.45	(2.25)**	.29	(1.40)	.32	(1.52)	.52	(1.88)*	.52	(1.93)
M-2	.24	(1.22)	.30	(1.50)	.15	(.74)	.14	(.70)	.12	(.41)	.13	(.45)
M-3	.25	(1.16)	.26	(1.24)	.34	(1.69)*	.33	(1.65)*	.22	(.67)	.18	(.56)
M, 5 yr	−.24	(.96)	−.38	(1.46)	−.15	(.57)	−.29	(1.07)	.00	. . .	−.04	(.10)
NIT	.41	(2.05)**	.45	(2.14)**	.36	(1.70)*	.43	(1.95)*	.05	(.17)	.01	(.03)
NIT, 3 yr	−.24	(1.05)	−.30	(1.30)	−.24	(1.02)	−.33	(1.38)	−.11	(.30)	.00	. . .

Denver =1 if family lives in Denver; O if in Seattle.
Dur. Marr. = years married at beginning of experiment.
Age-W = age of wife; Age H =age of husband.
Ed-W = Wife's education (years) Ed-H = husband's.
Children, n = number of children.
Young Children = 1 if a child under six years of age is present; O otherwise.

M-1: least generous training program.
M-2: more generous training program.
M-3: most generous training program.
M-5 yr: if training subsidy variable is for 5 years.
NIT = Pure NIT and NITxTR pooled.
NIT, 3 yr. = 1 if family was in NIT or NITXTR experimental status and in the 3-year duration group; O otherwise.

NOTE: This table is a replication of Table 5.1.A in Groeneveld, Hannan, and Tuma, "Marital Stability," *Final Report*, p. 367. The GHT columns refer to a continuous-time model; the other columns refer to a discrete-time model.

Using time intervals of six months (instead of one year as in Tables 3 and 4), it is possible to replicate closely the results of Groeneveld, Hannan, and Tuma when using the same data. See Table 7 for the replication of their results for all originally married couples, including those without children, for the first three years of the experiment.[39]

The outcomes of the experimental plans shown in Table 5 are not easy to summarize. No treatment variables are statistically significant among white and Chicano samples, and imposing zero coefficients on all five variables defining any of the three experimental plans, TR, NIT, or NIT × TR, does not significantly worsen the fitted relation. In terms of the pure NIT plans, six of the 13 coefficients are negative, showing a stabilizing effect on marriages, although all are statistically insignificant. Seven of the 13 are positive, showing a destabilizing effect, but only one is statistically significant: 0.88 for the 44 black families in the five-year NIT program in Denver. The pure NIT plan does not show a consistent destabilizing effect for any of the three ethnic groups.

The NIT × TR plan has a large and significant destabilizing effect on blacks. These plans have no statistically significant effects among whites or Chicanos, although the direction of the effects for whites is mainly positive. Finally, the five-year duration plans tend to be *less* destabilizing than the three-year plans in most comparisons.

Table 6 summarizes the separate experimental plans for each ethnic group, pooling the sites and durations to build up the sample size and to summarize an overall effect of each of the three experimental treatments. Of the nine experimental coefficients, only one is statistically significant, .57 for blacks in the NIT × TR plan. Of the three coefficients for the pure NIT, none is statistically significant, and one (for Chicanos) is negative. The pure NIT coefficient for blacks, .27, is large enough to cause concern, but it is not reliably estimated, and it is smaller in absolute value than the statistically insignificant negative coefficient, —.30, for Chicanos. A weighted average for the three ethnic groups, using the sample proportions of couples in each ethnic group as weights, is .10. For the relatively rare event of a marital dissolution, an effect of this magnitude has no practical significance.

Summary

The results shown in Tables 5 and 6 do not justify the conclusion that a negative income tax program, by itself, would lead to an increase in marital breakups among married couples with children. Three telling results argue against such a claim.

1. First, as shown in Table 5, the sample sizes for the cells that describe the pure NIT plan are not large enough to warrant any confidence in such a conclusion, unless the results for the different cities,

time durations, and ethnic groups were so consistent that the samples could be pooled. But the results are not consistent even with respect to sign.

2. Second, the summary estimate achieved by combining all durations and sites in Table 6 shows inconsistent signs and an overall small quantitative effect (.10) for the pure NIT treatment.

3. Third, the results have not been adjusted for attrition bias or for reconciliations. Attrition bias is, of course, unknown, and it is merely on the basis of prior theorizing that the adjustments suggested earlier diminished the dissolution rate among experimentals relative to controls. If the reader agrees that an adjustment is called for, perhaps a summary estimate would entail multiplying all the positive NIT coefficients by .95 and all the negative coefficients by 1.05. Reconciliations are observable during the course of the experiment, and although they have not been used in this paper, the findings of Groeneveld, Hannan, and Tuma, which we have corroborated, show that reconciliations are more prevalent among the experimental families. This indicates that a measure based on the fraction of time that the parents are separated is likely to show less instability than did the rate of first dissolutions, which Groeneveld, Hannan, and Tuma and this study have emphasized.

Several qualifications must be noted about these conclusions regarding the negative income tax and marital stability. One, which is probably not serious, is that the reanalysis has not examined the paradoxical result of Groeneveld, Hannan, and Tuma whereby the least generous negative income tax plans had the largest destabilizing effect, and the most generous plans the least destabilizing effect. As stated above, it is difficult to believe that the sample sizes justify these conclusions. Second, no explanation emerges for the significant destabilizing results for the combined negative income tax-training treatment. The training plan, by itself, had an even smaller destabilizing effect than did the pure negative income tax, on average and across all ethnic groups. So it is not plausible to portray the training program as the villain in promoting marital dissolutions. The destabilizing effect from the treatment that combined a negative income tax and training program, particularly among black families, remains not well explained.

Also unresolved is the issue of conflicting biases in short-duration experiments. Are the experimental outcomes exaggerated, via the intertemporal substitution effect? Or are they understated, via the lesser present value of the incentives? This issue is particularly interesting because Groeneveld, Hannan, and Tuma had emphasized that the dissolution effect was understated by a short-duration experiment. Their evidence was their report of a stronger destabilizing effect of the five-year plans, and their claim was that a permanent plan would have even larger destabilizing effects than the five-year plan. The results in Table 5 appear to refute these claims. The tendency for three-year plans to show larger annual rates of dissolution than the five-year plans is consistent with intertemporal substitution playing a significant role.

One obstacle to further analysis of this issue is the fact that the five-year controls were nonrandomly selected from among the control groups. The 20-year plans do not offer much help on this question. Overall, these groups had lower average annual dissolution rates over the years when they were assigned, which were years three through seven. However, they also were nonrandomly selected. Both the five-year controls and the 20-year groups demonstrated the trait of "stability" by virtue of their not having dropped out during the first several years of the experiment. There was no practical (or statistically significant) difference between the 20-year treatment and control groups (results not shown), but the sample sizes were small.

What explains the contrast between the large and dramatic destabilizing results of the earlier analysis compared to the smaller and inconsistent patterns shown in Tables 5 and 6? The analysis of this question is incomplete, but all of the following appear to contribute to the new mild results:

1. Separating the NIT plan from the NIT × TR plan;
2. Eliminating couples without children from the analysis;
3. Including the couples in the 20-year plans during the years in which these plans were in effect;
4. Permitting the 20-year couples to be part of their originally assigned plans during the years when the 20-year plan was not in effect;[40]
5. Including information on marital dissolutions even if they were recorded after the date of an attrition report.

The last item refers to the apparent decision of Groeneveld, Hannan, and Tuma to record the couple as having dropped out but not as having dissolved their marriage, if attrition was reported first. Our procedure helps in a small way to correct for the alleged attrition bias. There are more dropouts among controls, and if dropouts have high marital dissolution rates, the post-attrition information helps correct for the bias.

Probably the greatest difference between their conclusions and those of this study is that they emphasized results from the first three years of the experiment including the five-year negative income tax plans. It turns out that the results for the full five years of the experiment are less adverse regarding the effect of a negative income tax on marital stability. Also, the large impact of the five-year plans they report during the first three years are dissipated when the separate plans and extra years of the experiment are included.

The prevalence of reconciliations among the sample, particularly among the experimentals, may provide a clue to the high volume of dissolution and may suggest a way in which a negative income tax plan might deal with dissolutions. Consider that the families in the Seattle-Denver plans were eligible to receive a monthly payment if their incomes were sufficiently low. Surely they would realize that a departure by a spouse with earnings, particularly the husband, would lead to a quick and sharp increase in their monthly payment. The temptation to report frequent dissolutions, along with frequent reconciliations, may be strong on the part of a small percentage of the families. Only a few dissolutions are required to make a substantial difference in the rate, when the sample sizes are small and the rates are as low as 6 percent or less per year. AFDC might provide larger benefits to ''permanent'' dissolutions, but, as Groeneveld, Hannan and Tuma have suggested, the fixed costs of ''going on'' AFDC may dissuade mothers from doing so if the separation is believed to be temporary. Perhaps a negative income tax requires a longer waiting period before higher payments are made. Obviously, more than speculation is needed to determine if the phenomenon of ''temporary'' dissolutions explains the high dissolution rate among black couples covered by the NIT × TR plans. Our future work will examine this issue.

The basic finding is, however, not about reconciliations. Rather, the pure negative income tax plan had neither a practical nor a statistically significant destabilizing effect on the marriages of already married couples with children.

---

[1]Only two among many papers by the Seattle-Denver research staff will be cited at this point. The first published article, which was especially important for being first, was Michael T. Hannan, Nancy B. Tuma, and Lyle P. Groeneveld, "Income and Marital Events: Evidence from an Income Maintenance Experiment," *American Journal of Sociology, 82*, 1977, pp. 1186-1211. The final version of their findings is "Marital Stability," in *Final Report of the Seattle-Denver Income Maintenance Experiment, Volume 1, Design and Results*, SRI International, May 1983, Part V, pp. 257-383. Volume 1 will be cited hereafter as *Final Report*.

[2]Groeneveld, Hannan, and Tuma review much of the literature up to 1980 in *Final Report*, pp. 264-266. See also David Ellwood and Mary Jo Bane, "The Impact of AFDC on Family Structure and Living Arrangements," Report to U. S. Department of Health and Human Services, 1984.

[3]Andrew J. Cherlin, *Marriage, Divorce, and Remarriage*, Cambridge, Mass.: Harvard University Press, 1981, pp. 10-11, 23.

[4]One qualification is AFDC-UP, with UP standing for "unemployed parent," an optional program offering AFDC to poor married couples whose principal breadwinner is unemployed. Now adopted by half the states, the program nevertheless has a very small number of couples participating.

[5]See the arguments and citations for a positive effect of income on marital stability in Groeneveld, Hannan, and Tuma, *Final Report*, pp. 261-64.

[6]John Bishop, "Jobs, Cash Transfers, and Marital Instability: A Review and Synthesis of the Evidence," *Journal of Human Resources*, Summer 1980.

[7]Negative income tax plans of roughly the same level of generosity can differ in their income guarantees and benefit-reduction rates, but I will not discuss the differential effects on marital stability of these sorts of variations. The trade-off between guarantees and the benefit-reduction rates was not an important issue in the analysis of marital dissolutions in the negative income tax experiments.

[8]See the interesting exchange of questions and responses on this issue in a Senate hearing on welfare reform that is reported in Gilbert Y. Steiner, *The Futility of Family Policy*, Washington, D. C.: The Brookings Institution, 1982, pp. 101-102.

[9]Groeneveld, Hannan, and Tuma cite several studies that support this argument in their research review. See *Final Report*, pp. 265, 270. See also Ellwood and Bane, "The Impact of AFDC," 1984.

[10]I am grateful to James Albrecht, for aiding my consideration of this issue. See his "Hare [sic] Today, Gone Tomorrow: Divorce, Unemployment, and Other Sorry States," in K. Lang and J. Leonard, eds., *Unemployment and the Structure of Labor Markets*, London: Basil Blackwell, 1986.

[11]For further discussion of some of the features of the Seattle-Denver experiment that created incentives for creating new family units, see Gary Christophersen, *Final Report of the Seattle-Denver Income Maintenance Experiment, Volume 2*, Mathematica Policy Research, May 1983, especially pp. 37-51.

[12]*Final Report*, 266-269.

[13]Arlene Waksberg, "Overview of Master File System with Particular Attention to the Operational Flow of Family Composition Data," p. 24. This was originally published by SRI in January 1979 and is reprinted in the documentation for the Seattle-Denver data tapes provided by the National Archives. Waksberg noted that obtaining "Affidavits of Separation" was "done in a nonrigorous fashion" (p.24).

[14]The description of the training-and-counseling treatments used in the Seattle-Denver experiment does suggest their individuality along several dimensions. See Katherine P. Dickinson and Richard W. West, "Impacts of Counseling and Education Subsidy Programs," *Final Report*, especially pp. 201-216.

[15]*Final Report*, pp. 291-292.

[16]In the Seattle-Denver experiment a minimum monthly payment of $20 was paid to experimental families who filed their monthly reporting forms. Smaller payments were made to a subset of control families who were asked to file reports. These payments undoubtedly kept attrition lower than it otherwise would have been. See Christophersen, pp. 65-68.

[17]For the evidence supporting these generalizations about attrition, see Robert G. Spiegelman, "History and Design," *Final Report*, pp. 30-32.

[18]For references to personal problems, participation in public welfare programs, geographic mobility, and marital dissolution in connection with attrition, see David N. Kershaw and Jerilyn Fair, *The New Jersey Income-Maintenance Experiment, Vol. I,* New York: Academic Press, 1976, pp. 119-127.

[19]On page 239 of the microfiche description of the Seattle-Denver experiment that is provided by the National Archives we are told only that: "Later, a sample of the control families was selected to be interviewed for the same length of time as 5-year financials [5-year NIT experimentals]."

[20]See Philip K. Robins and Gary L. Steiger, "An Analysis of the Labor Supply Response of Twenty-Year Families in the Denver Income Maintenance Experiment," SRI unpublished paper, April 1980.

[21]See the longer version of this paper, available as a Discussion Paper from the Institute of Research on Poverty, University of Wisconsin, Madison, Wisconsin. This will hereafter be cited as Cain, "Discussion Paper."

[22]For a discussion of these reporting differences and the judgment that they led to a slight bias toward more reporting of splits by experimental families than by control families, see Waksberg, "Overview of Master File System."

[23]Groeneveld, Hannan and Tuma, *Final Report,* p. 357. On page 310 the authors suggest that "reasonable adjustments for attrition bias are on the order of 10 percent for blacks and 5 percent for whites." Also, the dissolution effect they report for all couples is about 5 percent higher than that for couples with children. Therefore, the researchers' estimates of 58 and 51 percent reported above correspond in their other reported results to estimates of 64 and 56 percent.

[24]Hannan, Tuma and Groeneveld, "Income and Marital Events," 1977, p. 120.

[25]Tuma, Hannan and Groeneveld, "Dynamic Analysis of Event Histories," *American Journal of Sociology, 84,* January 1979, pp. 835-836; and Groeneveld, Hannan and Tuma, "The Effects of Negative Income Tax Programs on Marital Dissolution," *Journal of Human Resources, 14,* Fall 1980, pp. 664-665.

[26]Isabel V. Sawhill, George E. Peabody, Carol A. Jones, Steven B. Caldwell, "Income Transfers and Family Structure," Washington, D.C.: The Urban Institute, September 1975.

[27]For the evidence and citations for these claims, see Cain, "Discussion Paper."

[28]Groeneveld, Hannan, and Tuma in particular expressed skepticism that the AFDC system had an important destabilizing effect on marriage. See *Final Report,* p. 266.

[29]For further discussion of Groeneveld, Hannan, and Tuma's rather complicated explanation of their findings regarding the different levels of negative income tax plans, see Cain, "Discussion Paper."

[30]*Final Report,* pp. 358-362; "Income and Marital Events," 1977, pp. 1208-1209.

[31]Steiner also questioned the "certainty" hypothesis, but it is not clear that he was referring to the short run of immediate payments. See Steiner, 1982, p. 109. On the other hand, if Groeneveld, Hannan, and Tuma claim that certainty affects the steady state dissolution rate in the short run, they must argue that the temporary wait for AFDC benefits is sufficient to permanently dissuade the mother from her intended "permanent" separation or divorce. Would the woman choose a "permanent" divorce if she can receive negative income tax payments for, say, three months but not so choose if she has to wait three months for AFDC benefits?

[32]Bishop argues for a stigma effect of transfer payments that destabilizes marriages, but his hypothesis is nearly the opposite of that of Groeneveld, Hannan, and Tuma. In Bishop's view, negative income tax payments stigmatize the husband, demeaning his role as a provider, and in this way promote marital breakups. See Bishop, "Jobs, Cash Transfers, and Marital Instability," 1980. In contrast, the Seattle-Denver researchers argue that because negative income tax payments have relatively little stigma, they will be chosen by a divorced mother as a source of income support that she has shunned when it is available through AFDC.

[33]Christophersen, *Final Report, Vol. 2,* pp. 10-12.

[34]We use the data for the same couples as were used by Groeneveld, Hannan, and Tuma, except that we restricted our analysis to couples with dependent children (under age 21) at the beginning of the experiment, and we discarded a few cases in which either spouse died. Groeneveld, Hannan, and Tuma had discarded cases in which the wife died. (I use the plural pronoun in discussing the reanalysis to acknowledge the contribution of Douglas Wissoker.) Although an analysis of a related outcome that measures the time

when children are without both parents as a result of a marital dissolution is underway, these results are not presented here. This latter outcome is based on the information on reconciliations, which will be only briefly referred to in this paper.

[35]The level of normal earnings, in seven categories, is defined as "expected family income for the year prior to the start of the experiment, and was derived from preenrollment interview data." Christophersen, *Final Report,* p. 61.

[36]Perhaps surprisingly, the (NIT × TR)-Control difference in the dissolution rate for the five-year group of blacks is not statistically significant at conventional levels, even though the NIT × TR rate, .131, is 70 percent higher than the Control rate, .077. The P-value for the two-sided test of significance is .155. The numbers of observations used in these tests of significance are derived from the person-years of record, which are about three times as large as the numbers of couples. Thus, the levels of significance may be overstated. For example, the marital records for 10 couples for one year should convey more information than the record of one couple for 10 years. If this view is correct, the criterion for judging a difference to be statistically significant should be more stringent than usual.

[37]More precisely, the multiplier equals *e* raised to the power of the coefficient. A coefficient of .10, for example, implies that the group's dissolution rate is 1.105 times as large as the control group's dissolution rate ($e^{010} = 1.105$). A coefficient as large as .76, however, implies a multiplier of 2.14, showing a 114 percent increase in the group's dissolution rate compared to the control group. A coefficient of -.12 indicates a multiplier of .887—about a 11 percent reduction in the group's dissolution rate compared to the control group.

[38]Paul D. Allison, "Discrete-time Methods for the Analysis of Event Histories," in S. Leinhardt, ed., *Sociological Methodology,* San Francisco: Jossey-Bass, 1982, pp. 61-98.

[39]In Table 7 Groeneveld, Hannan, and Tuma show four variables for the training programs, but they combine the NIT and (NIT × TR) programs, distinguishing only the three- and five-year durations by an additive three-year dummy variable. Their specification is approximately equivalent to one in which all nine NIT and (NIT × TR) variables in Table 5 are combined, which becomes equivalent to the Groeneveld, Hannan, and Tuma "NIT" variable, and in which all three-year NIT and three-year (NIT × TR) variables in Table 5 are combined, which becomes equivalent to the Groeneveld, Hannan, and Tuma "NIT, 3yr" variable.

[40]At least I believe this is a change from the Groeneveld, Hannan, and Tuma procedure; they state: ". . . we omitted the marital histories of the 20-year families after they were assigned to the 20-year treatment. Their marital histories prior to that time (about 2 years after enrollment) are included. In our analyses all experimental families who become 20-year families are classified as 5-year experimental families until the length of treatments is changed." *Final Report,* p. 287, footnote 1. However, Robins and Steiger, 1980, had claimed that the experimental families who became 20-year families were all originally assigned to the three-year experimental plan.

Discussion

David T. Ellwood

In reading Glen Cain's paper, I was reminded of Harry Truman's expressed desire for a one-handed economist. Cain has done a careful job of discussing all the "one hands" and "other hands" that can contaminate an experiment of this sort when looking at marital dissolution. And he shows us just how unstable the results of the Seattle-Denver income maintenance experiments really are. Yet in reading this paper one is left with the fundamental question: what should we believe about a negative income tax and marital stability? In the end I certainly come away convinced by Cain's assertion that the evidence that a negative income tax is strongly destabilizing is not decisive; but I cannot fully endorse the impression of Cain's last paragraph that the experiments showed neither a practical nor a statistically significant destabilizing effect. Rather I'd say the evidence is just too thin to draw firm conclusions.

Three questions are paramount as we evaluate the possible impact that a negative income tax might have. First, should we have expected the negative income tax to be stabilizing or destabilizing for marriage? Second, what, if anything, do the experimental results show? And finally, how likely is it that the experimental results are a good reflection of what would actually occur if a "permanent" nationwide negative income tax were adopted?

What Should We Have Expected?

I was not a participant or observer during much of the period when

*Associate Professor of Public Policy, John F. Kennedy School of Government, Harvard University.

the negative income tax was being proposed and debated. But my impression from my reading and discussions was that proponents generally expected the program would be stabilizing. Personally, I think the expectation should have been that the plans, at least as implemented in the experiments, would be destabilizing.

An economic model of divorce or separation would suggest that currently married couples compare the net benefits of being married to the net benefits of being apart. Compared to a situation where there are no benefits available to anyone, a negative income tax has an ambiguous impact in theory. It provides added income to both poor intact families and poor separated ones. But in practice, the negative income tax would almost certainly be more destabilizing than doing nothing.

The financial position of intact families is very different from the position of separated ones. Two-parent families are rarely poor, and when they are, their poverty tends to be short-lived. Single-parent families are typically poor, and the poverty often lasts much longer. Thus the expected benefit to single-parent families is far greater than that for two-parent families. Of course a lack of income may be a destabilizing factor in some divorces or separations, but for the most part lack of money is likely to be a far greater problem for the split family than for the intact one. Thus even though a negative income tax appears to be neutral, it is in fact a far greater subsidy to single-parent families than to two-parent families.

Of course the proper comparison is not between the negative income tax and nothing, it is between the negative income tax and the present system. If the effect of the negative income tax was to leave effective benefits for single-parent families unchanged and to increase the economic benefits only to two-parent families, the program ought to be mildly stabilizing. But the bulk of the tested programs offered benefits far more generous than those of the existing AFDC system. Moreover, the program provided far more information on available benefits and options than would generally be known among the general public with respect to the AFDC program. Thus, although these negative income tax programs could have been stabilizing, I think it was reasonable to expect they would have the opposite effect.

What Do the Experiments Show?

Cain is very effective in showing that the results are extraordinarily confusing and unstable. I've spent two days and nights, poring over Cain's detailed numbers looking for patterns, trying to pull out what message there is. In the end, I come away mostly frustrated, unable to say anything but the most equivocal statements.

The reason really is quite simple. The sample sizes are very small and divorce or separation is a relatively rare event. The entire control group of originally intact families in Seattle numbers 263, of which 31 split up and 53 dropped out of the experiment. Breaking these even into racial groups leaves one with almost no sample. Any further breaks leave almost nothing to examine. And attrition seems quite worrisome. As Cain makes quite clear, there are very good reasons to expect far less attrition among experimental families that split apart (since they benefit more from the negative income tax) than among those that remain stable.

What makes matters worse is that the experiments included not only several sites and racial groupings but also enormous variation in the treatments received. Participants received dramatically different levels of benefits. And some experimentals were offered a variety of training and counseling programs in addition to the negative income tax benefits or instead of them. With such thin data, it is almost impossible to disentangle any of the independent effects of one program or another.

Cain argues that we ought to look mostly at the groups that received a "pure" negative income tax with no training or counseling. This is one of the few parts of the paper I found quite unconvincing. Separating the "pure" negative income tax groups from the others thins an already thin sample. Cain finds no evidence that the training and counseling programs alone have much separate effect. And he cannot offer much a priori reasoning as to why we should expect an interactive training/negative income tax impact. The main reason for separately estimating a "pure NIT" effect and an "NIT/training" effect appears to be that the impacts seem larger for the latter group. But with such thin data, surely there are many divisions that would also show highly differential impacts.

I can surely understand why one would want to include separate treatment variables for the negative income tax and training, but I do not see why we should so severely limit our sample sizes in order to allow for an interactive effect of the NIT/training treatment combination. To my knowledge none of the labor supply models employed such a methodology, even though one could argue more directly for a possibly joint effect in that situation. Nor can we say, if a negative income tax plan were actually implemented, whether or not it would be accompanied by a training-like component.

Normally the way we deal with small samples is to pool. But we do so at our peril, of course. These data show little consistency across sites and treatments. Cain's "pure NIT" was stabilizing for blacks in Seattle but strongly destabilizing for them in Denver. Chicanos, on the other hand, were stabilized in Denver. Results for whites were similarly perplexing. As a result the standard errors of all the estimates were extremely high. One clearly cannot infer much from these data. The samples are simply too small and unstable to say anything definitive.

Yet the percentage point estimates are troubling. The overall effect of the negative income tax was to push up family splits among whites by 18 percent, and among blacks close to 50 percent. Even if one looks only at the "pure NIT" as Cain urges, destabilizing effects in the range of 15 to 30 percent are found for whites and blacks. The confusing Chicanos showed a moderately stabilizing pattern. Even though few results are significant, I conclude that there is almost no evidence in these results to suggest the Seattle-Denver income maintenance experiments were stabilizing; however, I do not think we can say just how large the destabilizing effects were. It is worth remembering that significant impacts were not found in other experimental sites.

Would the Results Be the Same for a National Negative Income Tax?

I see no reason to infer much from these results. Cain points out that a short-duration experiment could have a smaller than actual effect because people cannot count on the support indefinitely, or a larger than actual effect if people divorce now while there is an unusually generous basis for support. I strongly favor the latter hypothesis. Divorce or separation is certainly a "threshold" event where some impetus ultimately pushes people into action. It is also an event that may have a short time horizon. Many couples see separations as temporary or exploratory. Moreover, most separated women remarry or reconcile rather quickly. Finally, I doubt many women who go on welfare after a divorce see it as anything more than a temporary bridge. Data on the AFDC program show that formerly married women have the shortest durations on welfare.

The negative income tax may have been seen as a unique moment when a transition into another living arrangement was easier. And the information and attention that experimentals received may have brought the financial options into clearer focus. Some evidence for the proposition that information could have had an impact in and of itself comes from the fact that Groeneveld, Hannan, and Tuma found large destabilizing effects even for plans where benefits were no higher than under current AFDC programs, which pay only for single parents. Theory is unambiguous in suggesting that a program that leaves benefits unchanged for single parents while providing new benefits to two-parent families should be stabilizing. The fact that such plans were destabilizing suggests that information or some other factor contaminated the results. And one possible interpretation of the apparently higher impact of the negative income tax/training counseling combination could be that these programs helped people see how the negative income tax might help them in the short run. (I still do not think this explanation is plausible enough to justify special treatment of the option, however.)

In general, then, I think we learned very little from the negative income tax experiments with respect to divorce and separation. We learned that the experiments did not stabilize families and that they may have been destabilizing. But since I have argued that we should have expected that result anyway, I'm not sure that is very valuable information. I am less skeptical than Cain about the prospects for learning something about these events through experimentation. I think the biggest problem here was that sample sizes were small. But I do believe that these events are inherently difficult to study and are likely to be severely influenced by the experimental design itself, independent of the changed incentives that may be created. Yet social scientists interested in poverty must explore these issues, perplexing and ephemeral as they may seem, for family structure changes and poverty are inextricably and increasingly related.

Discussion

*Nancy Brandon Tuma**

Glen Cain's paper has "two main messages:" (1) "that the evidence [about the effects of the negative income tax experiments on marital stability] is not decisive, or even persuasive;" and (2) "that family issues like marital stability are not well-suited to experimental research." I will comment on each.

Is the Evidence Decisive or Persuasive?

Is the evidence about the effects of negative income tax treatments on marital stability decisive or persuasive? Cain says "no" to both parts of this question. I agree with him that the evidence is not decisive, but I disagree with him about whether it is persuasive.

A decisive result is rare in any experiment, whether it tests a new drug for treating cancer or a new weapons system. At best, most experimental results turn out to be "persuasive" or "suggestive." That is, they alter one's best guess (and hypotheses for the next study), but they are almost never so definitive that a next study is unnecessary.

Our analyses (I refer to those by Groeneveld, Hannan and myself, and especially those described in our final report) convinced me that the negative income tax treatments decreased the marital stability of low-income black and white couples.[1] Cain's reanalyses, which are, in fact, very similar to various analyses included in our final report, have not altered my conclusions. Although a detailed comparison of our 125-page final report and Cain's paper is not possible here, I will summarize what I consider to be the most salient points.

*Professor of Sociology, Stanford University.

Cain presents results for analyses that differ from ours in a number of relatively minor ways, many of which had already been explored in our report and were known (from our reported results) to decrease somewhat the negative income tax's effect on the marital breakup rate of black and white couples. When all of these minor changes are put together, Cain finds a positive but statistically insignificant effect of what he calls a "pure" negative income tax treatment on the marital breakup rate of black and white couples.

The statistical insignificance of Cain's finding is no surprise because the power of hypotheses tests about marital breakup rates using the Seattle-Denver data is low. In order to achieve *statistical* significance, both our analyses and Cain's found that the negative income tax treatments would have to increase the marital breakup rate by roughly 40 percent for black and white couples (and by over 80 percent for the much smaller sample of Chicano couples); hence, increases in the breakup rate that are smaller than 40 percent cannot be statistically distinguished from "no effect," although they may be big enough to be of considerable *social* significance.

Even if one accepts Cain's analytic decisions that act to reduce the negative income tax effects, his "pure" effect (see his table 6) is still positive and large enough to be noteworthy: his estimated increase in the marital breakup rate is 17 percent for whites and 31 percent for blacks. Moreover, the "impure" effect of combined negative income tax-training treatments is as large and positive as the "pure" effect for whites and much larger for blacks. Most people would not ignore the "impure" effect of the combined treatments, especially since the "pure" training effect is tiny for whites and moderate for blacks.

In addition, one may not want to accept Cain's analytic decisions for the following reasons:

(1) Cain omits childless couples because he believes that any negative income tax programs passed by Congress would exclude them from benefits. I contend that our job as social scientists is to analyze *all* of the data. We recognized that the presence of children might affect response, so we did estimate some models with separate effects for couples with and without children. We found (Groeneveld, Hannan, and Tuma 1983, table 5.8, pp. 298-99) that the negative income tax effects for couples with children were smaller than those given in our summary in the case of whites (a 36 percent increase rather than a 53 percent increase) but were about the same for blacks. That is, the negative income tax effect was in the "40 to 60 percent" range (a summary figure from our conclusion on which Cain focuses) for blacks, but a little less for whites.

(2) When Cain analyzes similar data using the same explanatory variables with a similar model, his estimates (table 7) are only about

80 percent as large as ours. (This applies to effects of nonexperimental variables as well as experimental treatments.) I suspect that his estimates are smaller than ours because he aggregates the data based on the date of the dissolution. (Cain aggregates to six-month intervals; we recorded events to the nearest day.) Although Monte Carlo studies are needed to say for sure, time aggregation probably biases estimates downward. Cain's decision certainly has no known scientific advantages.

(3) Due to the small sample size relative to the number of treatment and assignment variables, analysts of these data cannot cross-classify by all treatments and assignment variables. Cain chose to ignore one set of cross-classifications; we chose another. Naturally, results depend on these choices. Whose choice is better? Two differences in our choices stand out:

(a) Cain stresses a model that includes an interaction between training (actually, a mixture of three quite different treatments) and the negative income tax treatment (a grouping of 11 different financial plans). Like Cain, we estimated a model that interacted the negative income tax treatment with the training treatments, but we separated the three training treatments, which we regarded as quite different. We found that the set of interactions was not significant for blacks but was significant for whites (Groeneveld, Hannan, and Tuma 1983, table 5.B.2, pp. 371-72). We were skeptical about these results, however, because the pattern of effects for various treatments was unsystematic. We thought that the data were cross-classified so much that chance variations due to small cell sizes swamped any trends. Omitting the negative income tax-training interactions gives what we consider to be a clearer view of the overall effects on the negative income tax treatments.

(b) Cain handles plan length (length of treatment) differently than we did. In our view, having a five-year plan rather than a three-year plan is analogous to giving a drug to cancer patients in two strengths, the first more potent than the second. In this parallel situation, analysts do not regard the two treatments as entirely unrelated. Rather, they test whether the effects of the two doses differ. If patients given the stronger dose respond to the drug significantly, and patients given the weaker dose have a similar but smaller and insignificant response (essentially what we found), most analysts conclude that the drug *does* have an effect, but that one dose was too weak for its effect to be detected with the data available. This reasoning led us to stress the effects of the five-year plan.[2]

Cain's approach is quite different. In his table 5 he interacts plan length with site and his three treatment components: training, negative income tax, and negative income tax-training. I do not see any scientific reason for his approach here, but I would predict that spreading the treatment effects across 13 treatment variables is extremely unlikely to yield systematic or significant effects. In table 6 he omits not only the site and plan length interactions, but also a main effect for plan length. This is like treating weak and strong doses of a drug as equivalent. I do not see any scientific grounds for this.

Finally, there is a piece of evidence from our final report that Cain does not mention and that helps convince me that the negative income tax treatments did increase marital breakup rates. Namely, we also analyzed pooled data from the Seattle, Denver, and New Jersey experiments, which increases the overall sample size substantially. The larger sample increases the power of tests and greatly reduces the standard errors of estimated effects. These analyses gave estimates of significant, 25 to 35 percent increases in the marital breakup rates of white, black, and Hispanic couples (Groeneveld, Hannan, and Tuma 1983, table 5.11, p. 303). Further study is needed because we did find some important variations with site.[3] Still, the evidence from the pooled experimental data is persuasive that the negative income tax treatments tended to have *some* positive effect on marital breakup rates of low-income couples in diverse settings.

In comparing Cain's analyses and ours above, I have stressed differences that in principle can be evaluated objectively. Another difference may arise from our disciplinary perspectives. As an economist, Cain stresses monetary differences between the negative income tax treatments and welfare programs like aid to families with dependent children (AFDC). As sociologists, we consider nonmonetary as well as monetary differences in these programs. Much of our final report was devoted to analyses that tried to understand why a negative income tax program that was financially similar to AFDC increased marital breakup rates. Indeed, we thought this "message" was as important as numerical estimates of an overall negative income tax effect on marital stability, which is what Cain emphasizes. Since Cain "assumes away" this part of our message, I will restate it.

We argued that administration of AFDC and of the experimental negative income tax programs differed in several key ways that could cause differential response to the same monetary benefits. (1) Knowledge of benefits and rules is likely to be lower for AFDC than for the negative income tax programs, which were carefully explained initially and again a year later. (2) The costs in time, effort, and social embarrassment of getting benefits is greater with AFDC than with the negative income tax treatments. The latter, for example, required only a monthly mailed report of income and family composition, and the same report was to be sent whether or not a breakup occurred. (3) Promptness and the short-run certainty of receiving benefits after a breakup were greater in the negative income tax program than with AFDC, again because no special action on the part of a recipient was required after a breakup in order for benefits to begin or to be increased (if the couple was already receiving benefits). Unfortunately, with the data from the Seattle-Denver income maintenance experiment, one cannot assess the *relative* importance of these differences between AFDC and the experimental negative income tax programs. But we think that together they account for the relatively large increase in marital breakup rates under negative income tax treatments financially similar to AFDC.

Three final points about nonmonetary differences between AFDC and negative income tax programs deserve mention. First, although the administration of the *experimental* negative income tax programs made it easy for people to benefit from them, a *federal* program might not be administered in a similar way. Second, one could experimentally vary administrative features of a negative income tax program and study the consequences. Third, *if* administrative features are important, as we argued, and *if* the administrative features of a federal negative income tax program are different from those in the Seattle-Denver experiment, then neither our numerical estimates of negative income tax effects nor Cain's are a good basis for estimating the costs of a proposed federal program.

Are Family Issues Suited to Experimental Research?

What about Cain's other message? Should family issues be studied experimentally? Cain says "no," primarily, it appears, because he believes the cost of a well-designed experimental study would be "too high." Deciding if one agrees with Cain requires a cost-benefit analysis involving answers to three questions:

(1) What would a well-designed experimental study of family issues look like? What would it cost? No one has yet tried to *design* such a study, let alone estimate its *cost*. Thus, a very basic piece of evidence for Cain's view is missing.

(2) What would it cost to obtain the same information by other means? The most likely other source of such information would be analyses of nonexperimental data, for example, panel surveys. Not only would a good nonexperimental study of family issues be costly, but it quite possibly might be more costly than a well-designed experiment.[4]

(3) How valuable is knowledge about the relationship between social policies and family issues? Whatever the cost of a study of family issues, some people may think it is "too high" simply because they don't value the information it produces.

Since Cain has not yet given serious answers to (1) or (2), let alone said how much he thinks such information is worth to our society, I am totally unconvinced by his claim that the costs of a properly designed experimental study of family issues would be "too high."

I am convinced, however, that someone needs to think hard about how to design good experimental and nonexperimental studies of family issues, so that debate about the value of such studies can move from the level of rough and ready speculation to one with a sound scientific basis. And, while I am persuaded that the experimental negative income tax programs tended to decrease marital stability of low-income couples, I also think estimates of the magnitude of these effects (both ours and Cain's) are not sufficiently precise for policy planning. Moreover, the negative income tax experiments definitely did not give adequate information on the role of *nonmonetary features* of the treatments. If there is another set of negative income tax experiments someday, I hope that they will be designed not only to obtain more precise estimates of effects of plan generosity but also to study this important issue.

---

[1]Like Cain, we concluded that the negative income tax treatments did not alter the marital breakup rate of Chicanos couples. However, we did find that they markedly decreased the marital *formation* rate of unmarried Chicana women with children, and this is an alternative way of decreasing marital stability. So, while I think the evidence shows that the negative income tax treatments decreased the marital stability of Chicanos, this cannot be detected from analyses of data on marital breakups from the Seattle-Denver income maintenance experiment. See Groenveld, Hannan, and Tuma 1983.

[2]Cain suggests that his results may differ from ours partly due to different handling of those assigned to the 20-year plan. As far as I can tell from his paper, Cain treated them exactly the same as we did. In any case, I'm skeptical that somewhat different handling of fewer than 10 percent of the sample would cause appreciable differences, especially since our analyses focused on the first 36 months of data and the 20-year plan only began after about 30 months.

[3]The effects of the negative income tax treatments on marital breakup rates were somewhat smaller for whites in New Jersey than for whites in Seattle and Denver. Contrarily, the effects were much larger for Hispanics in New Jersey (mainly Puerto Ricans) than for Hispanics in Denver (mainly Chicanos). The variation with site could arise because of cultural, ethnic, or religious differences in the populations in the three sites. They could also be partly due to differences in state programs of aid to families with dependent children; these differences cause control group comparisons to differ even if the negative income tax plans are the same. However, the negative income tax treatments in New Jersey also differed in a number of ways from those in Seattle and Denver, so this is yet another possible reason for differences across sites. Still other reasons for site differences can be suggested.

[4]Since available nonexperimental data on family issues and income are still very inadequate, a good nonexperimental study would almost certainly involve costs of data collection as well as analysis. And, since the costs of the negative income tax experiments came disproportionately from data collection and analyses—not from administration of treatments (see Zellner and Rossi 1986)—a nonexperimental study might not cost much less than an experimental study with the same number of cases. Moreover, sample sizes must usually be much larger in a nonexperimental study than in an experimental study, in order to estimate effects with equal precision. As a result, a good nonexperimental study of family issues could be more costly than a well-designed experiment. This ignores likely biases in nonexperimental studies, which are even harder to handle than the two sources of bias that Cain associates with an experiment.

References

Cain, Glen G. ''The Issues of Marital Stability and Family Composition and the Income Maintenance Experiments,'' 1986, this volume.

Groeneveld, Lyle P., Michael T. Hannan, and Nancy Brandon Tuma. ''Marital Stability'' in *Final Report of the Seattle-Denver Income Maintenance Experiment, Volume 1: Design and Results,* Washington, DC: U.S. Government Printing Office, 1983.

Zellner, Arnold and Peter E. Rossi. ''Evaluating the Methodology of Social Experiments,'' 1986, this volume.