Economic Response to a Guaranteed Annual Income: Experience from Canada and the United States

Derek Hum and Wayne Simpson,
University of Manitoba

This article reviews research from the five income-maintenance experiments in Canada and the United States. After sketching the historical and political context of the experiments, we compare their designs and discuss some important analytic difficulties. Our primary focus is the work-incentive issue, both nonstructural estimates of the experimental effects and elasticity estimates of structural labor-supply functions. We provide initial estimates of nonstructural and structural models for the Canadian experiment. We discuss more briefly some non-work-response findings associated with a guaranteed annual income and offer some personal comments on social experimentation and the policy process.

I. Introduction

The United States and Canada are among the most prosperous industrial countries of the world. But both countries also experience significant poverty amid plenty. In both countries, there is mounting concern that cash transfers by government to the needy do little to diminish their numbers, nor do they lessen their dependence on welfare or encourage their transition to full-time employment. A consensus has emerged that the welfare system in each country is, in its own peculiar way, uncoordinated and inefficient. The past 2 decades were witness to much creative thinking and social experimentation in both countries.

We would like to thank Walter Block for his encouragement and direction in developing this article. We accept responsibility for the interpretation of research findings and any remaining errors.

[*Journal of Labor Economics*, 1993, vol. 11, no. 1, pt. 2]

0734-306X/93/1101-0019$01.50

Although Canadians and Americans highly value what they sometimes regard as a special relationship between their two countries, it does not mean that their histories, constitutional structure, attitudes, or policy approaches are identical. Indeed, given their similarity in so many other economic arrangements, the differences, particularly in social and cultural matters, are all the more subtle. These differences have recently been given center stage as a result of the Canada–United States Free Trade Agreement, particularly by Canada, which, being much the smaller partner, is especially worried that the agreement will, directly and indirectly, compel the two countries to align their social programs (Hum 1988*a*). There has been little convergence of income security programs in the past (Kesselman 1990). This feature stands out only because of the striking record of Canadian imitation of the United States in many other areas of economic policy—including personal and corporate taxation, deregulation, and economic stabilization. But whatever their common or separate roads in the past, it remains certain that both countries will continue to borrow ideas from each other. In the past 2 decades, no idea has passed more freely across the Canada–United States border than the guaranteed annual income or negative income tax.

"Income maintenance" is a broad category that can include any income-conditioned benefits, either in kind or in cash. A special form of income maintenance is the negative income tax (NIT), which provides a maximum cash benefit (G) to families with no other income and reduces the payment amount by a specific "tax" or "benefit-reduction" rate (t) for each dollar of other income received by the family. Since the family can never receive less than the amount G, this is tantamount to guaranteeing the family a minimum income transfer, hence the term "guaranteed annual income" (GAI). The NIT and GAI terms are often used interchangeably to refer to any income maintenance plan that delivers income-conditioned cash benefits by formula, usually through the tax-transfer system rather than through traditional welfare, which is based on discretionary determination of need.

The language undoubtedly reflects economist Milton Friedman's (1963) original suggestion for a negative income tax in which he proposed that a portion of the unused tax exemptions and deductions allowable under the personal income tax (PIT) be actually paid to individuals by government. Meantime, in Canada, a proposal for an unconditional income allowance, complete with cost calculations and administrative procedures, was detailed by a chartered accountant (Smith 1965). There have since been many variations on the theme of GAI or NIT, including some not so obvious. For example, it is easily demonstrated that a NIT is not significantly different from the many provincial and federal refundable income-tested tax credit programs now in place in Canada (Hum 1988*b*), and it is commonly acknowledged that Canada's Old Age Security (OAS) demogrant (transfer program) in combination with the Canada Pension Plan (CPP/QPP) and Guaranteed Income Supplement (GIS) programs effectively constitute a guaranteed annual income for the elderly.

The notion of a GAI or NIT was a radical suggestion in the early sixties. Little was known about its probable economic effects, particularly how a GAI might affect individual labor market behavior. Accordingly, both the United States and Canada embarked on a series of social experiments during the seventies in order to explore the economic and social consequences of a guaranteed income program. The experiments cost millions of dollars, lasted many years, and tested the negative income tax method of delivering transfers. At the heart of the experiments' objective in every case was the question of work incentives and, by extension, concern over the eventual cost of a nationwide GAI. It is quite clear that the more generous the GAI or NIT program, as specified by high support levels (G) and low tax rates (t), the larger will be the program costs to government. This results because nonworkers would receive larger payments, low-income workers would keep a larger fraction of their earnings, and a larger proportion of the population would be recipients since high guarantees and low tax rates have the effect of increasing the income eligibility threshold of the program. Consequently, attempting to eliminate poverty through a GAI can be costly, depending on the support level and tax rate chosen. Similarly, whether or not a GAI would induce able-bodied individuals to work less was an obvious policy concern, and this question required careful empirical research and precise measurement to provide the necessary answers.

All of the experiments on guaranteed income are now complete, and research findings are available on a variety of behavioral responses. The data and findings continue to produce new insights to this day. Because of its central focus, policy importance, and continuing controversy, we shall focus on the work incentives or labor-supply response to a GAI, although we shall also consider briefly other aspects of response. Needless to emphasize, all of the experiments were elaborately designed, produced massive amounts of information, and resulted (and continue to result) in enormous research effort. Furthermore, because each experiment differed with respect to design, sample composition, payment delivery method, and statistical methodology, the results of the different experiments are not directly comparable. In the case of work incentives, however, comparability across experiments and with nonexperimental evidence can be achieved by analyzing structural labor-supply models.

The policy debates and political history surrounding the GAI are too voluminous to survey. We therefore confine our attention to the social experiments in guaranteed income, not merely because of their significant symbolic nature, but also because of their unarguable importance as research projects breaking with past precedent as well as their spirit in encapsulating an entire generation of thinking about income maintenance possibilities and social reform.

The next section sketches the historical and political context of the origins of the income maintenance experiments in both the United States and Canada. This is followed by a brief outline of the designs of the various experiments, featuring their main similarities and unique aspects and indicating what the experiments hoped to learn respecting the question of work behavior responses. Section IV discusses some selected analytic difficulties associated with the analysis of experimental panel data of the sort produced by the experiments and what problems and benefits were encountered with these unique economic data sets. We then summarize and discuss the work-incentive issue, both in terms of nonstructural estimates of the experimental effects and in terms of the wage and income elasticity estimates of structural labor-supply functions. We present some initial estimates of nonstructural and structural models for the Canadian experiment, the details of which are placed in the Appendix to this article. We discuss more briefly some of the non-work-response findings associated with a GAI and offer some personal comments on social experimentation and the policy process in our conclusion.

II. Historical Background to the Experiments[1]

A. Origins of the American Experiments

President Lyndon Johnson called for a war on poverty in his state of the union address in 1964. In that same year the U.S. Congress established the Office of Economic Opportunity (OEO) as the vanguard for the antipoverty effort. The programs discussed by the OEO comprised three aspects, namely, public employment strategies, community action programs, and income maintenance. Although the role of income maintenance to combat poverty was readily accepted, a guaranteed income or negative income tax approach to delivering cash transfers proved more controversial.

The negative income tax idea initially met both opposition and neglect. Some, like Alvin Schorr, deputy director of research for OEO, favored an alternative proposal based on children's allowances payable to all families with children regardless of income.[2] In contrast, Joseph Kershaw, director of the research office of OEO, stressed the distributional efficacy of income-conditioned payments and recommended the NIT proposal to Sargent

[1] This sketch of the GAI experiments is intended for interest and continuity only. It does not pretend to be a complete critical political history of the experiments. Description of the U.S. experiments is taken from Basilevsky and Hum (1984). The Canadian chronology is adapted from Hum and Simpson (1991).

[2] Canada has had a system of universal family allowances since 1945.

Shriver, director of the OEO. Shriver was won over by the strong advocacy of the research group, and the antipoverty plan that was submitted to the White House in September 1965 contained the NIT as a component. The OEO also forwarded in October 1965 to the Bureau of the Budget a NIT proposal costing $4.7 billion as the centerpiece of its antipoverty plan. The White House, however, was preoccupied with the Vietnam War and the falling popularity of some of the OEO's social programs; it did not take the NIT proposal seriously (Levine 1975), and the only response from the president was to appoint a commission on income-maintenance programs (Lampman 1974).

Despite the lack of political enthusiasm, the negative income tax idea did not die. Partly because of OEO's unwavering faith and support, and partly because of the continuing war on poverty, the negative income tax was regarded by its proponents as an idea whose time had come. But besides OEO's support, additional factors contributed to the eventual series of negative tax experiments. The OEO continued to single out the NIT for attention as part of its mandate concerning antipoverty strategies. Additionally, the research staff and OEO bureaucrats were very heavily influenced by what Lampman (1974) has called the "ascending discipline of the Program Planning Budget System" (PPBS). Prominent within the OEO were key individuals—many recruited from RAND or the Pentagon, new to social welfare, and without sharply defined loyalties to specific agencies or proposals. These individuals accepted the application of evaluation techniques. Accordingly, the goal of eliminating poverty was stated in income-maintenance terms, alternative proposals were arrayed, and cost-effectiveness scores were assigned to different schemes on the basis of the "most bang for a billion bucks." Under this exercise the negative income tax received high marks and consequently had the effect of focusing further discussion on particular aspects of the NIT approach, such as the cost sensitivity and work-disincentive effect of guarantee amounts and tax rates. The effect of general cash transfer mechanisms on the work effort of the non-aged, able-bodied individuals therefore emerged as the (now-clarified) prime empirical issue.

Many critics of the negative income tax felt it would cost more than existing welfare programs since its objective is to extend cash payments to the working poor—a group ineligible for most other programs.[3] In contrast, the proponents of the NIT saw the major stumbling block as political. The belief, by politicians as well as the general public, that a NIT would promote idleness among the able-bodied poor was strongly held, and no amount of argument "without hard facts" was likely to dispel such beliefs. This then became the dominant issue—pushing all other disagreements concerning the cost of the NIT, the administrative practicality and

[3] This is generally true also for Canada.

mechanics of the scheme, the lack or otherwise of stigmatizing effects, and other issues into the background.

Viewing the central problem of the NIT in terms of the work disincentive effectively transformed the issue into one for which economists could claim special competence. In the jargon of economics, the NIT was restated as a controversy concerning wage rate (price) and income elasticities pertinent to labor-leisure choice. Economic theory provided a conveniently coherent model, and economists themselves readily demonstrated that existing data sources could not answer the incentives issue with confidence. Indeed, the early research produced wide differences in the estimates of labor-supply response, particularly in wage elasticities for married women (Cain and Watts 1973) that have endured.[4] However, the necessary information and evidence could be gained with an experiment. The proposition seemed so simple. Why not try it out? Conduct an experiment!

Credit for the initial idea goes to Heather Ross, an economics graduate student at the Massachusetts Institute of Technology who was working with the Council of Economic Advisers during the summer of 1965. Although Ross's specific proposal was not accepted, it received wide circulation within the OEO, according to Levine (1975, p. 17), and many econometricians strongly endorsed the idea of an experiment (Orcutt and Orcutt 1968). Proposal for an experiment received strong support from OEO, which initiated serious planning on the design for an experiment in 1966. The final proposals were endorsed by the OEO research staff as well, and Sargent Shriver added his approval in 1967. Shriver was able to counteract political opposition, and by the fall of next year families had been selected for enrollment in a negative income tax experiment, payments were being made, and the first of the large-scale social experiments in North America—the New Jersey Graduated Work Incentive Experiment—had begun. The undertaking was not called a negative income tax experiment but instead, for political purposes, a "work-incentive experiment," connoting a happy rather than unhappy anticipated outcome. As well, the experiment now emphasized the purely scientific dimensions of the project—as evidenced by the (deliberate) funding of the experiment through the Institute for Research on Poverty in Wisconsin.

The first income-maintenance experiment in the United States was therefore forged out of sharply different motives and interests. Undoubtedly, the antipoverty program was important in setting the climate for political and policy debate. Equally, the cost-effectiveness apparatus of the PPBS and the strong advocacy of OEO's research staff for the NIT were also ingredients. As well, academic econometricians eager to extend social science into the realm of controlled experimentation played an influential

[4] Reviews of the early, nonexperimental evidence include Keeley (1981); Killingsworth (1983); and Hum and Simpson (1991).

role (Lampman 1974, foreward). Nonetheless, it remains that no single statement can fully capture the subtleties of how and why the New Jersey experiment came to be. Nor did the matter end with the birth of an experiment, as Haveman and Watts (1976) observed: "[The] tension between the motivations of those who supported the experiment for 'general-political-demonstration' reasons and those who desired it for 'technical-economic-experimental' reasons persisted throughout the [New Jersey] experiment. It affected all of its primary characteristics from technical design to duration to selection of sites and finally to interpretation of results" (p. 427).

Other income-maintenance experiments in the United States rapidly followed. The OEO awarded a further grant to the Institute for Research on Poverty for a negative tax experiment in rural areas. The Department of Health, Education and Welfare (HEW) also funded one in Gary, Indiana, and others in Seattle, Washington, and Denver, Colorado. Each of these other experiments had a slightly different focus and often incorporated additional research objectives, but the New Jersey experiment remains distinctive in setting the precedent for the series of carefully controlled, scientific field tests of different benefit formulas on work behavior.

B. Origins of the Canadian Experiment

The discussions concerning the American war on poverty and the various proposals that evolved as part of its antipoverty strategy did not go unnoticed in Canada. The Canada Assistance plan (CAP) came into effect in 1967 and was to be the centerpiece of Canada's antipoverty efforts.[5] At about the time the first families were being enrolled in the New Jersey experiment in the summer of 1968, the Economic Council of Canada released its *Fifth Annual Review* (Economic Council of Canada 1968), telling Canadians about the extent of poverty in Canada. In November 1970, the Department of National Health and Welfare issued a white paper that emphasized the potential of a guaranteed income as an antipoverty measure but also worried about the disincentive economic effects. The white paper declared (p. 41): "An overall guaranteed income program . . . worthy of consideration [must] offer a substantial level of benefit to people who are normally in the labour market. Therefore, a great deal of further study and investigation, like the experiments now under way in New Jersey and Seattle in the United States, is needed to find out what effects such a program would have on people's motivation, on their incentives to work and save. Until these questions are answered, the fear of its impact on productivity will be the main deterrent to the introduction of a general overall guaranteed income plan" (Canada 1970).

[5] The CAP Act remains the most significant umbrella program for cost sharing of social assistance. For a detailed discussion of CAP, see Hum (1983).

The next year, 1971, saw the publication of the Croll Report (Canada 1971), which recommended that a GAI based on the NIT be implemented on a uniform, national basis and financed and administered by the government of Canada. The Castonguay-Nepveu Report (Quebec 1971) also appeared in 1971 and suggested an innovative two-part guaranteed-income program: one plan with a high support level and high tax rate for those unable to work, and a second plan with a lower support level and a lower tax rate for those with a significant attachment to the labor force. However, the impetus for experimentation and reform came from another quarter—federal provincial relations and the Constitution.

In 1971, a federal-provincial conference was held in Victoria in an attempt to rewrite and "patriate" the Canadian Constitution. The provinces and Canada appeared to reach agreement when Quebec declared that it could not support the "Victoria Charter" because, in part, it "failed to provide for a jurisdictional settlement in the field of social policy" and "no patriation of the Constitution would be possible until those concerns were satisfied" (Van Loon 1979). There was much discontent in federal-provincial relations after the Victoria Conference, and this surfaced in 1972 at the Conference of Provincial Welfare Ministers. Federal disappointment over the failure to patriate the Constitution (including an amending formula) was deep. Provincial dissatisfaction was fueled by the federal government's unilateral changes to unemployment insurance in 1971 and its proposed reform of family allowances. There was also resentment over federal intrusion into provincial jurisdiction with what provinces felt were ill-conceived and uncoordinated programs. Thus when the Conference of Provincial Welfare Ministers unanimously called for a joint review "to develop better mechanisms for achieving a rationalized social security system in Canada," the federal government quickly agreed (Johnson 1975, p. 457).

During 1971, Manitoba indicated serious interest in testing the guaranteed income approach, particularly as a demonstration project or administrative test. On June 4, 1974, Canada and Manitoba signed an Agreement Concerning a Basic Annual Income Experiment Project covering cost-sharing arrangements and the respective roles of the two governments. The agreement came about a year after the release of the orange paper (Canada 1973) and the start of the joint Federal-Provincial Review of Social Security.

The social security review and the Manitoba Basic Annual Income Experiment (dubbed "Mincome") were plainly linked in purpose and timing. The review is variously regarded as an attempt to supplant certain portions of the CAP legislation (Communique 1975) or even as a surrogate for constitutional discussions adjourned at Victoria (Van Loon 1979). The National Council of Welfare (1976, p. 1) bluntly asserted that "the goal of the social security review [was] the establishment of a guaranteed annual income." Similarly, Mincome was more than just an expensive exercise in econometrics. The joint news release (Department of National Health and

Welfare/Department of Health and Social Development 1974, p. 5) announcing the final approval of the experiment by Canada and Manitoba was quite clear about the role and purpose of the guaranteed income test. It proclaimed: "The Manitoba experiment is expected to make an important contribution to the review of Canada's social security system launched last April by all ten provinces and the federal government."

Manitoba's original support for the GAI was grounded in its strong interest in administrative and operational issues. Premier Edward Schreyer of Manitoba viewed the GAI as essentially involving "income-testing," and held that it "didn't differ at all [from] a negative income tax" (*Winnipeg Tribune* 1971). Furthermore, because the GAI "would . . . substitute for the Canada Assistance Plan Program," the Mincome "project would be established under the aegis of the Canada Assistance Plan." Premier Schreyer saw Manitoba's financial involvement at "something over $500,000" and the number of families involved "possibly 500" but "closer to 300." What emerged, however, was not the simple demonstration involving "300 families" and a half million dollars that Manitoba wanted but an extremely complicated experiment, modeled along the lines of the American efforts and concentrating on the issue of work response.

Unlike the American efforts, however, which all eventually released final reports and findings, the Canadian project languished. The project published no official findings concerning the labor market response of participants, and the vast amounts of data collected remain archived. The Mincome experiment died a quiet death in 1979, officially reported as a redirection of experimental objectives. It must be remembered that Mincome's official demise came toward the end of the seventies. The social security review had ended by then; there was no political support in the country for sweeping reforms of the type promised by a guaranteed income. The GAI concept itself had lost its allure.

During the next years the fate of the data set itself appeared uncertain. The manner in which the data were archived (unpublicized location, unknown means of access, etc.) was discouraging for the research community. Only recently has analysis of the Canadian experiment emerged from individual Canadian academics. Not surprisingly, discussion in Canada to this day concerning the effect of a guaranteed income on work behavior still relies heavily on American results.

III. Experimental Design and Expected Behavioral Response

The income-maintenance experiments remain the focus of research on the GAI, particularly concerning such economic issues as work incentives. While these experiments have many common elements, they also have unique features. As a prelude to our discussion of the experimental results in the next section, we review the design of the five income-maintenance experiments and the anticipated labor-supply response.

All of the experiments were designed to estimate the response of families to a permanent GAI program that would provide income maintenance payments, P, based on household income, Y, according to the formula

$$P = G - tY > 0 \quad \text{if } y < B = \frac{G}{t}, \tag{1}$$

where $Y = wh + y$, and where h is hours worked, w is the hourly wage, and y is other household income.[6] The level of support and program response will depend on the assigned guarantee level, G, and the tax (or benefit-reduction) rate, t.[7] Thus, to investigate program response, the experiments offered a variety of plans (combinations of G and t) to selected individuals, including a control group that remained on the existing welfare program but was monitored in the same fashion as those receiving GAI payments (the treatment group). In principle, this design offers a simple, direct comparison of the effects of a shift to alternative GAI plans.

In particular, the experiments sought to measure the labor-supply response or, most simply, the change in hours worked, h, caused by experimental intervention. A response was expected on the basis of conventional consumer theory, which may be summarized in terms of the static labor-supply model:

$$h = f[w, y] \geq 0. \tag{2}$$

Taking the total differential of equation (2) gives labor-supply response in terms of hours worked:

$$dh = \frac{\partial h}{\partial w} dw + \frac{\partial h}{\partial y} dy, \tag{3}$$

which depends on the change in the wage rate of each household member and the change in unearned income resulting from experimental intervention. This response may be rewritten to decompose the gross, or uncompensated, wage effect into a compensated wage (or substitution) effect and an income effect as follows:

[6] Other household income may include other earned income. Following other analyses of the experiments, we have ignored any cross-wage effects in household labor supply.

[7] For simplicity of exposition we ignore such complexities as the taxation of net worth. That is, we assume that the net worth of the family is zero in this example. As well, the administrative regulations of the payment delivery system differed markedly among experiments, and these details are also ignored here.

$$\frac{\partial h}{\partial w} = \frac{\partial f^s}{\partial w} + h\frac{\partial f}{\partial y}, \tag{4}$$

where $\partial f^s/\partial w > 0$ is the substitution effect and $h\partial f/\partial y < 0$ is the income effect. Substituting (4) into (3) gives

$$dh = \frac{\partial f^s}{\partial w}dw + [hdw + dy]\frac{\partial f}{\partial y}. \tag{5}$$

For a GAI plan that reduces after-tax wages, $dw < 0$, and that increases the unearned income guarantee, $dy > 0$, the signs in equation (4) imply that labor supply must fall, $dh < 0$. This constitutes the conventional textbook wisdom about the work incentive consequences of a GAI.[8] By introducing substantial variation in after-tax wage rates and income guarantees, the experiments thus became strategies to test this wisdom and to measure the magnitude of the labor-supply effects for different plans.

The guaranteed income experiments have many common features that facilitate comparison of the experimental designs and results. The common design elements can be traced to the original deliberations concerning the New Jersey Graduated Work Incentive Experiment. For example, Conlisk and Watts (1969) developed a sample design and assignment model for the New Jersey Experiment, which was used in all subsequent experiments. This model provides a formal technique, adapted from the classical experimental design literature, to allocate sample points, trading off the research benefits (in terms of a reduced variance of estimated response from a specified response function) against substantial transfer dollar costs per family selected to satisfy an overall budget constraint. Given a response function and total budget, the assignment model produces the sample allocation that yields the least prediction error. The sample assignment can then be adjusted for such considerations as anticipated attrition from the experiment and minimum cell size requirements. While indeed optimal in many respects, the assignment model in retrospect poses some problems for the analysis of response, particularly nonrandom assignment, which will be considered below.[9]

[8] As Levy (1979) points out, the argument should distinguish those already working from nonworkers, who will have weaker income effects, and new participants in a GAI arising from a higher breakeven income level. Moreover, for social assistance recipients facing tax rates very close to 100%, after-tax wages may rise as a result of a GAI for many program participants and provide a work incentive rather than a disincentive (Hum and Simpson 1991). Thus the textbook case is highly simplified and possibly misleading.

[9] See Keeley and Robins (1978); and Basilevsky and Hum (1984, chap. 3) for a critique of the Conlisk-Watts assignment model.

Table 1 presents many of the crucial design features of the five experiments, from which further similarities and some distinct features can be discerned. Each experiment concentrated on household units with low incomes. The New Jersey, Rural, Seattle-Denver, and Mincome experiments used similar income cutoffs (about 150% of the official poverty line), while the Gary experiment admitted households with incomes up to 240% of the poverty line and beyond. Each experiment included a number of plans, defined in terms of guarantee levels and tax rates, but Seattle-Denver included a declining tax rate, counseling, and training subsidy plans. The Gary experiment included social services counseling and day-care subsidy plans, and Mincome included a saturation site offering one plan to the entire community of Dauphin, Manitoba. The duration of the experiments was 3 years, but Seattle-Denver enrolled some households in 5- and 20-year plans to investigate the effect of experimental duration.

Seattle-Denver was by far the largest experiment with 4,800 participating families, almost as many as the other four experiments combined. It was also the most ambitious in terms of the variety of plans tested.

Moreover, each of the experimental sites provided a look at low-income households in a different setting and a different area of North America: New Jersey concentrated on inner-city households in an older industrial area; the Rural experiment looked at areas of widespread rural poverty (North Carolina) as well as poverty amid rural affluence (Iowa); Seattle-Denver looked at one West Coast city with considerable employment instability (Seattle because of its dependence on the aerospace industry) and another with greater employment stability (Denver); Gary examined black ghetto households, and particularly female-headed black households; and Mincome looked at low-income households on the Canadian prairies in both an urban (Winnipeg) and rural (the Dauphin saturation site) setting.[10]

Each of the experiments has now been examined in isolation, and various research reports have been issued. In all cases, individual researchers were permitted to develop and test their own models, select their own sample subset, choose their measure of response, and interpret their results. Yet the experiments also represent a series of closely related trials with many common design features that can inform us about behavioral response to a guaranteed annual income plan in the North American population. Although each experiment has unique features in terms of site selection, target population, and plan design, the common features can tell us a great deal about various aspects of response, the most important of which remains the topic of work incentives. We now turn to the analysis of labor-supply response in the five experiments.

[10] For further information on the sample design and assignment model in Mincome, see Hum, Laub, Metcalf, and Sabourin (1979).

Table 1
Summary of the Features of the Income Maintenance Experiments

	Plan				
	New Jersey	Rural (RIME)	Seattle-Denver	Gary	Mincome, Manitoba
Site(s)	Trenton, Patterson-Passaic, and Jersey City, N.J.; Scranton, Pa.	Duplin County, N.C.; Pocahontas and Calhoun Counties, Iowa	Seattle, Wash., Denver, Colo.	Gary, Ind.	Winnipeg and Dauphin, Manitoba
Eligibility	Intact households headed by able-bodied males 18–58 with at least one dependent and incomes < 150% of poverty line	Families with at least one dependent and incomes < 150% of poverty line	Families with at least one dependent and incomes < $11,000 (single-headed) or $13,000 (double-headed)	Black households, head 18–58 with at least one dependent and income < 240% of poverty line	Families with able-bodied heads under 58-years-old, incomes < $13,000 (family of four)
Sample size	1,357 households; 725 experimentals, 632 controls	809 families: 587 non-aged male-headed, 108 non-aged female-headed, 114 older heads	4,801 families (Denver 2,758, Seattle 2,043)	1,800 black households, 60% female-headed (125 households added with incomes above 240% of poverty line)	1,300 families and single individuals
Plans [not all t, G combinations included in each experiment; more generous plans (high G, low t) typically excluded]	8 plans; t = .3, .5, .7; G = .5, .75, 1.0, 1.25 of poverty line ($5,000 for family of 4)	8 plans; t = .3, .5, .7; G = .5, .75, 1.0 of poverty line	11 plans; t = .5, .7, .7*, .8* (* indicates tax rate declines per .025 per $100 income); G = .95, 1.2, 1.4 of poverty line; training counseling, training subsidies (50%, 100%)	4 plans; t = .4, .6; G = .75, 1.0 of poverty line; social services counseling, day care subsidies (35%, 60%, 80%)	Winnipeg: 7 plans; t = .35, .5, .75; G = $3,800, 4,800, 5,800 (family of four in 1975) Dauphin: 1 plan (saturated site); t = .5; G = $3,800
Duration/start up date	3 years/1968–69	3 years/1970	3, 5 years, 20 years (Denver only)/1969	3 years/1971	3 years/1975

SOURCES.—Pechman and Timpane (1975); Ferber and Hirsch (1978); Keeley (1981, chap. 5); Basilevsky and Hum (1984).
NOTE.—t refers to the experimental tax rate; G refers to the experimental income guarantee rate.

IV. Labor-Supply Response to a Guaranteed Income: Experimental Analysis and Response

The five experiments were designed to provide reliable and credible analysis of the response to a guaranteed annual income program to inform policy development. In particular, there was considerable controversy over prospective labor-supply response on the basis of nonexperimental evidence and political perception, as discussed in Section II. Data from the experiments were expected to resolve this controversy. In this section, we consider the analysis of labor-supply response in the experiments, the problems encountered, and the results obtained.

A. Analysis of Labor-Supply Response

The experiments provide panel data with substantial variation in certain critical variables (tax rates and guarantee levels) that are particularly useful in measuring labor-supply response. Consider a standard representation of annual hours worked, h_{it}, by individual i in time period t:

$$h_{it} = x_{it}\beta + \alpha_i + \eta_t + \xi_{it}, \quad i = 1, \ldots, n; t = 1, \ldots, T. \tag{6}$$

where x_{it} represents observable determinants of hours worked (experimental status, wage rate, household income, etc.), α_i represents unobservable individual effects (aptitude, ambition, etc.), η_t represents time effects (economic growth, technical change, etc.), and ξ_{it} is a standard disturbance term. Estimates of equation (6) may be biased if, for example,

$$E[\alpha_i \mid x_i] \neq 0,$$

a concern of many analysts since the early literature (Garfinkel 1973; Greenberg and Kosters 1973). If the individual effects are fixed, however, we can estimate an equation of the form

$$h_{it} - h_i = [x_{it} - x_i]\beta + [\eta_t - \eta] + [\xi_{it} - \xi_i] \tag{7}$$

without bias from this source, where h_i, x_i, and η represent the mean over all time periods or some specific time period (such as the preexperimental observation) for individual i and $\eta_t - \eta$ can be represented as fixed time effects by dummy variables for panel data of moderate duration (3 years). Moreover, an experimental design that provides for random assignment of households to treatment and control plans will ensure independence of determinants $x_{it} - x_i$ and errors $\xi_{it} - \xi_i$ so as to provide unbiased estimates of labor-supply response behavior with relatively simple and well-understood statistical methods.

Analysis of labor-supply response can be conventionally divided into two main types, based on equation (7): (*a*) nonstructural, analysis-of-variance (ANOVA) methods and (*b*) structural labor-supply models. The ANOVA methods simply involve the differentiation of treatment and control groups by means of dummy variables, often dropping the time effects and differentiating distinct plans (Hall [1973], for New Jersey; Ashenfelter [1978], for Rural; Robins and West [1978, 1980], for Seattle-Denver) or by means of spline series (Watts, Poirier, and Mallar 1976). Structural labor-supply models, however, are derived from an economic model such as equation (5), where

$$[x_{it} - x_i] \equiv [dw, hdw + dy]$$

to capture changes in after-tax wages and unearned income (evaluated at preexperimental hours worked), which will be primarily experienced by the treatment group (Keeley, Robins, Spiegelman, and West [1978], for Seattle-Denver; Hum and Simpson [1991], for Mincome).

Structural labor-supply models require more careful specification of the structure of labor-supply response and increase the possibility of specification error or misinterpretation of labor-supply response behavior. Analysis of variance models avoid this problem but cannot be easily generalized for social policy analysis when only dummy variables are used to distinguish experimental plans. The estimated response necessarily applies only to the specific experimental programs tested, and its implications for the evaluation of any prospective guaranteed income program with quite different features—such as the universal income security program recently proposed by the Macdonald Commission in Canada (Canada 1985)—are unclear. Only modified ANOVA (or analysis of covariance; ANCOVA) models that include the experimental design parameters (G and t), such as the spline functions estimated for New Jersey, can predict response to such a program. Estimates from structural labor-supply models, however, can be applied to the evaluation of social policy in general as well as the guaranteed annual income concept (Keeley 1981). Thus, both ANOVA and structural models provide useful information about labor-supply response, and estimates of each type are considered below.

B. Problems in the Analysis of the Experimental Data

Experimental data potentially resolve some very serious problems in the measurement of labor-supply response to a guaranteed annual income plan. Yet several problems remain, the most important of which appear to be nonrandom selection and nonparticipation. We shall briefly outline the nature of these problems for analysis of the experimental data and then discuss the evidence on the impact of these problems on estimated labor-supply response.

The Conlisk-Watts (1969) assignment model used by all five experiments favors inexpensive observations to improve estimation reliability when the experimental budget is constrained. Nonrandom assignment will therefore occur because the cost of a treatment observation depends on household income that, in turn, affects labor supply, as can be seen from equation (2). Thus, families with low preexperimental income are less likely to be allocated to generous plans (low t, high G), but their low income likely means that they also supply little labor prior to the experiment. In this case, the allocation of families to the treatment group is not independent of labor supply and, hence, not independent of the error term in equation (7). This will introduce bias to ordinary least squares (OLS) estimates of equation (7) whether it be an ANOVA or structural equation. Keeley and Robins (1978) argue that the only way to correct for the bias from nonrandom assignment would be to include all assignment variables as control variables in the estimated model. The resulting estimates of the experimental effect would be unbiased conditional on the particular assignment made. But, since assignment varies according to family size and composition as well as preexperimental income, the number of assignment categories will likely be very large, and this tactic could seriously reduce the reliability of the estimates (Keeley 1981).

The decision whether or not to participate in the GAI experiment depends to some extent on the expected financial gains. Families below the break-even level, B, for the assigned plan face an experimentally altered tax-transfer system that corresponds to the textbook-theoretical case of an unanticipated shift in the budget constraint of the household. Problems of nonparticipation arise for treatment-assigned households at or above B for prevailing (preexperimental) labor supply, however, because they are not affected by their assigned plan at the margin. These households may or may not choose to participate, depending on the size of the compensated wage effect and the proximity of their income to B (Ashenfelter 1980). The problem may be further complicated by nonrandom assignment because break-even status and labor supply depend on preexperimental income and family size, which also affect assignment.

Many families in both the treatment and control groups left the experiments before their completion. In New Jersey, for example, 374 of 1,357 families enrolled (28%) did not complete the experiment; in Mincome, 427 of 1,187 families (36%) enrolled in Winnipeg failed to complete all surveys. If this attrition were random, then the only concern would be loss of efficiency from declining sample size. However, decisions to leave the experiment likely depend on the financial incentives to stay, which vary with break-even status and the tax rate (eq. [1]). Since the assignment of plans depends on preexperimental income and labor supply, the incentives to remain in the experiment will also be related to these variables, introducing another potential source of sample selection bias to the analysis of labor-supply response. In sum, despite the potential and promise of experimental data to answer questions about the likely work response to a guaranteed-income program, the experiments are not without their own analytic headaches.

C. The Experimental Evidence for Labor-Supply Response

As discussed in Section IVA, both ANOVA and structural models provide useful information about the labor-supply response to a GAI using the experimental data. The ANOVA model gives a direct answer to the question, "Was there an experimental response (i.e., does labor supply response differ between the treatment and control groups)?" The structural model, in contrast, answers the question, "What was the experimental response in terms of conventional (substitution and income) labor-supply effects?" The models should provide similar answers concerning labor-supply response to a guaranteed-income program, but the results from the structural model may be more useful in other social policy evaluation.

In table 2 we summarize the evidence from a variety of studies on the difference in mean annual hours worked between the treatment and control groups in the five experiments. For the U.S. experiments, the surveys of

Table 2
Nonstructural Labor-Supply Response Estimates from the Five Experiments (in Annual Hours Worked)

	Husbands		Wives		Single Female Heads	
Experiment/Author	Estimates	%	Estimates	%	Estimates	%
New Jersey:						
Keely (1981)	−116	7	−75	33	...	
Robins (1985)	−34	2	−56	25	...	
Burtless (1986)	−21	1	−56	25	...	
Rural:						
Keeley (1981)	?	9	?	29*	...	
Robins (1985)	−56	3	−178	28	...	
Burtless (1986)	−56	3	−178	28	...	
Seattle-Denver:[a]						
Keeley (1981)	−147	8*	−139	21*	−155	15*
Robins (1985)	−113	7*	−141	21*	−163	16*
Burtless (1986)	−144	8	−107	17	−85	9
Gary:						
Keeley (1981)	−80	5	−9	3	−102	28
Robins (1985)	−35	2	−58	20	−37	10
Burtless (1986)	−114	7	+14	5	−112	30
All U.S. experiments:						
Robins (1985)	−89	5	−117	21	−123	13
Burtless (1986)	−119	7	−93	17	−133	17
Mincome:						
Appendix	−17	1[b]	−15	3	−79	7

[a] 3-year experiment only.
[b] Includes single individuals (21% of all men in sample).
* Statistical significance at the 5% level or lower. In some cases, statistical significance is not reported or is mixed (the result is an average of several results, some of which are significant). Burtless (1986) does not report statistical significance.

experimental evidence by Keeley (1981, chap. 5), Robins (1985), and Burtless (1986) are presented. For the Canadian experiment, the results are from the Appendix to this article. The results indicate, as expected, that hours worked will decline with the introduction of a guaranteed-income program, although the size of the decline depends on several factors. The reduction in hours worked is very small for men, never exceeding 9%, but larger for women.[11] Weighted averages of the U.S. results by Robins and Burtless imply a reduction in hours worked of about 6% for husbands, 19% for wives, and 15% for single mothers. Note, however, that only the results for Seattle-Denver are statistically significant; as far as can be determined from published reports, the estimates for New Jersey, Gary, and the Rural experiments are generally insignificant. Response in the Canadian experiment is similarly modest—1% for men, 3% for wives, and 5% for unmarried women—and statistically insignificant when time effects are properly included as in equation (7).[12] (See Table A1 in the Appendix for results from Mincome.)

One area of concern may be the potential bias arising from nonrandom sample allocation and participation in the experiments. Ashenfelter (1980, 1983) for the Seattle-Denver experiment and Sabourin (1985) for Mincome both find that participation is primarily determined by eligibility (break-even status) rather than choice (labor-supply response), implying that participation behavior should have little effect on measured labor-supply response in table 2. Hausman and Wise (1979), using a probability model of attrition in conjunction with random effects models of individual response, find no attrition bias for a structural model but some evidence of bias for nonstructural (ANOVA) models in the Gary experiment. Robins and West (1986) combine evidence from the Social Security Administration earnings records and the Seattle-Denver data base to test various hypotheses regarding attrition bias in an ANOVA framework, but they conclude that attrition bias is not likely a serious problem. Ashenfelter and Plant (1990), however, find evidence of systematic attrition from the Seattle-Denver Experiment and conclude that nonparametric estimates of labor-supply response are sensitive to attrition behavior. Hum and Simpson (1991, chap. 7) find that assignment variables (preexperimental income, break-even status, and family size) are insignificant in the ANOVA model once time effects have been considered. While there is no clear evidence that allocation and participation bias are a serious concern in the experiments, further

[11] We do not consider the variation in response by race. For a summary of the U.S. evidence on this issue, see Robins (1985); or Burtless (1986). The Canadian experiment did not stratify observations by race.

[12] Hum and Simpson (1991) find significant reductions in mean hours worked for the experimental groups, but these effects disappear when time dummies are included.

research is warranted in view of the recent evidence from Ashenfelter and Plant (1990).[13]

It is difficult to conclude whether the experimental effects in table 2, even the statistically significant ones, are large or not. The answer to that question depends on the generosity of the guaranteed-income plan offered, and the results in table 2 are estimates of the response to specific experimental treatments; that is, the results are some weighted average of all programs tested at each site. The results merely imply that there would be a reduction in work hours if something like the "average plan" were implemented. Moreover, it is very difficult to compare these results with nonexperimental evidence based on structural labor-supply models.

The structural models are ostensibly based on equation (5), although specific regression models of this sort were only estimated for Seattle-Denver (Keeley et al. 1978) and Mincome (see the Appendix). For example, the model specified by Keeley et al. combines equations (5) and (7) to obtain

$$\begin{aligned}\Delta h &\equiv h_e - h_p \\ &= \beta_0 + \beta_1 \Delta w + \beta_2[h_p \Delta w + \Delta y] + Z\alpha + \xi,\end{aligned} \tag{8}$$

where Δh is the change in hours worked between the experimental and preexperimental periods, Δw is the change in after-tax wages, $h_p \Delta w + \Delta y$ is the change in income evaluated at preexperimental hours, and Z is a set of control variables. The compensated wage (substitution) effect β_1 and the income effect β_2 are estimated directly. Keeley et al. actually estimate a related equation of the form

$$h_e = \beta_0 + \beta_1 \Delta w + \beta_2[h_p \Delta w + \Delta y] + \beta_3 h_p + Z\alpha + \nu \tag{9}$$

using a Tobit regression, where h_p is included as an explanatory variable to correct for any bias caused by the substitution of preexperimental labor supply for permanent labor supply in the determination of the change in income. Other estimates, such as those presented for New Jersey, Gary, and the Rural experiments by Robins (1985) are typically indirect estimates of parameters of the structural model that may not be strictly comparable to those from Seattle-Denver and Mincome.[14]

[13] Other problems, such as the duration of the experiment, have been analyzed elsewhere by the comparison of 3-, 5-, and 20-year samples in the Seattle-Denver Experiment (Robins 1984). The results do not suggest that the experimental estimates are biased, although attrition from the 20-year sample led to its abandonment after 12 years.

[14] For further discussion of the methodologies used in the New Jersey, Gary, or the Rural experiment, see Keeley (1981); or Moffitt and Kehrer (1981).

Table 3
Structural Labor-Supply Response Estimates from the Five Experiments

Experiment/Group	Substitution Elasticity	Income Elasticity
Husbands:		
New Jersey	.09	−.02
Rural	.09	.00
Seattle-Denver	.09	−.14
Gary	.06	−.08
All United States	.08	−.10
Mincome	−.07	−.03
Wives:		
New Jersey	−.08	−.28
Rural	.28	.01
Seattle-Denver	.14	−.12
Gary	.37	.26
All United States	.17	−.06
Mincome	−.08	.07
Single female heads:		
New Jersey		
Rural		
Seattle-Denver	.12	−.15
Gary	.14	−.20
All United States	.13	−.16
Mincome	−.17	−.01

SOURCES.—Robins (1985) for the U.S. experiments; the Appendix (table A3) for the Canadian experiment.

Table 3 provides a summary of the evidence on labor-supply response consistent with structural models. Elasticity estimates are presented to facilitate comparison across experiments and with nonexperimental research and to provide the most useful information for current policy analysis. The results from the GAI experiments, in contrast with the nonexperimental literature, provide very uniform and quite low elasticity estimates of labor-supply response. The experimental evidence therefore corroborates the nonexperimental evidence of very inelastic male labor-supply response (Keeley 1981; Killingsworth 1983; Pencavel 1986; Hum and Simpson 1991). Robins's (1985) weighted average for all experiments is 0.08 for the substitution elasticity and 0.10 for the income elasticity.

Unlike the bulk of the nonexperimental evidence,[15] however, the experimental evidence for wives and single mothers indicates similarly inelastic response. Robins's estimates are 0.17 and 0.13 for the substitution elasticities for wives and single mothers and 0.06 and 0.16 for their re-

[15] Reviews include Keeley (1981); Killingsworth (1983); Killingsworth and Heckman (1986); and Hum and Simpson (1991).

spective income elasticities. Unfortunately, the statistical significance of the estimates in table 3 is either unavailable (because the substitution is calculated from the Slutsky formula, for example) or not reported in most cases. Robins (1985) notes, however, that 20% of his average estimated substitution and income effects for individual U.S. experiments contradict theoretical prediction (positive substitution effects and negative income effects); only in the Seattle-Denver experiment are results uniformly consistent with standard theory, although the contradictory results elsewhere may often be statistically insignificant and hence interpreted as zero.

Recent research by Mroz (1987) and MaCurdy, Green, and Paarsch (1990) explain why labor-supply elasticity estimates have been exaggerated in studies using nonexperimental data, particularly for married women. Mroz finds evidence of severe misspecification bias in most studies of working wives. The few nonexperimental studies that pass Mroz's specification tests corroborate the low elasticity estimates found using experimental data. MaCurdy et al. show that estimation procedures to correct for biases arising from nonlinear budget constraints, commonly carried out on nonexperimental data sets from married women in the last decade, are themselves biased toward higher wage effects.

Preliminary evidence from the Canadian experiment agrees with evidence from the U.S. experiments that labor-supply elasticity estimates are much smaller than corresponding estimates from such nonexperimental data as the Mincome preexperimental survey (Prescott, Swidinsky, and Wilton 1986; Hum and Simpson 1991). We use a within-groups estimator based on equation (5) to explain the change in hours worked for panel t relative to a base period (the average of all panels):

$$h_{it} - h_i = \beta_0 + \beta_1[w_{it} - w_i] + \beta_2[h_i(w_{it} - w_i) + v_{it} - v_i] + \beta_3[c_{it} - c_i] + \sum_{i=2}^{4} \gamma_i t_i + \xi, \tag{10}$$

where $w_{it} - w_i$ is the change in after-tax wages from the base period, $v_{it} - v_i$ is the change in income from a base period, $c_{it} - c_i$ is the change in the number of preschool children in the household from the base period, t_i is a set of dummy variables where $t_i = 1$ for the 1975 panel ($i = 2$), 1976 panel ($i = 3$), and 1977 panel ($i = 4$), and ξ is the residual term. If unobserved individual effects and experimental design effects from nonrandom assignment and participation are fixed, then the within-groups estimator should yield unbiased estimates of the substitution effect β_1 and the income effect β_2. The model is same as Keeley et al. (1978; see eq. [8] above), with the addition of a variable to reflect any changes in the number of preschool children in the household, but they opt for equation (9), which requires inclusion of a complete set of control variables. Moreover, as Mroz (1987) demonstrates, the Tobit estimator they use has not proven reliable in estimating labor-supply elasticities for nonexperimental data. Hum and Simpson (1991) find that the Keeley model generally produces larger estimated substitution elasticities than a within-groups estimator, a result that agrees with Mroz.

Our results from Mincome are presented in table A2 in the Appendix, and the elasticities are calculated in table A3. The initial results indicate negative compensated wage elasticities, contrary to expectations, but they are generally insignificant and should likely be regarded as zero. Since many studies of the other experiments produced perverse, but small and often insignificant, wage and income effects, our results are consistent with results from New Jersey, Gary, and the Rural experiment, if not Seattle-Denver. The one consistently important factor in the labor-supply response of married men and women in table A2 is the presence of preschool children in the home. Preschool children significantly increased the labor supply of the husband and reduced the labor supply of the wife by roughly the same amount. Indeed, the preliminary results from Mincome indicate that changes in family composition may have far more impact on labor supply than a guaranteed-income program.[16] Some changes in family structure, such as marital dissolution, may not be independent of changes in the tax-transfer system, however, as we consider in the next section.

V. Responses Other than Labor Supply

The NIT experiments understandably concentrated on the work response to a guaranteed income, not only because of its significance for calculating program costs should a nationwide GAI be implemented as a tactic for combating poverty but also because the fear of wholesale work withdrawal from the able-bodied population was perceived to be the ultimate stumbling block. At the same time, a number of behavioral responses other than labor supply were investigated. The list of topics examined may be loosely classified in terms of (*a*) consumption pattern studies, in which expenditure by experimental participants on such goods as clothing, housing, consumer durables, food, health care, debt, and asset accumulations are examined; (*b*) human capital investment studies, in which responses such as child

[16] One unresolved question from our results for Mincome is why wives in the experimental group worked fewer hours (table 2) if their substitution and income elasticities are perverse (table 3). If their substitution elasticity is negative and their income elasticity is positive, then lower after-tax wages and higher virtual income for the experimental group should lead to more hours worked, not fewer. One consideration is that the reported effects in tables 2 and 3 for Mincome are small and statistically insignificant (i.e., really all zero), but the incompatible point estimates must also reflect the inclusion of changes in the number of preschool children in the structural model (table A2). Thus the possible endogeneity of family structure, including fertility, needs to be examined in future work.

care utilization, counseling, school attendance, nutrition, migration, and geographical mobility constitute the focus of interest; and (*c*) noneconomic, or "sociological" responses, a term we shall use to embrace such wide-ranging topics as psychological well-being, marital stability, delinquency, political participation, educational aspirations, and family life. Included here might even be such administrative concerns as misreporting behavior, participant comprehension of program rules, and the like. Thus a rough distinction is possible between consumption and human capital studies in terms of "short-run" versus "longer-run" responses.

Clearly, the above list of topics is wide-ranging, and any attempt to provide a rigid classification would appear futile. As well, many of the studies are not corroborated and are exceedingly difficult to assess. Not only are the non-work-response results more complex in terms of an expected behavioral result, they are also necessarily much more diffuse since there is often no common point of departure, theoretical structure, or even well-defined empirical technique. Nonetheless, the above loose classification can serve to highlight a number of concerns from a policy perspective.

The consumption studies are of interest because of the light it might shed on whether GAI recipients alter their expenditure patterns in a "socially acceptable" manner. In reviewing some of the consumption studies from the Rural experiment, Michael (1978) argued that the investigators were not armed with a clear-cut social issue or an urgent programmatic issue, unlike the question of work withdrawal and its impact on program costs. Masters (1978) disagreed, as did Baumol (1974) in an earlier review of the New Jersey results: "Those who fear the worst of a [GAI] may hold the hypothesis that a large part of the payments will be wasted by the recipients—either being spent on drugs, drinks, and gambling or being dissipated in increased leisure time unproductively used" (p. 253). Alternatives to this view would include the possibility that a GAI will not interfere materially with life-styles at all and that non-work-response effects are minimal.

Baumol's quotation is useful in reminding us that much of the passion and controversy surrounding a GAI centers around what society is willing to accept as a socially approved response to unconditional cash payments. In the case of labor markets, this is clearly revealed as a work disincentive issue. In the case of consumption, this appears more subtly as disapproval over the way GAI monies are spent. As Masters (1978, p. 172) engagingly puts it: "The labour supply analysis is relevant for the stereotype of the poor as lazy bums. The expenditure analysis could be relevant for the stereotype of the poor as profligate boozers."

The consumption studies are also of potential policy importance in deciding between delivery of in-kind benefits versus cash transfers, especially in areas such as housing and food and possible child care or education.

However, many have come to the view that, on such matters, research programs other than experimentation would probably provide better estimates of behavioral reactions (Hanushek 1986). Canadians and Americans will have different policy preferences on this topic as well.

Rather less was studied in the experiments concerning human capital investment than was accomplished for the consumption studies. This is perhaps understandable given the very short duration of each experiment, which could be expected to be more problematic for human capital investment response than either work effort or consumption of nondurables. Metcalf (1973, 1974) has especially emphasized the fact that a limited duration experiment may underestimate long-term income effects and overestimate long-run price effects. In any case, it is fair to conclude that the various consumption and human capital investment studies from the NIT experiments have had little impact on policy. This is because, in general, the studies show that the experiments had little or no discernible impact on consumption and investment decisions, or, where any response was detected, it was either slight, mildly beneficial, or, in the case of housing, it merely altered the timing of already planned purchases (Hanushek 1986). Furthermore, the tone of all these studies, taken together, would suggest that the NIT payments were spent in much the same manner as money income received in other ways.

The various sociological responses are also hard to assess, partly because economists often question the reliability of the scale measures usually constructed. As well, it is generally thought that many of the issues examined under the rubric of sociological responses—such as psychological well being and marital interaction—are not well served by the NIT experimental designs and that any information obtained on these topics is simply a welcome bonus. The one exception is the matter of marital instability or family dissolution.

The impact of a GAI on marital disruption and family composition has become a major controversy and probably now tops the agenda of the policy debate concerning a GAI among policymakers and academics alike. The reasons are quite transparent. First, the cost implications strike a familiar chord; families induced to split in order to receive larger benefits will add to overall program costs, the same fear that motivated the concern over work disincentives. Second, a GAI program that actually encourages marriages and families to break up is not acceptable either to policymakers nor the general public. The third factor is the controversial nature of recent research on this subject.

The initial findings from the New Jersey experiment that GAI payments might influence to some degree the breakup or durability of a family was lost amid the rush of findings on work response. However, the startling findings reported by Hannan, Tuma, and Groeneveld (1977, 1978) and Groeneveld, Tuma, and Hannan (1980) that a NIT program dramatically increased marital dissolutions has been recently challenged by Cain (1986). Reexamining the same evidence, Cain and Wissoker (1990*a*) find only mild or insignificant effects on marital instability. As Murray (in this issue) notes with obvious reference to the American controversy, "The dust will settle eventually" (p. S233). Meantime, Allen (in this issue) is left to explore the same difficult questions for Canada in less than ideal circumstances, using Canadian census data and concentrating on welfare in general rather than on a GAI program. The Canadian experimental evidence respecting a GAI and marital stability is just emerging, and it is too early to tell how "the snow will pack" eventually. The preliminary results would suggest a moderate response of marital dissolution to NIT payments (Choudhry 1989; Hum and Choudhry 1992, in press). This debate goes on, even among the principals (Cain and Wissoker 1990*b;* Hannan and Tuma 1990), and will doubtlessly continue among policymakers for some time, especially as the controversy over work incentives abates. The battleground of the GAI may be expected to shift to the link between welfare structure and family composition, a much more intuitive and accessible affair than technical squabbles over wage and income elasticity estimates.

VI. Concluding Remarks

Canada and the United States have followed similar paths with respect to the guaranteed annual income over the last quarter century. Dissatisfaction with public assistance programs led both countries to consider a guaranteed-income plan. This consideration was serious enough to motivate large-scale social experiments to determine the economic impact of such a program. The analysis of the experiments and the GAI issue is by no means exhausted, but a large volume of research has now accumulated to pinpoint our common progress toward understanding the economic effects of a GAI plan. We have tried, in the limited space available, to present and assess that evidence from American sources and to provide some initial comparable results from the Canadian experiment, paying particular attention to the primary policy and research topic, namely, labor-supply response or work incentives.

If we were asked to summarize "in 25 words or less" what has been learned from the experiments about the economic effects of a GAI plan we would respond: "Few adverse effects have been found to date. Those adverse effects found, such as work response, are smaller than would have been expected without experimentation." Indeed, in the emerging consensus among economists that elasticities of labor-supply response are smaller than previously estimated, particularly for married women, we argue that the experimental evidence has played an important role (Hum and Simpson 1991). That consensus should influence future policy debate over the costs of a GAI, as well as other social policy reform. Burtless (1986), for example, estimates that the cost in terms of earnings reduction from work disincentives of a fairly generous GAI (75% of the poverty line with a tax rate of 50%) would be between 30 and 60 cents for every dollar transferred to the poor, based on the Seattle-Denver results. Since the other experiments find smaller labor-supply response, this likely provides a high estimate of the cost of a prospective GAI program, but an estimate that is much smaller, and much more precise, than nonexperimental evidence would provide. Whether such costs are "high," and whether such costs are "worth it," depends on the assessor and, ultimately, on political assessment of public support for a GAI. But that assessment will be better informed now that the experiments have reported, if (when?) welfare reform and the GAI return to the policy limelight.

Another question that might be asked is whether the GAI experience contains lessons for the evaluation of social policy. We would argue that valuable lessons have been learned on a variety of issues from experimental design to analytic methods. Lessons continue to be learned from reassessment of the experiments and applied to other, more modest, evaluations of social policy such as employment and training policy in the United States. To those who argue that the GAI debate and income maintenance experimentation was not worth the money, we would simply observe that the money involved was small in relation to annual expenditures on social programs in Canada and the United States. If the GAI experience can sharpen policy debate and help us to avoid ill-advised social policy decisions in the future, then it was likely a solid investment.

Appendix

Empirical Results from the Mincome Experiment

Our basic equation for the analysis of experimental effects is given by equation (7) in Section IV A:

$$h_{it} - h_i = [x_{it} - x_i]\beta + [\eta_t - \eta] + [\xi_{it} - \xi_i] \tag{7}$$

for individual i and time period t, where y is annual hours worked and $\eta_{it} - \eta$ is represented as fixed time effects by three annual dummy variables. Our analysis varies the specification of $[x_{it} - x_i]$ as follows:

(i) nonstructural models:

$$(a)\ [x_{it} - x_i] = T_i = 1 \quad \text{if treatment group,}$$
$$= 0 \quad \text{if control group,}$$

$$(b)[x_{it} - x_i] = [P_1, \ldots, P_8], \text{ where}$$
$$P_j = 1 \quad \text{if assigned to plan } j,$$
$$= 0 \quad \text{if otherwise;}$$

(ii) structural models, based on equation (5):

$$[x_{it} - x_i] = [w_{it} - w_i, h_i(w_{it} - w_i) + v_{it} - v_i, c_{it} - c_i],$$

where w_{it} is the imputed after-tax wage, h_i is the preexperimental hours worked by that individual, v_{it} is the virtual income of the household, and c_{it} is the number of children in household i at time t. Thus, $w_{it} - w_i$ represents the change in after-tax wages while $v_{it} - v_i$ represents the change in virtual income. For the experimental group, the mean of $w_{it} - w_i$ is negative while for the control group it is positive; for the experimental group $v_{it} - v_i$ is positive while for the control group it is negative.[17] These variations, in addition to the panel character of the data, should permit more precise estimates of wage and income effects than nonexperimental data.

Note that, if

$$h_i = \sum_{t=1}^{T} h_{it} \qquad \text{and} \qquad x_i = \sum_{t=1}^{T} x_{it}$$

represent the mean of all time periods under consideration in the panel, then we have the conventional within-group estimator. This means that there is no need to include baseline control variables as in Keeley et al. (1978) to correct for nonrandom assignment and participation. Unbiased estimates of the experimental effects rest only on the assumption that these problems are fixed effects, or approximately so, and not on ad hoc specification of observable control variables. Hence our model is more credible in that it is not susceptible to misspecification of baseline control variables. This may be a significant advantage since, as we discuss in Section IV B, there are many potential control variables arising from the assignment model and family eligibility conditions (Keeley 1981).

The nonstructural estimates of labor-supply response, corresponding to (i)(*a*) and (i)(*b*) above, are presented in table A1 for all men, wives, and single female heads of households. To inspect the argument that there is a "phase-in" effect and a "phase-out" effect to the experiment, we also estimate the model deleting the first and last experimental years in table A1. Thus, in the final two columns of table A1 we use only the midyear of the experiment as in Keeley et al. (1978) for Seattle-Denver.

The results, which are summarized in table 2 in the main text, indicate fairly weak response to the experimental treatment. In fact, although the results indicate modest reductions in hours worked for all groups as predicted, the experimental effects are uniformly insignificant at the 5% level.[18] The fixed time effects are quite large, negative and significant for men, indicating that omission of these variables would lead to an overestimate of the experimental effect by confounding it with a general decline in hours worked unrelated to the experimental treatment.

[17] See Hum and Simpson (1991, table 8-2) for further details.

[18] This includes an F-test for the significance of P_j ($j = 1, \ldots, 8$) as a group.

Table A1
Nonstructural Estimates of the Labor-Supply Response in the Mincome Experiment for Men, Wives, and Single Female Heads

	1974–77		1974, 1976	
	(1)	(2)	(3)	(4)
Men:				
Constant	123.7*	123.7*	101.8*	101.8*
T	−14.6	...	−20.3	...
P_1	...	−16.4	...	−12.7
P_2	...	13.3	...	−3.2
P_3	...	.04	...	42.9
P_4	...	−27.0	...	−59.5
P_5	...	−20.1	...	−51.7*
P_6	...	29.9	...	144.2
P_7	...	16.9	...	10.2
P_8	...	.2	...	−1.1
1975 = 1	−108.2*	−108.2*	...	...
1976 = 1	−170.4*	−170.4*	−190.1*	−190.1*
1977 = 1	−187.1*	−187.1*	...	...
F	10.8*	4.4*	24.2*	7.4*
N	1,284		806	
Wives:				
Constant	32.7	32.7	30.6	30.6
T	−16.0	...	−13.5	...
P_1	...	−24.4	...	−56.3
P_2	...	−3.8	...	2.1
P_3	...	10.0	...	113.4
P_4	...	.5	...	11.5
P_5	...	−11.8	...	−19.6
P_6	...	−56.3	...	−134.1
P_7	...	15.8	...	120.7
P_8	...	5.7	...	−10.6
1975 = 1	−6.1	−6.1	...	...
1976 = 1	−54.5	−54.5	−52.5	−52.5
1977 = 1	−39.0	−39.0	...	...
F	1.5	.8	2.9	2.0
N	1,176		664	
Single female heads:				
Constant	84.5*	84.5*	67.5*	67.5*
T	−43.7	...	−114.7	...
P_1	...	1.6	...	29.3
P_2	...	−49.3	...	−114.4
P_3	...	−89.7	...	−186.9*
P_4	...	−73.2	...	−234.2
P_5	...	−5.2	...	−35.4
P_6	...	−27.7	...	−32.1
P_7	...	59.2	...	42.9
P_8	...	−49.2	...	−105.0
1975 = 1	−58.9	−58.9	...	...
1976 = 1	−104.1	−104.1	−55.2	−55.2
1977 = 1	−84.5	−84.5	...	...
F	2.3	1.0	6.9*	2.2*
N	580		368	

NOTE.—$T = 1$ if experimental group; $P_1 = 1$ if $G = 3{,}800$, $t = .35$; $P_2 = 1$ if $G = 4{,}800$, $t = .35$; $P_3 = 1$ if $G = 3{,}800$, $t = .50$; $P_4 = 1$ if $G = 4{,}800$, $t = .50$; $P_5 = 1$ if $G = 5{,}800$, $t = .50$; $P_6 = 1$ if $G = 3{,}800$, $t = .75$; $P_7 = 1$ if $G = 4{,}800$, $t = .75$; $P_8 = 1$ if $G = 5{,}800$, $t = .75$. The results reported in table 2 for Mincome are the mean of the simple treatment effects in the second and fourth columns.

* Significant at the 5% level.

The structural estimates of labor-supply response, corresponding to (ii) above, are presented in table A2 for husbands, wives, and single female heads. Again, we have deleted 1975 and 1977 in the final two columns to account for possible "phase-in" and "phase-out" effects. The results of table A2 indicate weak income and compensated wage effects, as expected from the results in table A1. The elasticity estimates are calculated in table A3 and summarized in table 3 in the main text. The income elasticity estimates are uniformly small, generally insignificant, and mostly negative, as expected. The compensated wage effects, however, are generally negative, contrary to expectations, although generally insignificant as well. The weak income effect and the weak compensated wage effect are obviously consistent with the weak labour supply response in table A1.

Table A2
Structural Estimates of the Labor-Supply Response in the Mincome Experiment for Husbands, Wives, and Single Female Heads

	1974–77	1974, 1976
Husbands:		
Constant	100.4*	99.3*
$w_{it} - w_i$	−31.4	−37.2
$h_i(w_{it} - w_i) + v_{it} - v_i$	−.01	−.01
$c_{it} - c_i$	106.8*	126.7*
1975 = 1	−113.3*	. . .
1976 = 1	−146.6*	−198.6*
1977 = 1	−141.5*	. . .
F	9.8*	13.5*
N	852	476
Wives:		
Constant	127.6*	128.8*
$w_{it} - w_i$	−66.1*	−36.7
$h_i(w_{it} - w_i) + v_{it} - v_i$	.01	.04*
$c_{it} - c_i$	−106.6*	−109.7*
1975 = 1	−137.7	. . .
1976 = 1	−181.5	−257.7*
1977 = 1	−191.3	. . .
F	5.1*	7.2*
N	884	474
Single female heads:		
Constant	93.7	116.9*
$w_{it} - w_i$	−86.2*	−84.1
$h_i(w_{it} - w_i) + v_{it} - v_i$	−.03	.002
$c_{it} - c_i$	−176.1	−189.1*
1975 = 1	123.3	. . .
1976 = 1	138.5	−233.8*
1977 = 1	−112.9	. . .
F	5.0*	7.1*
N	468	296

NOTE.—The wages are imputed from a wage regression corrected for sample selection bias arising from the omission of nonworkers and including the following variables: education, education squared, work experience, experience squared, age, age squared, age times education, and time dummies as above (1975 = 1, 1976 = 1, 1977 = 1).
* Significant at the 5% level.

Table A3
Labor-Supply Elasticity Estimates from Results in Table A2 for Husbands, Wives, and Single Female Heads

	1974–77	1974, 1976
Husbands:		
Own-wage	−.07	−.07
Income	−.03	−.03
Wives:		
Own-wage	−.12	−.05
Income	.03	.11
Single female heads:		
Own-wage	−.16	−.19
Income	−.07	.05

NOTE.—Elasticity estimates are calculated on mean wages, income, and hours worked in the sample. Elasticities reported in table 3 in the text are based on the simple average of the first and second columns.

Finally, we also include changes in the number of children in the family as an indicator of changes in family structure, primarily arising from births. An additional child significantly increases the labor supply of the husband and significantly reduces the labor supply of the wife by roughly the same amount. Clearly, additional children alter the allocation of time within the household and should not be ignored in assessing labor-supply response.

Our structural labor-supply results are consistent with recent findings of weak labor-supply response to tax-transfer changes, although much more remains to be done. In particular, the entire issue of family structure and labor supply would seem to be a fertile area for further investigation of labor-supply behavior using experimental data. These preliminary results are intended to be comparable to earlier studies, ignoring such important features of household labor supply behavior as cross-wage effects and marital stability. In subsequent work, we intend to extend our analysis to investigate these issues.

References

Allen, Douglas W. "Welfare and the Family: The Canadian Experience." In this issue.

Ashenfelter, Orley. "The Labor Supply Response of Wage Earners." In *Welfare in Rural Areas,* edited by J. Palmer and J. Pechman. Washington, D.C.: Brookings Institution, 1978.

———. "Discrete Choice in Labor Supply: The Determinants of Participation in the Seattle and Denver Income Maintenance Experiments." Working Paper no. 136. Princeton, N.J.: Princeton University, Industrial Relations Section, 1980.

———. "Determining Participation in Income-tested Social Programs." *Journal of the American Statistical Association* 78, applications section (1983): 517–25.

Ashenfelter, Orley, and Plant, W. Mark. "Nonparametric Estimates of the Labor-Supply Effects of Negative Income Tax Programs." *Journal of Labor Economics* 8, no. 1, pt. 2 (1990): S396–S415.

Basilevsky, Alexander, and Hum, Derek. *Experimental Social Programs and Analytic Methods.* New York: Academic Press, 1984.

Baumol, William. "An Overview of the Results on Consumption, Health, and Social Behavior." *Journal of Human Resources* 9 (1974): 253–64.

Burtless, Gary. "The Work Response to a Guaranteed Income: A Survey of Experimental Evidence." In *Lessons from the Income Maintenance Experiments,* edited by Alicia Munnell, pp. 22–52. Boston: Federal Reserve Bank of Boston, 1986.

Cain, Glen. "The Issues of Marital Stability and Family Composition and the Income Maintenance Experiments." In *Lessons from the Income Maintenance Experiments,* edited by Alicia Munnell, pp. 60–93. Boston: Federal Reserve Bank of Boston, 1986.

Cain, Glen, and Watts, Harold. *Income Maintenance and Labor Supply.* New York: Academic Press, 1973.

Cain, Glen, and Wissoker, Douglas. "A Reanalysis of Marital Stability in the Seattle-Denver Income-Maintenance Experiment." *American Journal of Sociology* 95 (1990): 1235–69. (*a*)

———. "Response to Hannan and Tuma." *American Journal of Sociology* 95 (1990): 1299–1314. (*b*)

Canada. *Income Security for Canadians.* National Health and Welfare white paper. Ottawa: Queen's Printer, 1970.

———. *Poverty in Canada* (the Croll Report). Special Senate Committee on Poverty. Ottawa: Queen's Printer, 1971.

———. *Working Paper on Social Security in Canada (the Orange Paper). National Health and Welfare.* Ottawa: Queen's Printer, 1973.

———. *Report of the Royal Commission on the Economic Union and Development Prospects for Canada* (the MacDonald Commission), vols. 1–3. Ottawa: Minister of Supply and Services, 1985.

Choudhry, Saud. *Income Maintenance Experiments and Household Transition Dynamics: A Temporal Treatment of Incremental Effects.* Ph.D. thesis, University of Manitoba, 1989.

Communique of meeting of federal and provincial ministers of welfare, April 30–May 1, 1975.

Conlisk, J., and Watts, H. "A Model for Optimizing Experimental Designs for Estimating Response Surfaces." *American Statistical Association Proceedings,* Social statistics section 64 (1969): 150–56.

Department of National Health and Welfare/Department of Health and Social Development. "$17 Million Basic Income Experiment to Start in Manitoba." Joint News release, Department of National Health and Welfare/Department of Health and Social Development, February 22, 1974.

Economic Council of Canada. *The Problem of Poverty: Fifth Annual Review.* Ottawa: Economic Council of Canada, 1968.

Ferber, Robert, and Hirsch, Werner. "Social Experimentation and Economic Policy: A Survey." *Journal of Economic Literature* 16 (1978): 1379–1414.

Friedman, Milton. *Capitalism and Freedom.* Chicago: University of Chicago Press, 1963.

Garfinkel, Irwin. "On Estimating the Labor-Supply Effects of a Negative Income Tax." In *Income Maintenance and Labor Supply,* edited by Glen Cain and Harold Watts, pp. 205–64. New York: Academic Press, 1973.

Greenberg, David, and Kosters, Marvin. "Income Guarantees and the Working Poor: The Effect of Income Maintenance Programs on the Hours of Work of Male Family Heads." In *Income Maintenance and Labor Supply,* edited by Glen Cain and Harold Watts, pp. 14–101. New York: Academic Press, 1973.

Groeneveld, L.; Tuma, N.; and Hannan, M. "The Effects of Negative Income Tax Programs on Marital Dissolution." *Journal of Human Resources* 15 (1980): 654–74.

Hall, Robert. "Wage, Income, and Hours of Work in the U.S. Labor Force." In *Income Maintenance and Labor Supply,* edited by Glen Cain and Harold Watts, pp. 102–62. New York: Academic Press, 1973.

Hannan, Michael, and Tuma, Nancy. "A Reassessment of the Effect of Income Maintenance on Marital Dissolution in the Seattle-Denver Experiment." *American Journal of Sociology* 95 (1990): 1270–98.

Hannan, Michael; Tuma, Nancy; and Groenveld, Lyle. "Income and Marital Events: Evidence from an Income-Maintenance Experiment." *American Journal of Sociology* 82 (1977): 1186–1211.

———. "Income and Independence Effects on Marital Dissolution: Results from the Seattle and Denver Income-Maintenance Experiments." *American Journal of Sociology* 84 (1978): 611–33.

Hanushek, Eric. "Non-Labor-Supply Response to the Income Maintenance Experiments." In *Lessons from the Income Maintenance Experiments,* edited by Alicia Munnell, pp. 106–21. Boston: Federal Reserve Bank of Boston, 1986.

Hausman, Jerry, and Wise, David. "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment." *Econometrica* 47 (1979): 455–73.

Haveman, Robert H., and Watts, Harold W. "Social Experimentation as Policy Research: A Review of Negative Income Tax Experiments." *Evaluation Studies* 1 (1976): 406–31.

Hum, Derek. *Federalism and the Poor: A Review of the Canada Assistance Plan.* Toronto: Ontario Economic Council, 1983.

———. "Harmonization of Social Programs under Free Trade." In *Free Trade and Social Policy,* edited by Glenn Drover, pp. 26–47. Ottawa: Canadian Council on Social Development, 1988. (*a*)

———. "Integrating Taxes and Transfers." *Canadian Tax Journal* 3 (1988): 671–90. (*b*)

Hum, Derek, and Choudhry, Saud. "Income, Work and Marital Dissolution: Canadian Experimental Evidence." *Journal of Comparative Family Studies* (1992, in press).

Hum, D.; Laub, M.; Metcalf, C.; and Sabourin, D. "The Sample Design and Assignment Model of the Manitoba Basic Annual Income Experiment." Technical Report no. 2. Winnipeg: Mincome Manitoba, 1979.

Hum, Derek, and Simpson, Wayne. *Income Maintenance, Work Effort, and the Canadian Experiment.* Ottawa: Economic Council of Canada, 1991.
Johnson, A. W. "Canada's Social Security Review (1973–75): The Central Issues." *Canadian Public Policy* 1 (1975): 456–72.
Keeley, Michael. *Labor Supply and Public Policy.* New York: Academic Press, 1981.
Keeley, Michael, and Robins, Philip. "The Design of Social Experiments: A Critique of the Conlisk-Watts Assignment Model." Research Memorandum no. 57. Menlo Park, Calif.: Stanford Research Institute, Center for the Study of Welfare Policy, 1978.
Keeley, M.; Robins, P.; Speigelman, R.; and West, R. "The Estimation of Labor Supply Models Using Experimental Data." *American Economic Review* 68 (1978): 873–87.
Kesselman, Jonathan. "Income Security via the Tax System: Canadian and American Reforms." Discussion Paper no. 90-31. Vancouver: University of British Columbia, 1990.
Killingsworth, Mark. *Labor Supply.* Cambridge: Cambridge University Press, 1983.
Killingsworth, Mark, and Heckman, James. "Female Labor Supply: A Survey." In *Handbook of Labor Economics,* vol. 1, edited by Orley Ashenfelter and Richard Layard, pp. 103–204. Amsterdam: North-Holland, 1986.
Lampman, Robert. "The Decision to Undertake the New Jersey Experiment." In *Final Report of the New Jersey Experiment,* vol. 4. Madison, Wis.: Institute for Research on Poverty and Mathematica, 1974.
Levine, Robert. "How and Why the Experiment Came About." In *Work Incentive and Income Guarantees,* edited by J. A. Pechman and P. M. Timpane. Washington, D.C.: Brookings Institution, 1975.
Levy, Frank. "The Labor Supply of Female Household Heads, or AFDC Work Incentives Don't Work Too Well." *Journal of Human Resources* 14 (1979): 76–97.
MaCurdy, Thomas; Green, David; and Paarsch, Harry. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *Journal of Human Resources* 25 (1990): 415–90.
Masters, Stanley H. "Comment on Robert Michael: The Consumption Studies." In *Welfare in Rural Areas,* edited by John Palmer and Joseph Pechman, pp. 171–73. Washington, D.C.: Brookings Institution, 1978.
Metcalf, Charles. "Making Inferences from Controlled Income Maintenance Experiments." *American Economic Review* 63 (1973): 473–83.
———. "Predicting the Effects of Permanent Programs from a Limited Duration Experiment." *Journal of Human Resources* 9 (1974): 530–55.
Michael, Robert. "The Consumption Studies." In *Welfare in Rural Areas: The North Carolina-Iowa Income Maintenance Experiment,* edited by John Palmer and Joseph Pechman, pp. 149–71. Washington, D.C.: Brookings Institution, 1978.
Moffitt, Robert, and Kehrer, Kenneth. "The Effect of Tax and Transfer Programs on Labor Supply: Evidence from the Income Maintenance Experiments." *Research in Labor Economics* 4 (1981): 103–50.
Mroz, Thomas. "The Sensitivity of an Empirical Model of Married

Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica* 55 (1987): 765–99.

Murray, Charles. "Welfare and the Family: The American Experience." In this issue.

National Council of Welfare. *Guide to the Guaranteed Income.* Ottawa: National Council of Welfare, 1976.

Orcutt, G., and Orcutt, A. "Incentive and Disincentive Experimentation for Income Maintenance policy Purposes." *American Economic Review* 58 (1968): 754–72.

Pechman, Joseph, and Timpane, Michael. "Introduction and Summary." In *Work Incentive and Income Guarantees,* edited by J. A. Pechman and P. M. Timpane, pp. 1–14. Washington, D.C.: Brookings Institution, 1975.

Pencavel, John. "Labor Supply of Men: A Survey." In *Handbook of Labor Economics,* vol. 1, edited by Orley Ashenfelter and Richard Layard, pp. 3–102. Amsterdam: North-Holland, 1986.

Prescott, David; Swidinsky, Robert; and Wilton, David. "Labour Supply Estimates for Low-Income Female Heads of Households Using Mincome Data." *Canadian Journal of Economics* 19 (1986): 134–41.

Quebec. *Income Security: Report of the Royal Commission of Inquiry on Health and Social Welfare* (the Castonquay-Nepveu Report). Quebec City: Government of Quebec, 1971.

Robins, Philip. "The Labor Supply Response of Twenty-Year Families in the Denver Income Maintenance Experiment." *Review of Economics and Statistics* 66 (1984): 491–95.

———. "A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments." *Journal of Human Resources* 20 (1985): 567–82.

Robins, Philip, and West, Richard. *Participation in the Seattle and Denver Income Maintenance Experiments and Its Effect on Labor Supply.* Palo Alto, Calif.: Stanford Research Institute, Center for the Study of Welfare Policy, 1978.

———. "Program Participation and Labor-Supply Response." *Journal of Human Resources* 15 (1980): 499–523.

———. "Sample Attrition and Labor Supply Response in Experimental Panel Data." *Journal of Business and Economic Statistics* 4 (1986): 329–38.

Sabourin, Donald. "Participation in Income-tested Social Programs: Evidence from the Mincome Experiment." Winnipeg: University of Manitoba, Institute for Social and Economic Research, 1985.

Smith, D. B. "A Simplified Approach to Social Welfare." *Canadian Tax Journal* 13 (1965): 260–65.

Van Loon, R. "Reforming Welfare in Canada." *Public Policy* 27 (1979): 469–504.

Watts, Harold; Poirier, Dale; and Mallar, Charles. "Sample, Variables, and Concepts Used in the Analysis." In *The New Jersey Income-Maintenance Experiment,* vol. 2: *Labor-Supply Responses,* edited by Harold Watts and Albert Rees. New York: Academic Press, 1976.

Winnipeg Tribune. "Schreyer Outlines Guaranteed Annual Income Test." Interview with Premier Ed Schreyer. (July 10, 1971), p. 16.