EVALUATING THE WELFARE STATE

James J. Heckman
Jeffrey A. Smith

Working Paper **6542**

NBER WORKING PAPER SERIES

EVALUATING THE WELFARE STATE

James J. Heckman
Jeffrey A. Smith

Working Paper 6542
http://www.nber.org/papers/w6542

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1998

Portions of this paper were first presented as the Barcelona Lecture to the World Econometric Society (Heckman, 1990b) and some of this material draws from Heckman (1992), Heckman (1993), Heckman, Smith and Clements (1993, published 1997), and Heckman and Smith (1993, 1995). The ideas not presented in the other papers were first presented in informal discussions at the CEMFI conference in Madrid, Spain, September, 1993, and at a workshop at the Institute for Research on Poverty at the University of Wisconsin, Madison, February, 1994, and have been shared with colleagues at the University of Chicago. The research in this paper was supported by NSF-SBR-93-21-048 and by grants from the Russell Sage Foundation and the American Bar Foundation. We thank the editor of the Econometric Society Monograph Series, Alberto Holly, and an anonymous referee for helpful comments, as well as Gary Becker, Olav Bjerkholt, Dragan Filipovich, Lars Hansen, Hidehiko Ichimura, Lance Lochner, Tom MaCurdy, Derek Neal, Robert Pollak, Jose Scheinkman, Chris Taber, Ed Vytlacil and seminar participants at the Oslo meeting in March, 1995, Columbia University, the University of Chicago, University College London, the Canadian Econometric Study Group in Ottawa in September, 1995, and Washington University, St. Louis. Ed Vytlacil is singled out for a particularly close reading. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Evaluating the Welfare State
James J. Heckman and Jeffrey A. Smith
NBER Working Paper No. 6542
May 1998
JEL Nos. H43, C93, C14

ABSTRACT

A variety of criteria are relevant for evaluating alternative policies in democratic societies composed of persons with diverse values and perspectives. In this paper, we consider alternative criteria for evaluating the welfare state, and the data required to operationalize them. We examine sets of identifying assumptions that bound, or exactly produce, these alternative criteria given the availability of various types of data. We consider the economic questions addressed by two widely-used econometric evaluation estimators and relate them to the requirements of a comprehensive cost-benefit analysis. We present evidence on how the inference from the most commonly used econometric evaluation estimator is modified when the direct costs of a program are fully assessed, including the welfare costs of the taxes required to support the program. Finally, we present evidence of the empirical inconsistency of alternative criteria derived from evaluations based on "objective" outcomes, on self-selection and attrition decisions, and on self-reported evaluations from questionnaires when applied to a prototypical job training program.

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
jimh@cicero.spc.uchicago.edu

Jeffrey A. Smith
Department of Economics
University of Western Ontario
Social Science Centre
London, Ontario N6A 5C2
CANADA
and NBER
smith5@sscl.uwo.ca

Ragnar Frisch was a leading advocate of national economic planning in the service of the welfare state. His Nobel lecture (1970) stressed the value of interactions between economists and politicians in arriving at politically acceptable and economically viable national plans. A major theme of his lecture was that economists should act in the public interest and in so doing should recognize the diversity of policy objectives advocated by different groups in democratic societies. He made the important distinction between maximizing the mythical welfare function assumed in classical welfare economics and in the classical policy analysis of Tinbergen (1956) and reconciling and satisfying the diverse perceptions and values held by citizens of modern states. He stressed the role of economists in informing policymakers and the general public about the relevant economic tradeoffs and the costs of alternative policies.

Frisch's faith in the power of economics to supply the information required to make informed public choices seems wildly optimistic today, yet his message remains relevant. Economists are still asked to inform the general public and policymakers about the likely consequences of alternative social programs. Social welfare functions do not govern decision making in any democratic society and it is clear that a variety of criteria are relevant for evaluating alternative policies in democratic societies composed of individuals and groups with diverse values and perspectives.

Coercive redistribution and intervention are defining activities of the welfare state. Principled redistribution and intervention are based on interpersonal comparisons made by governments and groups of individuals in society. Different public policies typically have different consequences for different citizens. Enumerating and evaluating these consequences is an important part of social decision making, and different criteria have been suggested.

This essay presents these criteria and considers the data required to operationalize them. Some are very difficult to empirically implement and so cannot serve as practical guides to social decision making. Other criteria can be implemented, especially if economists have access to microdata, but such data must be supplemented by knowledge of – or assumptions about – the choice process of the agents being studied or the dependence across potential outcomes associated with different policies and by assumptions that relate partial equilibrium evaluations to general equilibrium evaluations. We examine alternative sets of identifying assumptions that bound, or exactly produce, the alternative criteria used in evaluating the welfare state.

Frisch was well aware that economists often need to supplement the available data with assumptions in order to evaluate policies. His Nobel lecture emphasized the "Lure of Unsolvable Problems" and he advocated evaluation of policies by making bold assumptions if necessary. His famous article on circulation planning emphasizes this point (Frisch, 1934). This essay examines how bold the assumptions have to be to answer major economic evaluation questions given the type of data available in many countries.

This paper considers the general policy evaluation problem but focuses most of its attention on a specific version of it which is widely studied in econometrics under the rubric of "the analysis of treatment effects" (See, *e.g.*, Heckman and Robb, 1985). In this version, a policy has a voluntary component, sometimes called a program, and persons choose to participate in it. A job training program or a tuition subsidy for college attendance are examples. There may be an involuntary component to the policy as well, such as paying the taxes to finance the voluntary component or facing the wages produced by an increase in the supply of trained workers resulting from the program. Having access to data on the outcomes of participants and nonparticipants simplifies some aspects of the policy evaluation problem compared to the case where only participants or nonparticipants are observed, but also raises the problem of self-selection bias.

From information on the program participation decisions of eligible persons, it is sometimes possible to infer their preferences for the outcomes produced by the program. This information is of value in its own right. A strictly libertarian evaluation of a program stops with determining individual subjective valuations for the program being evaluated. Evaluation of the welfare state requires more information. "Objective" evaluations of outcomes supplement revealed preference information to form the basis for policy discussions and interpersonal comparisons. Even if such "objective" information is available about outcomes under each policy, knowledge of individual preferences for specific programs helps in constructing, or bounding, the distribution of outcomes across alternative policy regimes that is required to implement some of the criteria examined in this paper. In addition, if general equilibrium effects can be safely ignored, knowledge of the outcomes of self-selected nonparticipants sometimes identifies the distribution of outcomes for society at large in the absence of the program produced by a policy.

In the course of examining these issues, we consider which policy questions conventional econometric "treatment effect" estimators answer by embedding them in a simple general equilibrium framework. Most of the standard econometric estimators identify parameters of only limited economic interest, but in certain special cases they provide partial answers to economically interesting questions.

We use data from a social experiment designed to evaluate the gross impact of training on the earnings and employment of participants to examine the conflict or consistency among the various criteria that have been proposed to evaluate policies. If all the criteria are in agreement, their multiplicity is not a matter of practical importance. We examine participant evaluations as revealed by their attrition from the program instituted by a policy, by their responses to questionnaires and by econometric analyses of outcome and participation equations under different identifying assumptions. We find that participant self-assessments disagree with revealed preferences as manifested by choices, and with impacts objectively estimated using experimental data. The criteria do not all agree and there is scope for seriously conflicting assessments of a program. We also present evidence that favorable cost-benefit assessments of government training programs are considerably weakened once the full cost of raising government revenue is accounted for.

This paper is organized in the following way. Section 1 presents alternative criteria for evaluating the welfare state and the data required to operationalize them. Both the general case and the specific case of evaluating a program into which agents self-select are examined. Section 2 considers how alternative assumptions about decision processes and access to different sources of micro data aid in the construction of the evaluation criteria. Section 3 examines the economic questions addressed by two widely-used econometric evaluation estimators and relates them to a portion of what is required in a comprehensive cost-benefit analysis. Evidence is presented on how the inference from the most commonly-used econometric evaluation estimator is modified when direct costs of the program are fully assessed, including the welfare costs of taxes raised to support the program. Section 4 presents empirical evidence on the consistency of alternative criteria derived from evaluations based on "objective" outcomes, evaluations inferred from self-selection and attrition decisions and self-reported evaluations elicited from questionnaires on a prototypical job training program. The paper concludes with a summary in Section 5.

1. Alternative Criteria for Evaluating the Welfare State

(A) The Origin of the Demand for Evaluations

Coercive redistribution and intervention are essential activities of the welfare state. Adopting any particular policy with a redistributional component involves weighing the subjective assessments held by different groups of the outcomes created by the policy using the political process. Coercion arises because the perceived benefit from a policy does not always equal or exceed its cost for all members of a society. If there were no coercion, redistribution and intervention would be voluntary activities and, apart from the free rider problem, there would be no scope for government activity in orchestrating redistribution and conducting interventions. There would be no need to publicly justify voluntary trades among individuals.

If government is producing a service for which there are good market substitutes, there is no need to resort to an elaborate evaluation procedure for the service. Market prices provide the right measure of marginal gain and cost unless the usual problems of increasing returns, externalities or public perception that private preferences are defective lead to mistrust of the signals produced from the market mechanism. The argument that justifies the welfare state denies the use of prices and private evaluations as the sole criterion for evaluation of governmental activities, but recognizes that they may be relevant inputs into the general policy evaluation process.

The demand for publically-documented evaluations arises from a demand for information by rival parties in a democratic state. Even libertarians who do not accept coercion and who oppose government intervention evaluate policies in order to participate in the political dialogue of the welfare state.

The claims that markets fail or that consumer judgements are faulty are often made without a factual basis. If these claims are false, the case for a welfare state is weakened. In this paper, we accept the reality of the welfare state, without necessarily endorsing the arguments for it. We do not consider the quality of the evidence supporting the premises of the welfare state. Instead, we consider the evaluation of specific policy proposals within its framework.

(B) Alternative Criteria for Evaluating Programs

Let the outcome in the presence of policy j for person i be Y_{ji} and let the personal preferences of person i for outcome Y be denoted $U_i(Y)$. A policy effects a redistribution from taxpayers to beneficiaries, and Y_{ji} represents the flow of resources to i under policy j. Some persons can be both beneficiaries and tax payers. All policies we consider are assumed to be feasible. In the simplest case, Y_{ji} is net income after tax and transfers, but it may also be a vector of incomes and benefits, including provisions of in-kind services. Many criteria have been proposed to evaluate policies. Let "0" denote the no-policy state, and initially abstract from uncertainty. The standard model of welfare economics postulates a social welfare function W that is defined over the utilities of the N members of society:

(1) $$W(j) = W(U_1(Y_{j1}), ..., U_N(Y_{jN})).$$

Policy choice based on a social welfare function picks that policy j with the highest value for $W(j)$. A leading special case is the Benthamite social welfare function:

(2) $$B(j) = \sum_{i=1}^{N} U_i(Y_{ji}).$$

Criteria (1) and (2) implictly assume that social preferences are defined in terms of the private preferences of citizens as expressed in terms of their own consumption. (This principle is called welfarism. See Sen, 1979.) They could be extended to allow for interdependence across persons so that the utility of person i under policy j is $U_i(Y_{j1}, ..., Y_{jN})$ for all i.

Conventional cost-benefit analysis assumes that Y_j is scalar income and orders policies by their contribution to aggregate income:

(3) $$CB(j) = \sum_{i=1}^{N} Y_{ji}.$$

Analysts who adopt criterion (3) implicitly assume that outputs are costlessly redistributed among persons via a social welfare function, or else accept GNP as their measure of value for a policy.

While these criteria are traditional, they are not universally accepted and they do not answer all of the interesting questions of political economy or "social justice" that arise in the political arena of the welfare state. In a democratic society, politicians and advocacy groups are interested in knowing the proportion of people who benefit from policy j as compared to policy k:

$$PB(j\,|j,k) = \frac{1}{N}\sum_{i=1}^{N} 1(U_i(Y_{ji})) \geq U_i(Y_{ki})), \tag{4}$$

where "1" is the indicator function: $1(A) = 1$ if A is true; $1(A) = 0$ otherwise. In the median voter model, a necessary condition for j to be preferred to k is that $PB(j\,|j,k) \geq 1/2$. Many writers on "social justice" are concerned about the plight of the poor as measured in some base state k. For them, the gain from policy j is measured in terms of the income or utility gains of the poor. In this case, interest centers on the gains to specific types of persons, *e.g.*, the gains to persons with outcomes in the base state k less than $\underline{y}$: $\Delta_{jki} = Y_{ji} - Y_{ki}\,|Y_{ki} \leq \underline{y}$, or their distribution

$$F\left(\Delta_{jk} | Y_k = y_k, y_k \leq \underline{y}\right), \tag{5}$$

or the utility equivalents of these variables. Within a targeted subpopulation, there is sometimes interest in knowing the proportion of people who gain relative to specified values of the base state k:

$$\Pr\left(\Delta_{jk} > 0 | Y_k \leq \underline{y}\right). \tag{6}$$

In addition, measures (2) and (3) are often defined only for a target population and not the full taxpayer population.

The existence of merit goods like education or health implies that specific components of the vector Y_{ji} are of interest to certain groups. Many policies are paternalistic in nature and implicitly assume that people make the wrong choices. "Social" values are placed on specific outcomes, often stated in terms of thresholds. Thus one group may care about another group in terms of whether they satisfy an absolute threshold requirement:

$$Y_{ji} \geq \underline{y} \qquad \text{for } i \in S,$$

where S is a target set toward which the policy is directed, or in terms of a relative requirement compared to a base state k:

$$Y_{ji} \geq Y_{ki} \qquad \text{for } i \in S.$$

Uncertainty introduces additional considerations. Participants in society typically do not know the consequences of each policy for each person. A fundamental limitation in applying these criteria is that, ex ante, these consequences are not known and, ex post, one may not observe all potential outcomes for all persons. If some states are not experienced, the best that agents can do is to guess about them. Even if, ex post, agents know their outcome in a benchmark state, they may not know it ex ante, and they may always be uncertain about what they would have experienced in an alternative state.

In the literature on welfare economics and social choice, one form of decision-making under uncertainty has been extensively investigated. The "Veil of Ignorance" of Vickrey (1945) and Harsanyi (1955; 1975) postulates that decision makers are completely uncertain about their positions in the distribution of outcomes under each policy, or should act as if they are completely uncertain, and they should use expected utility criteria (Vickrey-Harsanyi) or a maximin strategy (Rawls, 1971) to evaluate their welfare under alternative policies. This form of ignorance is sometimes justified as an "ethically correct" position that captures how an "objectively detached" observer should evaluate alternative policies even if actual participants in the political process use other criteria. An approach based on the veil of ignorance is widely used in practical work in evaluating different income distributions (See Sen, 1973), and requires information only about the marginal distributions of outcomes produced under different policies.

A less high-minded, but empirically more accurate, description of social decision making recognizes that persons act in their own self-interest, and have some knowledge about how they will fare under different policies, but allows for the possibility that persons only imperfectly anticipate their outcomes under different policy regimes. The outcomes in different regimes may be dependent so that persons who benefit under one policy may also benefit under another. However, agents may not possess perfect foresight. Letting I_i denote the information set available to agent i, he (she) would evaluate policy j against k using that information. Let $F(y_j, y_k | I_i)$ be the distribution of outcomes (Y_j, Y_k) *as perceived* by agent i. Under an expected utility criterion, person i prefers policy j over k if

$$E(U_i(Y_j) | I_i) > E(U_i(Y_k) | I_i).$$

Letting θ_i parameterize heterogeneity in preferences, so $U_i(Y_j) = U(Y_j; \theta_i)$, and using integrals to simplify the expressions, the proportion of people who prefer j is

(7) $$PB(j|j,k) = \int 1(E(U(Y_j|\theta)|I) > E(U(Y_k|\theta)|I))dF(\theta, I),$$

where $F(\theta, I)$ is the joint distribution of θ and I in the population whose preferences over outcomes are being studied.[1] In the special case where $I_i = (Y_{ji}, Y_{ki})$, so there is no uncertainty about Y_j and Y_k,

(8) $$PB(j|j,k) = \int 1(U(y_j; \theta) > U(y_k; \theta)) dF(\theta, y_j, y_k).$$

Expression (8) is an integral version of (4) when outcomes are perfectly predictable and when preference heterogeneity can be indexed by vector θ.

Adding uncertainty to the analysis makes it fruitful to distinguish between ex ante and ex post evaluations. Ex post, part of the uncertainty about policy outcomes is resolved although individuals do not, in general, have full information about what their potential outcomes would have been in policy regimes they have not experienced and may have only incomplete information about the policy they have experienced (*e.g.* the policy may have long run consequences extending after the point of evaluation). It is useful to index the information set I_i by t, I_{it}, to recognize that information about the outcomes of policies may accrue over time. Ex ante and ex post assessments of a voluntary program need not agree. Ex post assessments of a program through surveys administered to persons who have completed it (see Katz, Gutek, Kahn and Barton, 1975), may disagree with ex ante assessments of the program. Both may reflect honest valuations of the program but

[1] We do not claim that persons would necessarily vote "honestly", although in a binary choice setting they do and there is no scope for strategic manipulation of votes. See Moulin (1983). PB is simply a measure of relative satisfaction and need not describe a voting outcome where other factors come into play.

they are reported when agents have different information about it. Before participating in a program, persons may be uncertain of the consequences of participation. A person who has completed program j may know Y_j but can only guess at the alternative outcome Y_k which they have not experienced. In this case, ex post "satisfaction" for agent i is synonymous with the following inequality:

(9) $$U_i(Y_{ji}) > E(U_i(Y_{ki})\,|I_{it}),$$

where t is the post-program period in which the evaluation is made. In addition, survey questionnaries about "client" satisfaction with a program may capture subjective elements of program experience not captured by "objective" measures of outcomes that usually exclude psychic costs and benefits.

In order to operationalize these notions empirically, it is useful to distinguish the effects of a policy as it impacts the tax collection system from its effects operating through direct program participation. To this end, it is useful to isolate policy outcomes from alternative revenue-neutral programs under the same tax structure and consider the consequences of alternative tax structures separately. In most of the empirical work reported below, we are only able to measure impacts of specific programs on direct participants. We abstract from the tax consequences of financing a program except when we consider the effect of accounting for full social costs in a cost-benefit analysis of a training program. This approach is justified by two distinct arguments: (a) that the tax consequences of the program being evaluated are slight; or (b) that the program's "clients" differ from the taxpayers and we only wish to measure the welfare gains of the "clients".

(C) The Domain of the Microeconomic Literature on Self-Selection and "Treatment Effects"

The classical macroeconomic general equilibrium policy evaluation program considered by Knight (1921), Tinbergen (1956), Marschak (1953) and Lucas and Sargent (1981) forecasts and evaluates the impacts of policies that have never been implemented. To do this requires knowledge of policy-invariant structural parameters and a basis for making proposed new policies comparable to old ones.[2]

The common form of the microeconomic evaluation problem is apparently more tractable. It considers evaluation of a program in which participation is voluntary although it may not have been intended to be so. Persons are offered a service and may select into the program to receive it. Eligibility for the program may be restricted to subsets of persons in the larger society. Many "mandatory" programs have as an option that persons may attrite from them or fail to comply with program requirements. Participation in the program is equated with direct receipt of the service and payments of taxes and general equilibrium effects of the program are ignored.[3]

In this formulation of the evaluation problem, the no-treatment outcome distribution for a given program is used to approximate the distribution of outcomes in the no-program state. That is, the outcomes of the untreated within the context of an existing program are used to approximate outcome distributions when there is no program. This approximation rests on two distinct arguments: (a) that general equilibrium effects inclusive of taxes and spillover effects on factor and output markets can be ignored; and (b) that the problem of selection bias that arises from using self-selected samples of participants and nonparticipants to estimate population distributions can be ignored or surmounted.

More precisely, let j be the policy regime we seek to evaluate. Eligible person i in regime j has two potential outcomes: (Y_{ji}^0, Y_{ji}^1), where the superscripts denote nondirect participation ("0") and direct participation ("1"). Noneligible persons have only one option: Y_{ji}^0. These outcomes are defined at the equilibrium level of participation under program j. All feedback effects are incorporated in the definitions of the potential outcomes.

Let subscript "0" denote a policy regime without the program. Let $D_{ji} = 1$ if person

[2] A quotation from Knight is apt "...The existence of a problem in knowledge depends on the future being different from the past, while the possibility of a solution of the problem depends on the future being like the past". (Knight, 1921, p. 313.)

[3] The contrast between micro and macro analysis is overdrawn. Baumol and Quandt (1966), Lancaster (1971) and Domencich and McFadden (1975) are micro examples of attempts to solve what we have called a macro problem. Those authors consider the problem of forecasting the demand for a new good which has never previously been purchased.

i participates in program j. A crucial identifying assumption that is implicitly invoked in the microeconomic evaluation literature is

(A-1) $$Y_{ji}^0 = Y_{0i},$$

and hence that $F(y_j^0 | D_j = 0, X) = F(y_0 | D_j = 0, X)$ for $y_j^0 = y_0$ given conditioning variables X. The outcome of nonparticipants in policy regime j is the same in the no policy state "0" or in the state where policy j is operative. This assumption is consistent with a program that has "negligible" general equilibrium effects and where the same structure of tax revenue collection is used in regimes j and "0".

An additional assumption sometimes invoked is that

(A-2) $$Y_{ji}^1 = Y_{ji},$$

where Y_{ji} is the outcome if the program is universally applied. This entails a different kind of general equilibrium assumption – this time about expansion of program j to universal coverage. Making (A-1) and (A-2) together strains the imagination, for if a program is small enough that (A-1) is plausible, its universal expansion may make it so large that (A-2) is not plausible. Nonetheless, taken together, these assumptions strengthened with additional assumptions about agent self-selection rules, enable analysts to generalize from self-selected samples within a given policy regime to choices across policy regimes. Assumption (A-2) is rarely used and plays only a minor role in this paper. Assumption (A-1) plays a much more substantial role in this paper and in the microeconomic evaluation literature.

From data on individual program participation decisions, it is possible to infer the implicit valuations of the program made by persons eligible for it. These evaluations constitute all of the data needed for a libertarian program evaluation, but more than these are required to evaluate programs in the interventionist welfare state. For certain decision rules, it is possible to use the data from self-selected samples to bound or estimate the joint distributions required to implement criteria (4) or (7), as we demonstrate below.

The existence of a voluntary participation component for a program under policy j creates an option value which for eligible person i is

(10) $$Max\{Y_{ji}^0, Y_{ji}^1\} - Y_{ji}^0.$$

By (A-1) this is the same as $Max\{Y_{0i}, Y_{ji}^1\} - Y_{0i}$, which if strengthened by (A-2) is $Max\{Y_{0i}, Y_{ji}\} - Y_{0i}$. The distribution of the value of this option for those who take it is

$$F(y_j^1 - y_j^0 \,\big|\, Y_j^1 > Y_j^0). \tag{11}$$

For persons interested in the equity of program provisions, it is of interest to examine the dependence between the options offered and the non-participation outcomes, which are assumed to approximate the no-policy outcomes.

People who fear "cream skimming" by program administrators whose performance is evaluated on the basis of the outcomes of the participants they select, claim that Y_j^1 and Y_j^0 are strongly positively dependent and that the gross value added, $\Delta_j = Y_j^1 - Y_j^0$, is unrelated or negatively related to Y_j^0. To address these concerns, it is necessary to know the joint dependence between Y_j^0 and Y_j^1 and compute the dependence between Δ_j and Y_j^0.

(2) The Data Needed to Evaluate the Welfare State

To implement criteria (1) and (2), it is necessary to know the distribution of outcomes across the entire population and to know the utility functions of individuals. In the case where Y refers to scalar income, criterion (3) only requires GNP (the sum of the program j outcome distribution). If interest centers solely on the distributions of outcomes of direct program participants, the measures can be defined solely for populations with $D_j = 1$. Criteria (4), (5), (6) and (8) require knowledge of outcomes and preferences across programs. Criterion (7) requires knowledge of the joint distribution of information and preferences across persons. Tables 1A and 1B summarize the criteria and the data needed to implement them.

This paper has little to say about estimating preference functions or preference heterogeneity. We refer readers to Heckman (1974a) and the comprehensive survey by Browning, Hansen, Heckman and Taber (1997), who document the empirical importance of preference heterogeneity. Our focus is on estimating the distributions of outcomes across policy states as a first step toward empirically implementing the full criteria. This more modest objective can be fit into the framework of Section 1 by assuming that utilities are linear in their arguments and identical across persons.

This section considers the problem of constructing the distribution of (Y_j^0, Y_j^1), the distribution of potential outcomes within policy regime j in which direct participation is voluntary. Extension of the estimates of this distribution to other policy regimes follows by invoking the assumptions discussed in Section 1. We discuss how the widely-invoked implicit assumption that responses to program treatment are homogeneous across persons greatly simplifies the construction of the joint distribution of potential outcomes and how explicit assumptions about the structure of voluntary program participation rules aids in identifying or reducing the uncertainty about the distributions of outcomes. We consider the information available from cross section data, from social experiments, from panel data and from repeated cross section data.

The Microeconomic Evaluation Problem

To simplify the notation, we drop the policy regime subscript j. All of the distributions we consider in this section are measured within that regime. The extrapolation of within-regime measures to across-regime measures is made using assumptions (A-1) and (A-2) discussed in Section 1. In a regime with voluntary participation, we have access to

$$F(y_t^1 | D = 1, X) \text{ and } F(y_t^0 | D = 0, X), \tag{12}$$

the distributions of outcomes for participants and nonparticipants at time t, respectively. These embody both the direct and indirect effects of the program.

The fundamental evaluation problem arises from the fact that we do not observe (Y_t^0, Y_t^1) for anyone - just one coordinate or the other of this pair. Given knowledge of individual preferences, and their joint distribution with the outcomes, all of the policy criteria discussed in Section 1 and summarized in Tables 1A and 1B can be implemented. Here we focus on recovering $F(y_t^0, y_t^1, D|X)$, from which all of the distributions discussed in Section 1 can be recovered. For evaluating the criteria only for program participants, it is enough to know $F(y_t^0, y_t^1|D = 1, X)$ – the potential outcomes for participants – or various marginal distributions formed from this distribution.

As previously noted, the different evaluation criteria require different data for their empirical implementation. Cost-benefit analysis can in principle be performed using a before-after analysis on aggregate time series data. However, if a program has a small impact on the economy and other policies are instituted coincident with the program being evaluated, or if the time series is nonstationary, aggregate data are not a reliable source of information.

The missing counterfactual for cost-benefit criterion (3) is the mean $E(Y_t^0)$, or $E(Y_t^0 \,|D = 1, X)$ if the evaluation is conducted solely for participants. $E(Y_t^1 \,|D = 1, X)$ is produced from data on program participants. Benthamite criterion (2) is more demanding and requires $F(y_t^0 \,|X)$, or $F(y_t^0 \,|D = 1, X)$ if the criterion is defined only for participants. The voting criterion (8) requires $F(y_t^1, y_t^0 \,|X)$, or $F(y_t^1, y_t^0 \,|D = 1, X)$ if the criterion is defined only for participants.

In this section, we consider how to use cross section data, data from ideal social experiments, panel data and repeated cross section data to construct the different evaluation criteria. Panel data can be used as repeated cross sections and repeated cross sections can be used as cross sections. Thus it is natural to start with the cross section case and then determine how access to other sources of data aids in securing identification of the evaluation criteria presented in Section 1.

A. Cross Section Data

From cross section data on $F(y_t^1|D = 1, X)$, $F(y_t^0|D = 0, X)$ and $\Pr(D = 1 \,|X)$, we cannot directly construct the joint distribution, $F(y_t^1, y_t^0, D \,|X)$. Using $F(y_t^0|D = 0, X)$ to proxy $F(y_t^0|D = 1, X)$ runs the risk of selection bias. A variety of different identifying assumptions have been used to recover the counterfactual distribution $F(y_t^0|D = 1, X)$ or the joint distribution $F(y_t^1, y_t^0, D \,|X)$. To simplify the notation in this subsection, we drop the "t" subscript, and assume that (Y^0, Y^1) are measured after the program intervention.

(i) Conditional Independence

One assumption that underlies the *method of matching* postulates conditioning variables X such that

(I-1a) $$F(y^0|D=1,X) = F(y^0|D=0,X) = F(y^0 \mid X).$$

If this assumption is valid, we can safely use nonparticipants to measure what participants would have earned had they not participated, provided we condition on X. Using " $\perp\!\!\!\perp$" to denote independence, this identifying assumption is equivalent to $Y^0 \perp\!\!\!\perp D|X$. To ensure that (I-1a) has an empirical counterpart, we also assume that

(I-1b) $$0 < \Pr(D=1\,|X) < 1$$

over the support of X. This condition assures that both sides of (I-1a) are well defined, *i.e.*, that for each X, there are both participants and nonparticipants.[4] For computing counterfactual means, a simpler requirement is:

(I-2) $$E(Y^0|D=1,X) = E(Y^0|D=0,X).$$

This method underlies the intuitive principle of "controlling on observables" to eliminate selection bias. (Heckman and Robb, 1985.)

Identification assumption (I-1a) implies that $\Pr(D=1|X,Y^0) = \Pr(D=1\,|X)$, *i.e.*, that Y^0 does not determine participation in the program, although it does not exclude the possibility that participation in the program is based on Y^1. If we strengthen (I-1a) to read

(I-3) $$(Y^0,Y^1) \perp\!\!\!\perp D|X,$$

we can recover $F(y^1\,|\,X)$ for the support of X satisfying (I-1b). Thus for the entire population or for the sample conditional on $D=1$, we can construct the cost-benefit criterion and the Benthamite criterion but not the voting criterion, because there is no information on the joint distribution of (Y^0,Y^1).

To recover the joint distribution, we need some way to associate values of Y^0 with Y^1. The dummy endogenous variables model (Heckman, 1978) assumes that

[4]Failure to satisfy this condition is an important source of failure in the use of matching to evaluate job training programs. See Heckman, Ichimura, Smith and Todd (1994, 1996).

(I-4) $$Y^1 = \alpha + Y^0,$$

where α is a constant or a function of X. Defining "α" as the treatment effect, this assumption imposes homogeneity of response to treatment. Everyone with the same X value benefits or loses by the same amount. A generalization of this method developed in Heckman and Smith (1993) and Heckman, Smith and Clements (1997) assumes that the quantiles of Y^1 and Y^0 are the same for each person with the same X. Equating quantiles across the two marginal distributions we form pairs

(13) $$\{(y^0(q), y^1(q)) | \inf_{y^0} F(y^0|X) > q \text{ and } \inf_{y^1} F(y^1|X) > q, 0 \leq q \leq 1\}.$$

Conditional on X, the quantile ranks are preserved, but the effect of treatment is not necessarily the same at all quantiles. More generally, we could assume that the quantiles are mapped in a general way $q_1 = \varphi(q_0)$, where q_1 is a quantile of $Y^1$and q_0 is a quantile of Y^0. The gain to moving from "0" to "1"is

(14) $$\Delta(q_0) = Y^1(\varphi(q_0)) - Y^0(q_0),$$

where $Y_1(\varphi(q_0))$ is the q_1^{th} quantile of Y_1 expressed as a function of q_0 and $Y_0(q_0)$ is the q_0^{th} quantile of Y_0.

If φ is a random function, then the mass at q_0 is distributed to different values of q_1 and $\varphi(q_0)$ has an interpretation as a probability density. If φ is a uniform density mapping q_0 to q_1 over the interval $[0, 100]$ for all q_0, Y^1 and Y^0 are stochastically independent. Provided the mapping φ is known, the assumption of conditional independence is sufficient to identify the joint distribution $F(y^1, y^0, D \mid X)$.[5]

An alternative assumption about the dependence across outcomes is that $Y^1 = Y^0 + \Delta$, where Δ is stochastically independent of Y^0 given X, *i.e.*,

(I-5) $$Y^0 \perp\!\!\!\perp \Delta \mid X.$$

This assumption states that the gain from participating in the program is independent of the base Y^0. If (I-3) and (I-5) are invoked jointly, we can identify $F(y^0, y^1 \mid X)$ from the cross section outcome distributions of participants and non-participants and estimate the

[5]When φ is random, and the random variables are discrete, the matrix mapping probability of Y^0 into Y^1 must be a Markov matrix to preserve probability. For continuous distributions we need a Markov operator.

joint distribution by deconvolution methods.[6]

To see how to use this information, note that

$$Y = Y^0 + D\Delta.$$

From $F(y \mid X, D = 0)$, we identify $F(y^0 \mid X)$ as a consequence of (I-3). From $F(y \mid X, D = 1)$ we identify $F(y^1 \mid X) = F(y^0 + \Delta \mid X)$. If Y^0 and Y^1 have densities then, as a consequence of (I-5), the densities satisfy

$$f_1(y^1 \mid X) = f_\Delta(\Delta \mid X) * f_0(y^0 \mid X)$$

where "$*$" denotes convolution. The characteristic functions of the three random variables satisfy

$$E(e^{i\ell Y^1} \mid X) = E(e^{i\ell\Delta} \mid X)E(e^{i\ell Y^0} \mid X).$$

Since we can identify $F(y^1 \mid X)$, we know its characteristic function. By a similar argument we can recover $E(e^{i\ell Y^0} \mid X)$. Then

$$E(e^{i\ell\Delta} \mid X) = \frac{E(e^{i\ell Y^1} \mid X)}{E(e^{i\ell Y^0} \mid X)}, \tag{15}$$

and by the inversion theorem, (see, *e.g.*, Kendall and Stuart, 1977), we can recover the density $f_\Delta(\Delta \mid X)$. We know the joint density

$$f(\Delta, y_0 \mid X) = f_\Delta(\Delta \mid X) f(y^0 \mid X) .$$

From the definition of Δ, we obtain

$$f(y^1 - y^0 \mid X) f(y^0 \mid X) = f(y^1, y^0 \mid X).$$

[6] Barros (1987) uses this assumption in the context of an analysis of selection bias.

Thus we can recover the full joint distribution of outcomes and the distribution of gains.

Under assumption (I-3), assumption (I-5) is testable. The ratio of two characteristic functions in (15) is not necessarily a characteristic function. If it is not, the estimated density f_Δ recovered from the ratio of the characteristic functions need not be positive and the estimated variance of Δ can be negative.[7]

In a regression setting in which means and variances are assumed to capture all of the relevant information, this approach is equivalent to the traditional normal random coefficient model. Letting

$$\begin{aligned} Y^1 &= \mu_1(X) + U_1 \qquad & E(U_1 \,|X\,) = 0 \\ Y^0 &= \mu_0(X) + U_0 \qquad & E(U_0 \,|X\,) = 0, \end{aligned}$$

this version of the model may be written as

$$\begin{aligned} (16) \qquad Y &= \mu_0(X) + D(\mu_1(X) - \mu_0(X) + U_1 - U_0) + U_0 \\ &= \mu_0(X) + D(\mu_1(X) - \mu_0(X)) + D(U_1 - U_0) + U_0 \\ &= \mu_0(X) + D\bar{\alpha}(X) + D\varepsilon + U_0, \end{aligned}$$

where $\bar{\alpha}(X) = \mu_1(X) - \mu_0(X)$ and $\varepsilon = U_1 - U_0$. By virture of (I-3), $(U_0, U_1) \perp\!\!\!\perp D\,|X$.

We may use nonparametric regression methods to recover $\mu_0(X)$ and $\mu_1(X) - \mu_0(X)$ or we may use ordinary parametric regression methods assuming that $\mu_1(X) = X\beta_1$ and $\mu_0(X) = X\beta_0$. Equation (16) is a components of variance model and a test of (I-5) is that

$$\begin{aligned} Var(Y\,|D=1,X\,) &= Var(Y^0 + \Delta\,|D=1,X) = Var(Y^0\,|X\,) + Var(\Delta\,|X) \\ &\geq Var(Y\,|D=0,X) = Var(Y^0\,|X) \,. \end{aligned}$$

Under standard conditions each component of variance is identified and estimable from the residuals obtained from the nonparametric regression of Y on D and X.

An alternative approach relies only on the information contained in the marginal distributions obtained using the conditional independence assumption to bound the joint distribution conditional on D. The Frechet (1951) bounds inform us that

$$(17) \; Max\{F(y^0\,|X\,) + F(y^1\,|X\,) - 1, 0\} \leq F(y^0, y^1\,|X\,) \leq Min\{F(y^0\,|X\,), F(y^1\,|X\,)\}.$$

[7]For the ratio of characteristic functions, $r(\ell)$, to be a characteristic function, it must satisfy the requirement that $r(0) = 1$, that $r(\ell)$ is continuous in ℓ and $r(\ell)$ is nonnegative definite. This identifying assumption can be tested using the procedures developed in Heckman, Robb and Walker (1990).

These bounds are purely statistical and assume no information about agent behavior. Combining the bounds with (I-1b) and (I-3) allows us to bound the ($D = 1$) joint distribution $F(y^1, y^0 \,|X)$. Heckman and Smith (1993) and Heckman, Smith and Clements (1997) demonstrate that in most applications these bounds are not very informative.[8]

(ii) Information From Revealed Preference

An alternative approach with a long history in economics uses information on agent choices to recover the population distribution of potential outcomes.[9] Unlike the method of matching, the method based on revealed preference capitalizes on a close relationship between (Y^0, Y^1) and program participation. Participation includes voluntary entry into a program or attrition from it.

The prototypical framework is the Roy (1951) model. In that setup,

$$D = 1(Y^1 \geq Y^0). \qquad (18)$$

If we postulate that the outcome equations can be written in a separable form, so that

$$\begin{aligned} Y^1 &= \mu_1(X) + U_1 \qquad & E(U_1 \,|X) = 0 \\ Y^0 &= \mu_0(X) + U_0 \qquad & E(U_0 \,|X) = 0, \end{aligned}$$

then $Pr(D = 1 \,|X) = \Pr(Y^1 - Y^0 \geq 0 \,|X) = \Pr(U_1 - U_0 \geq -(\mu_1(X) - \mu_0(X)))$. Heckman and Honoré (1990) demonstrate that if $X \perp\!\!\!\perp (U_1, U_0)$, $Var(U_1) < \infty$ and $Var(U_0) < \infty$, and (U_1, U_0) are normal, the full model $F(y^0, y^1, D \,|X)$ is identified even if we only observe Y^0 or Y^1 for any person and there are no regressors and no exclusion restrictions. If instead of assuming normality, we assume that the supports of $\mu_1(X)$ and $\mu_0(X)$ overlap or contain the supports of U^1 and U^0, the full model $(\mu_1(X), \mu_0(X))$ and the joint distribution of

[8] An exception is that the bounds for small probability events are informative.

[9] Heckman (1974a,b) demonstrates how access to censored samples on hours of work, wages for workers, and employment choices identifies the joint distribution of the value of nonmarket time and potential market wages under a normality assumption. Heckman and Honoré (1990) consider nonparametric versions of this model without labor supply.

U_1, U_0 are nonparametrically identified up to location normalizations. Precise conditions are given in Theorem A-1 in Appendix A.

The crucial feature of the Roy model is that the decision to participate in the program is made solely in terms of potential outcomes. No new unobservable variables enter the model that do not appear in the outcome equations.[10] In this case, information about who participates also informs us about the distribution of the value of the program to participants $F(y^1 - y^0 \,|Y^1 > Y^0, X)$. Thus, we acquire the distribution of implicit values of the program for participants, which is all that is required in a libertarian evaluation of the program. However, as we have stressed repeatedly, evaluation of the welfare state requires information about "objective" outcomes and their distributions that are needed to make the interpersonal comparisons that are an essential feature of the welfare state. Only in the Roy model do the "objective" and "subjective" outcomes coincide.

If the Roy model is extended to allow for variables other than Y^0, Y^1 (and the observed conditioning variables) to determine participation, then the decision rule is changed to $D = 1(IN > 0)$ where $IN = \eta(Y^1, Y^0, V, X)$, and it is not possible to identify the joint distribution $F(u_0, u_1)$ even if the unobservables V, U_0 and U_1 are independent of X. Under conditions similar to those presented in Theorem A-1, Heckman (1990a) demonstrates that in this more general case, provided that some structure is placed on η, we can nonparametrically identify $F(y^0, D\,|X)$ and $F(y^1, D\,|X)$ but not the full joint distribution $F(y^0, y^1, D\,|X)$. A generalization of his proof is given in Theorem A-2 of Appendix A. As soon as the tight link in the Roy model between participation and potential outcomes is broken, we confront the standard evaluation problem that failure to observe both coordinates of (Y^0, Y^1) precludes identification of their joint distribution. To identify the full joint distribution of potential outcomes, we can assume the same dependence across quantiles as

[10]We could augment decision rule (18) to be $D = 1(Y^1 - Y^0 - k(Z) \geq 0)$. Provided that we measure Z and condition on it, and provided that $(U_1 - U_0) \perp\!\!\!\perp (X, Z)$, the model remains nonparametrically identified. The crucial property of the identification result is that no new unobservable enters the model through the participation equation. However, if we add Z, subjective valuations of gain $(Y^1 - Y^0 - k(Z))$ no longer equal "objective" measures $(Y^1 - Y^0)$.

was previously discussed in the case of econometric matching methods or we can apply the Frechet inequalities to bound the joint distributions from the nonparametrically-determined marginals.

Thus far we have considered the case that in advance of participating in a program, persons know their own (Y^0, Y^1). If decision rule (18) is operative in the participant population, this implies that

$$\Pr(Y^1 \geq Y^0 \,|Y^0 = y^0, D = 1, X) = 1.$$

This is a strong form of stochastic dominance. All of the mass of the Y^1 distribution conditional on $Y^0 = y^0$ is to the right of y^0 in the participant population.

More generally, persons may not know (Y^0, Y^1) but may base their participation decisions on unbiased guesses (Y^{0*}, Y^{1*}) about them. We can model this in the following way:

$$Y^{0*} = Y^0 + \varepsilon_0 \text{ and } Y^{1*} = Y^1 + \varepsilon_1,$$

where $E(\varepsilon_0, \varepsilon_1) = (0, 0)$, $(\varepsilon_0, \varepsilon_1) \perp\!\!\!\perp (Y^0, Y^1)$, and $\varepsilon_0 \perp\!\!\!\perp \varepsilon_1$.

In this case, if $D = 1(Y^{1*} > Y^{0*})$, conditioning on realized values produces positive regression dependence between Y^1 and Y^0 so that $\Pr(Y^1 \leq y^1 \,|Y^0 = y^0, D = 1, X)$ is non-increasing in y^0 for all y^1. This in turn implies that Y^1 is right-tail increasing in Y^0. That is, $\Pr(Y^1 > y^1 \,|Y^0 > y^0, D = 1, X)$ is non-decreasing in y^0 for all y^1. Intuitively, the higher the value of y^0, the more the mass in the conditional Y^1 distribution is shifted to the right so that "high values of Y^0 go with high values of Y^1". Y^1 being right tail increasing given y^0 implies that Y^1 and Y^0 (given $D = 1$) are positive quadrant dependent, so that $\Pr(Y^1 \leq y^1 \,|Y^0 \leq y^0, D = 1) \geq \Pr(Y^1 \leq y^1 \,|D = 1, X)$ and $\Pr(Y^0 \leq y^0 \,|Y^1 \leq y^1, D = 1, X) \geq \Pr(Y^0 \leq y^0 \,|D = 1, X)$.[11] Common measures of dependence like the product-moment correlation, Kendall's τ and Spearman's ρ are all positive when there is

[11] These implications are strict except in the case where Y^0 and Y^1 are binary random variables. In this case, Tong (1980) shows that these notions of dependence are all equivalent.

positive quadrant dependence. Even under imperfect information, rationality in the form considered here can restrict the nature of the dependence between Y^1 and Y^0 given $D = 1$. Evidence against such dependence is evidence against the income-maximizing Roy model. Even if Y^0 and Y^1 are negatively correlated in the population, they are positively correlated given $D = 1$ if agents are income maximizers. This insight motivates our imposition of positive dependence between Y^0 and Y^1 in participant populations ($D = 1$) to recover the joint distribution $F(y^0, y^1 \mid D = 1, X)$ in the empirical analysis reported in Section IV.

(iii) Identification Through the Instrumental Variable Moment Condition and Extensions of the Condition

Taking (16) as a point of departure, it is possible under conditions we now specify to apply the method of instrumental variables to estimate $E(Y^1 - Y^0 \mid D = 1, X)$ and $E(Y^1 - Y^0 \mid X) = E(\Delta \mid X)$. This allows implementation of the cost-benefit criterion provided that instrumental variables Z exist that satisfy the following conditions:

(I-6a) $\quad E(U_0 + D(U_1 - U_0) \mid X, Z) = 0$
for identifying $E(Y^1 - Y^0 \mid X)$, or

(I-6b) $\quad E(U_0 + D(U_1 - U_0 - E[U_1 - U_0 \mid D = 1, X]) \mid X, Z) = 0$
for identifying $E(Y^1 - Y^0 \mid X, D = 1)$.

A second condition is that D depends on Z:

(I-7) $\Pr(D = 1 \mid X, Z = z) \neq \Pr(D = 1 \mid X, Z = z')$ for some $z \neq z'$ for all X.

Under condition (I-6a) we may write

$$E(Y \mid X, Z) = \mu_0(X) + E(\Delta \mid X) \Pr(D = 1 \mid X, Z) .$$

Or, under condition (I-6b) we obtain

$$E(Y \mid X, Z) = \mu_0(X) + E(\Delta \mid D = 1, X) \Pr(D = 1 \mid X, Z).$$

Thus the population moment equation that identifies $E(\Delta \mid X)$ under (I-6a) and (I-7) is

(19) $$E(\Delta \mid X) = \frac{E(Y \mid X, Z = z) - E(Y \mid X, Z = z')}{\Pr(D = 1 \mid X, Z = z) - \Pr(D = 1 \mid X, Z = z')}$$

and the population moment equation that identifies $E(\Delta \mid X, D = 1)$ under (I-6b) and (I-7) is the same:

(20) $$E(\Delta \mid X,\ D = 1) = \frac{E(Y \mid X, Z = z) - E(Y \mid X, Z = z')}{\Pr(D = 1 \mid X, Z = z) - \Pr(D = 1 \mid X, Z = z')}.$$

To satisfy condition (I-6b), it is required that a standard instrumental variable condition be satisfied: $E(U_0 \mid X, Z) = 0$ and in addition that

$$E(U_1 - U_0 \mid D = 1, X, Z) = E(U_1 - U_0 \mid D = 1, X)\ .$$

Notice that condition (I-6b) is still satisfied if $U_1 \equiv U_0$ (so the response to treatment in (16) is the same for everyone as assumed in Heckman (1978)). This condition is also satisfied if

(I-8) $$(U_1 - U_0) \perp\!\!\!\perp (X, Z, D).$$

As a consequence of (I-8),

$$\Pr(D = 1 \mid X, Z\ , Y^1 - Y^0) = \Pr(D = 1 \mid X, Z\ , U^1 - U^0) = \Pr(D = 1 \mid X, Z\ ).$$

Identifying assumption (I-8) would be satisfied if agents cannot predict $(U_1 - U_0)$ at the time they make their decisions to participate in the program but they know X and Z. Condition (I-6b) would also be satisfied if (I-8) is weakened to a statement about mean independence:

(I-8)′ $$E(U_1 - U_0 \mid X, Z, D = 1) = E(U_1 - U_0),$$

which would be satisfied if the unobserved components of the gain do not determine program participation.[12] Condition (I-8) does not rule out that Y^0 determines D but if it does, it is required that given X, Z and Y^0, Δ does not determine D.

Under condition (I-8), $E(Y^1 - Y^0 \mid D = 1,\ X) = E(Y^1 - Y^0 \mid X\ )$. The effect of "treatment on the treated" is the same as the effect of taking a person from the population

[12]See Heckman (1997a) for further discussion of these conditions.

at random and assigning that person to treatment. Moreover, (I-8) ensures that (I-6a) is satisfied as well.

Notice that condition (19) is still satisfied if (I-6a) is weakened to

(I-6a)′ $$E(U^0 + D(U_1 - U_0)\,|X, Z) = M_1(X)$$

and condition (20) still holds if (I-6b) is weakened to

$$E(U_0 + D(U_1 - U_0 - E(U_1 - U_0\,|D = 1, X))\,|X, Z) = M_2(X).$$

The $M_1(X)$ and $M_2(X)$ terms difference out in the instrumental variable moment conditions (19) and (20) respectively.

Invoking (I-6) and (I-7) under assumption (A-1) and (A-2), we can answer the cost-benefit questions for the entire population (if we assume (I-6a)) and for populations for which $D = 1$ (if we assume (I-6b)). These assumptions are not strong enough to identify the Benthamite or the voting criteria. To recover the full joint distribution of (U_1, U_0, D) requires strengthening these assumptions. The conditional independence assumption that justifies matching (I-3) would suffice.

Thus in place of (I-3), which is defined solely in terms of variables X in the outcome equations, we may assume that access to a variable Z produces conditional independence:

(I-9) $$(U_0, U_1) \perp\!\!\!\perp D \mid X, Z$$
$$\text{but } (U_0, U_1) \not\perp\!\!\!\perp D \mid X.$$

Equivalently, we may write

(I-9)′ $$(Y_0, Y_1) \perp\!\!\!\perp D \mid X, Z$$
$$\text{but } (Y_0, Y_1) \not\perp\!\!\!\perp D \mid X.$$

Under these assumptions we may recover the marginal and joint distributions as discussed in the subsection on conditional independence. Interpreted in this way the instrumental variables method generalizes the matching method and extends the identification analysis based on conditional independence in terms of variables in the outcome equation to utilize a larger conditioning set beyond those variables.[13]

[13] Heckman, Ichimura and Todd (1997a,b) and Heckman, Ichimura, Smith and Todd (1994, 1996) extend matching to consider variables in the program participation equation that are not in the outcome equation.

B. Social Experiments

We consider randomization administered at two different points: (a) at entry or the stage where persons have applied and been accepted into a program; and (b) at eligibility. As noted in Heckman (1992) and Heckman and Smith (1993, 1995), social experiments with randomization administered at the stage where persons have applied and been accepted into a program recover two marginal distributions conditional on $D = 1$:

(21) $$F(y^1\,|D = 1, X) \quad \text{and} \quad F(y^0\,|D = 1\,, X).$$

From such an experiment, we obtain a truncated sample and experiments administered at this stage do not identify $\Pr(D = 1\,|X)$. (See Heckman, 1992, and Moffitt, 1992) The identifying assumptions that justify this method are:

(I-10) Randomization does not change the program being studied (no randomization bias) and no close substitutes for the treatment are available to persons randomized out (no substitution bias).

Heckman (1992) and Heckman and Smith (1993) discuss the need for the absence of substitutes for the program being evaluated and the failure of the no randomization bias assumption. Heckman, Hohmann, Khoo and Smith (1997) provide evidence on the importance of substitution bias in an evaluation of a major job training program.

From the conditional distributions, it is possible to recover the information required to construct the participant versions of the cost-benefit criterion,

$$E(Y^1 - Y^0\,|D = 1, X)\ ,$$

and the Benthamite criterion. Without further assumptions, social experiments do not recover the conditional distribution

(22) $$F(y^0, y^1\,|D = 1, X).$$

Any one of several additional assumptions can be used to supplement the information available from social experiments. The joint distribution (22) can be bounded from the experimentally determined marginals using the Frechet bounds (Heckman and Smith, 1993; Heckman, Smith and Clements, 1997). Assumptions can be made about the association of quantile ranks (dependence) between outcomes across distributions to recover $F(y^0, y^1 \mid D = 1, X)$. An alternative assumption is (I-5).

With these assumptions, we can construct or bound all of evaluation criteria presented in Section 1 for the conditional (on $D = 1$) distribution. Under conditional independence assumption (I-3), it is possible to recover the complete marginal distributions $F(y^1 \mid D = 1, X) = F(y^1 \mid X)$ and $F(y^0 \mid D = 1, X) = F(y^0 \mid X)$ and bound $F(y^0, y^1 \mid X)$ using the Frechet bounds; or to identify $F(y^0, y^1 \mid X)$ by (a) making an assumption connecting the quantiles of the two marginal distributions or (b) assuming as in (I-5) that gains Δ are unrelated to the base state Y^0.

If decision rule (18) is postulated, we may use the Roy model (under the conditions specified in Theorem A-1) to identify $F(y^0, y^1 \mid X)$ from the conditional distributions $F(y^0 \mid D = 1, X)$ and $F(y^1 \mid D = 1, X)$. Under assumptions (A-1) and (A-2), we can answer the evaluation questions comparing policy j with policy "0" that were posed in Section 1 for the entire population and the conditional population.

Under more general participation rules, we may apply Theorem A-2 to data from a social experiment with randomization administered at the point of entry into the program to identify $F(y^1, D \mid X)$ and $F(y^0, D \mid X)$ for both $D = 1$ and $D = 0$. Thus we can construct the cost-benefit and Benthamite criteria for the general population and for the participant populations, but not the general voting criterion or any other criterion requiring the joint distribution of outcomes.

One advantage of social experiments over conventional micro data augmented with the conditional independence condition (I-3) is that experiments expand the range of the support over which the parameters can be identified. Thus, suppose that $Support(X \mid D = 1) \neq Support(X \mid D = 0)$. For the domains of X where there is no common support, Theorems A-1 and A-2 do not apply and we cannot use conditional independence assumption (I-3). Randomization guarantees that in the population generating the experimental samples $Support(X\,|D = 1)$ is the same for participants and randomized-out persons. Thus randomization ensures that the support conditions of Theorems A-1 and A-2 are satisfied for the population of participants. However, it may still happen that the support of X for the population for which $D = 1$ is not the same as the support of X for the whole population. Then even with experimental data, the parameters of interest are only identified over the available support. For both experimental and non-experimental data, it may be necessary to sample more widely on X coordinates to recover parameters defined for the entire population. Experiments have the advantage that they allow identification of impacts even for persons with values of X such that $\Pr(D = 1\,|X) = 1$, which is not possible using non-experimental methods because there is no comparison group.

If randomization is performed on eligibility for the program, we recover $F(y^0\,|X)$, $F(y^1\,|D = 1\,, X)$ and $F(y^0\,|D = 1\,, X)$. (See Heckman 1992 and Heckman and Smith, 1993). In addition, we recover $\Pr(D = 1\,|X)$, at least for those values of X possessed by eligible persons. Many would regard $F(y^0\,|X)$ as a better approximation to the no-policy outcome distribution than the approximation embodied in assumption (A-1). Although both approximations ignore general equilibrium effects, $F(y^0\,|X)$ avoids self-selection bias. Randomization at eligibility does not recover the full joint distribution of outcomes unless additional assumptions of the type previously discussed are invoked. Table 3 summarizes the information obtained from the two types of experiments.

C. Panel Data

Panel data provide a new source of identifying information. Participation or nonparticipation outcomes in one period can proxy participation or nonparticipation outcomes in another period. Restoring the t subscript, panel data allow us to make the following approximations for person i:

(I-11a) $$Y^1_{t'i} \doteq Y^1_{ti} \qquad t \neq t'$$

or

(I-11b) $$Y^0_{t'i} \doteq Y^0_{ti} \qquad t \neq t'.$$

Provided that the approximations are valid ("$\doteq$" is "="), we can substitute for the missing counterfactual outcome *for each person* and identify the joint distribution of (Y_t^0, Y_t^1) for different conditioning sets. We can answer all of the questions posed in Section 1 for period t versions of the criteria presented there. It is the ability to directly estimate the dependence across potential outcomes without invoking additional assumptions that is the distinguishing feature of panel data.

When adding a temporal dimension to the analysis, it is useful to distinguish reversible from irreversible programs. Human capital or personal investment programs have certain irreversibility features, but it is typically assumed that they have no effect on preprogram outcomes.[14] For such programs, we require $t' < t$ in (I-11), where t is the period of participation. Reversible program switch on and off and have no lasting effects. Examples may include job subsidies or unemployment insurance benefits. With reversible programs we can go forward or backward in time in the search for valid counterfactual state, so that we may have $t' < t$ or $t' > t$ in (I-11). We first consider reversible programs.

Reversible Programs

Nonstationarity in the external environment, the effects of aging and life cycle investment, and idiosyncratic period-specific shocks render assumptions (I-11a) and (I-11b) suspect. To circumvent these problems, the identifying assumptions are usually reformulated at the population level and conditioning variables X are assumed that "adjust" $Y_{t'}^0$ and Y_t^0 and $Y_{t'}^1$ and Y_t^1 to equality in distribution or conditional mean and allow for idiosyncratic fluctuations. For simplicity, we only conduct a two-period analysis, but to estimate the necessary adjustments may require more data.[15] The modified identification conditions become

$$\text{(I-11a)}' \qquad F(y_{t'}^0, y_t^1 \mid X) = F(y_t^0, y_t^1 \mid X) \qquad y_{t'}^0 = y_t^0$$

[14]If agents anticipate participation in the program, they may take actions in the preprogram period that distinguish them from nonparticipants. The assumed absence of anticipatory behaviour is central to received models of program evaluation.

[15]See Heckman and Robb (1985, pp. 210-215), where these adjustments are discussed in detail.

and

(I-11b)′ $$F(y_t^0, y_{t'}^1 \mid X) = F(y_t^0, y_t^1 \mid X) \qquad y_{t'}^1 = y_t^1.$$

Weaker versions of (I-11a)′ and (I-11b)′ that are more commonly used are

(I-11a)″ $$E(Y_{t'}^0 \mid X) = E(Y_t^0 \mid X)$$

and

(I-11b)″ $$E(Y_{t'}^1 \mid X) = E(Y_t^1 \mid X).$$

The outcome variables may need to be adjusted for deterministic trends. Heckman and Robb (1985) consider cases where common deterministic trends affecting mean outcomes in both the participation and nonparticipation states can be eliminated using multiperiod and multicohort data, assuming that they are restricted to be low order functions of time or age.[16]

The potential cost of using this information on the missing counterfactual outcomes is the possibility of selection bias. Persons who don't participate in t and participate in t' may be atypical of those who participate in t', especially if their nonparticipation in t is linked to the value of the outcome variable in t. Specifically, we can use (I-11a) to construct all of the counterfactuals in period t conditional on $D = 1$ without any further adjustment if it is further assumed in the reversible case that:

(I-12a) $$F(y_{t'}^0, y_t^1 \mid D_{t'} = 0, D_t = 1, X) = F(y_t^0, y_t^1 \mid D_t = 1, X) \qquad \text{for } y_{t'}^0 = y_t^0.$$

We can use (I-11b) to construct all of the counterfactuals in period t conditional on $D = 0$ without any further adjustments if it is assumed that

(I-12b) $$F(y_t^0, y_{t'}^1 \mid D_t = 0, D_{t'} = 1, X) = F(y_t^0, y_t^1 \mid D_t = 0, X) \qquad \text{for } y_{t'}^1 = y_t^1.$$

Much less often is it also assumed that $F(y_{t'}^0, y_t^0 \mid D_{t'} = 0, D_t = 0, X) = F(y_t^0 \mid D_t = 0, X)$ for $y_{t'}^0 = y_t^0$ or $F(y_{t'}^1, y_t^1 \mid D_{t'} = 1, D_t = 1, X) = F(y_t^1 \mid D_t = 1, X)$, $y_{t'}^1 = y_t^1$ although these assumptions seem equally plausible and are testable. They would require that the $Y_{t'}^1$ and Y_t^1 are perfectly dependent as are outcomes $Y_{t'}^0$ and Y_t^0.

[16] In the method of difference-in-differences, it is assumed that a common trend operates on all persons irrespective of their participation status. The trend is eliminated from the means by comparing participant change to nonparticipant change. More generally, nonparticipants in t and t' can be used to identify the common trend.

For means, the weaker versions of (I-12a) and (I-12b) are, respectively,

(I-12a)′ $E(Y_{t'}^0 \mid D_{t'} = 0, D_t = 1, X) = E(Y_t^0 \mid D_t = 1, X)$ (for (I-11a)″)

and

(I-12b)′ $E(Y_{t'}^1 \mid D_t = 0, D_{t'} = 1, X) = E(Y_t^1 \mid D_t = 0, X)$ (for (I-11b)″).

These are strong implicit behavioral assumptions. Assumption (I-12a) and (I-12a)′ require that persons who participate in t but not in t' have the same no-treatment mean outcome in t' as persons who take treatment in period t would have in t. It rules out that the switch from $D_{t'} = 0$ to $D_t = 1$ is caused by differences in Y^0 between t' and t. More precisely, it excludes $Y_{t'}^0$ as a determinant of $D_{t'}$. Assumptions (I-12b) and (I-12b)′ are comparable assumptions about the lack of influence of $Y_{t'}^1$ in determining participation in t'.

One way to justify these identifying assumptions is to postulate a strengthened form of the conditional independence assumption used to justify matching:

(I-13) $(Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp (D_{t'}, D_t) \mid X, \quad t \neq t'.$

This condition rules out any dependence between D_t and $D_{t'}$ and the components of $(Y_{t'}^{D_{t'}}, Y_t^{D_t})$ that cannot be predicted by X. This assumption rules out selection on any unobserved components of potential outcomes. It is inconsistent with the Roy model. A weaker version of (I-13) is that conditional on D_t and X, $(Y_{t'}^{D_{t'}}, Y_t^{D_t})$ are independent of $D_{t'}$:

(I-14) $(Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp D_{t'} \mid X, D_t.$

This condition rules out any dependence between the components of $(Y_{t'}^{D_{t'}}, Y_t^{D_t})$ that cannot be predicted by D_t and X and the random variable $D_{t'}$. (I-12a)′ and (I-12b)′ can be justified by these assumptions.

We could augment (I-13) or (I-14) to include matching variables Z not included in X. Thus it may happen that (I-13) does not hold but

(I-13)′ $(Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp (D_t, D_{t'}) \mid X, Z.$

Similarly, (I-14), may be invalid but it may happen that

(I-14)′ $(Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp D_{t'} \mid X, D_t, Z$

is valid. We may also invoke other assumptions patterned after our cross section analysis to recover the missing counterfactual state. We could model participation in periods t and t' using dynamic selection models. Each cross section estimator has a panel data counterpart which, for the sake of brevity, we do not develop in this paper.

If the date of enrollment into the program is endogenous, it is incorrect to simply condition on it and conditions (I-13) and (I-14) have to be strengthened in order to avoid building an explicit model of the date of enrollment.[17] Let τ be the date of enrollment into the program. Then to use (I-12a) and (I-12b) without modification, we need to augment the conditional independence assumptions to read

(I-13)″ $\quad (Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp (D_{t'}, D_t, \tau) \mid X, \qquad t \neq t'$

or in the weaker form

(I-14)″ $\quad (Y_{t'}^{D_{t'}}, Y_t^{D_t}) \perp\!\!\!\perp (D_{t'}, \tau) \mid X, D_t, \qquad t \neq t'.$[18]

These conditions rule out dependence between potential outcomes and the set of participation variables conditional on X (I-13)″ or dependence between potential outcomes and non-t participation variables conditioned on X and D_t (I-14)″. Under either set of assumptions, we can ignore the date of enrollment as a factor in producing the counterfactual distributions.

Other types of identifying assumptions can be invoked. Cameron and Heckman (1991) develop a multivariate version of the Roy model that explicitly models τ and show that its parameters can be identified. These models are closely related to standard panel data attrition models. (See, *e.g.,* Ridder, 1991.)

The Irreversible Case

In the irreversible case, there are no counterparts for (I-11b)′, (I-11b)″, (I-12b) or (I-12b)′ because there are no observations on treated persons in the preprogram period t'. First

[17] In a fully dynamic model in which enrollment dates are endogenous, the date of enrollment would be a further source of information about revealed preferences, which we do not pursue in this paper. Qualitatively, it conveys information on subjective evaluations in the same way attrition and self-selection decisions convey information about choices.

[18] The required modification for conditional means is obvious and hence omitted.

consider the case where program enrollment date τ is fixed and common for all persons. The probability space is restricted so $\Pr(D_{t'} = 1 \mid X) = 0$ and no value of $Y_{t'}^1$ is defined. $F(y_t^0, y_t^1 | D_t = 1, X)$ can be identified from $F(y_{t'}^0, y_t^1 \mid D_t = 1, X)$ if the preprogram outcomes of participants have the same relationship to program outcomes in t as their nonprogram outcomes in period t. (This is just assumption (I-12a).) We cannot use (I-12b) to construct $F(y_t^0, y_t^1 \mid D_t = 0, X)$ because no value of $Y_{t'}^1$ is defined. In the irreversible case, we have a truncated sample.

If we invoke a conditional independence assumption and assume a counterpart to (I-12) defined for the reversible case:

(I-15) $$(Y_{t'}^0, Y_t^1) \perp\!\!\!\perp D_t \mid X,$$

we can identify the full joint distribution.[19] Otherwise, we can only identify the criteria for the population conditional on $D_t = 1$. Since we know $(Y_{t'}^0, Y_t^1)$ conditional on $D_t = 1$ and X, we can use a vector generalization of Theorem A-2, presented in Appendix A as Theorem A-3, to identify $F(y_t^0, y_t^1 \mid X)$ and $F(y_t^0, y_t^1, D_t \mid X)$. What is required is a set of X values where $\Pr(D_t = 1 \mid X) = 0$. Under the assumptions made in Theorem A-3, it is possible to recover the full distribution of outcomes even in the reversible case.

If τ is not the same for everyone, and is random, but $t' < \tau < t$, then to use (I-15) we need to assume

(I-16) $$(Y_{t'}^0, Y_t^1) \perp\!\!\!\perp D_t, \tau \mid X$$

or

(I-16)′ $$(Y_{t'}^0, Y_t^1) \perp\!\!\!\perp D_t \mid \tau, X.$$

These assumptions enable us to ignore the date of enrollment as a determinant of outcomes in constructing the counterfactual distributions.

Heckman and Robb (1985) discuss more general uses of panel data to proxy unobservables to eliminate selection bias. The leading cases are fixed effect or autoregressive models that transform equations by differencing or generalized differencing to eliminate unobserved

[19] This assumption could be augmented to allow for Z to be added to the conditioning set so that we have $(Y_{t'}^0, Y_t^1) \perp\!\!\!\perp D_t \mid X, Z$ but (I-15) is invalid.

components that produce selection bias. All of the conventional "proxy variable" econometric methods that eliminate selection bias through some transformation of the original equations can be shown to be equivalent to constructing counterfactual outcomes, *i.e.*, producing predicted values of the outcomes needed to form the missing component of the counterfactual. More generally, if the original equations are subject to transformations, the previously stated identification conditions apply to the transformed equations. See Heckman (1997b).[20] A summary of the main identification results for joint distributions and means and marginal distributions that exploit panel data are given in Tables 4 and 5, respectively.

D. Repeated Cross Section Data

Heckman and Robb (1985) demonstrate that all panel data identification assumptions about means, variances and covariances have counterparts in repeated cross section data. Conditional mean versions of all of the identification assumptions presented in Section C have counterparts in repeated cross sections of unrelated persons sampled from the same populations. We first consider the reversible case.

Identification conditions (I-11a)$''$ and (I-11b)$''$ can be defined for a common population and do not require that the same persons be followed over time. The same is true for (I-12a)$'$ and (I-12b)$'$ and the other identifying assumptions for conditional means. However, it now becomes necessary to classify persons in t' as program participants or nonparticipants in t. This is not so easy to do in the repeated cross section case because different persons are sampled in t and t'. What is lost when the analyst is restricted to using repeated cross section data is the ability to construct joint distributions (Y_t^0, Y_t^1) without invoking the assumptions made in Section A.

[20] For example, in the method of fixed effects without regressors, $Y_{it}^1 = \alpha_i + \varphi_i + \varepsilon_{it}$ and $Y_{it}^0 = \varphi_i + \varepsilon_{it}$, where $E(\varepsilon_{it}) = 0$ and $\varepsilon_{it} \perp\!\!\!\perp \varphi_i$, $Y_t^0 \doteq Y_{t'}^0$ and $Y_{it}^1 - Y_{it'}^0 = \alpha_i + \varepsilon_{it} - \varepsilon_{it'}$. If $\Pr(D_i = 1 \mid \varepsilon_{it} - \varepsilon_{it'}) = P$, which is not a function of $\varepsilon_{it} - \varepsilon_{it'}$, we can identify $E(\alpha_i \mid D = 1) = E(Y_{it}^1 - Y_{it'}^0 \mid D = 1)$. Observe that P can depend on φ_i. See Heckman and Robb (1985). Heckman (1997b) demonstrates that $Y_{it'}^0$ is properly interpreted as a proxy for Y_{it}^0. If there are regressors, we can modify this example to allow for use of X-adjusted $Y_{it}^1 - Y_{it'}^0$, $[(Y_{it}^1 - X_{it}\beta) - (Y_{it'}^0 - X_{it'}\beta)]$, where for convenience we assume a common β.

Without invoking additional assumptions about dependence between the two potential outcomes, the identifying assumptions for conditional means only enable us to recover the cost-benefit and the Benthamite criteria and not the voting criteria, which is based on the full joint distribution of potential outcomes. The essential benefit of panel data—that they afford nonparametric identification of the joint distribution of potential outcomes under the identifying assumptions made in Section C—is lost when the analyst only has access to repeated cross section data.[21] A summary of the main cases for panel data repeated cross sections is presented in Table 5.

(3) The Relationship Between Traditional Cost-Benefit Analysis and The Parameters Widely Used In The Econometric Evaluation Literature

In this section we relate the parameters estimated in the micro-econometric evaluation literature to the parameters needed to perform cost-benefit analysis. We present empirical evidence on the importance of accounting for the direct costs of a program and the marginal welfare costs of taxation in assessing the net benefits of a policy. We follow the literature in cost-benefit analysis and assume that the policy being evaluated has a voluntary component and that valid evaluations of a policy can be derived from looking at the impact of the policy on self-selected participants and nonparticipants.

We postulate the following framework. For a given program associated with policy j, there are two discrete outcomes corresponding to direct receipt of treatment ($D_j = 1$, for program participation) or not ($D_j = 0$), and a set of program intensity variables φ_j defined under policy j that affect outcomes in the two states and the allocation of persons to "treatment" or nontreatment. The program intensity variables φ_j may be discrete or continuous. Policy "0" is a no intervention benchmark with program intensity φ_0.

Assuming that costless lump-sum transfers are possible, that a single social welfare function governs the distribution of resources and that prices reflect true opportunity costs,

[21]The modification of the analysis in this subsection to account for irreversibility is straightforward and is omitted.

traditional cost-benefit analysis (see, *e.g.*, Harberger (1971) or Boadway and Bruce (1984)) seeks to determine the impact of programs on the total output of society. Efficiency becomes the paramount criterion in this framework, with the distributional concerns assumed to be taken care of through lump sum transfers and taxes. In this framework, impacts on total output, as in the evaluation criterion (3), are the only objects of interest in evaluating policies.

For policy j let Y_{ji}^1 and Y_{ji}^0 be individual output for person i in the direct participation ($D_j = 1$) and direct non-participation ($D_j = 0$) state, respectively. The vector of program intensity variables φ_j operates on all persons within the context of program j, although its effect need not be uniform. It determines, in part, participation in the program. We may write $D_j(\varphi_j)$ as the indicator for participating in program j when program intensity is φ_j. To simplify notation we keep implicit any conditioning on personal characteristics that may affect both participation and outcomes. We define $c_j(\varphi_j)$ as the social cost of φ_j denominated in units of output. In general, policies could be designed for specific persons but we do not consider that possibility here. We assume that $c_j(0) = 0$ and that c is convex and increasing in φ_j. The value φ_0 defines another benchmark policy, "0", in which there is no program and therefore no participants. This policy has associated cost function $c_0(\varphi_0)$.

When $\varphi_j = 0$, there might be effects of policy j on output that distinguish that policy from the no policy regime "0". A law that is universally assented to and accepted may raise output at no cost (*e.g.*, adopting a convention about driving on the right hand side of the road). Output could be different in a policy without the law (policy "0") but the direct costs of enforcement would be the same under both policies.

Letting $N_1(\varphi_j)$ be the number of direct program participants and $N_0(\varphi_j)$ be the rest of the population, the total output of society under policy j at program intensity level φ_j is

$$N_1(\varphi_j)E(Y_j^1 \,|D(\varphi_j) = 1, \varphi_j) + N_0(\varphi_j)E(Y_j^0 \,|D(\varphi_j) = 0, \varphi_j) - c(\varphi_j),$$

where $N_1(\varphi_j) + N_0(\varphi_j) = \bar{N}$ is the total number of persons in society. "φ_j" appears twice in the conditioning arguments: as a determinant of D_j and as a determinant of the output levels in the different states. Vector φ_j is general enough to include financial incentive variables as well as mandates that assign persons to a particular treatment state. Recall that we keep conditioning on personal characteristics implicit.

Assume for simplicity the differentiability of the treatment choice and mean outcome functions and further assume that φ_j is a scalar, a simplifying assumption that is easily relaxed. The change in output in response to a marginal increase in the policy intensity parameter φ_j from any given position is:

$$\begin{aligned} M(\varphi_j) &= \frac{\partial N_1(\varphi_j)}{\partial \varphi_j}\left[E\left(Y_j^1|D_j(\varphi_j)=1,\varphi_j\right)-E\left(Y_j^0|D_j(\varphi_j)=0,\varphi_j\right)\right] \\ &+N_1(\varphi_j)\left[\frac{\partial E\left(Y_j^1|D(\varphi_j)=1,\varphi_j\right)}{\partial \varphi_j}\right] \\ &+N_0(\varphi_j)\left[\frac{\partial E\left(Y_j^0|D(\varphi_j)=0,\varphi_j\right)}{\partial \varphi_j}\right]-c_j'(\varphi_j). \end{aligned}$$

The first term arises from the change in the number of participants induced by the policy change. The second and third terms arise from changes in output among participants and nonparticipants induced by the policy change. The fourth term is the marginal direct output cost of the change in the intensity of policy φ_j.

In principle, this measure could be estimated from time-series data on the change in aggregate GNP occurring after the policy intensity parameter is varied. Under the assumption of a well-defined social welfare function with interior solutions and the additional assumption that prices are constant at initial values, an increase in GNP at base period prices raises social welfare.[22]

If marginal program intensity changes under policy regime j have no effect on intra-sector mean output, the bracketed expressions in the second and third terms are zero. In this case, the parameters of interest are:

[22]See, *e.g.*, Laffont (1989, p. 155), or the comprehensive discussion in Chipman and Moore (1976).

(i) $\dfrac{\partial N_1(\varphi_j)}{\partial \varphi_j}$ — the number of people induced into program j by the change in φ_j,

(ii) $E\left(Y_j^1 | D_j(\varphi_j) = 1, \varphi_j\right) - E\left(Y_j^0 | D_j(\varphi_j) = 0, \varphi_j\right)$ — the mean output difference between participants and nonparticipants.

(iii) $c_j'(\varphi_j)$ — the direct social marginal cost of policy j at program intensity level φ_j.

It is revealing that nowhere on this list are the parameters that receive the most attention in the econometric policy evaluation literature. (See, *e.g.*, Heckman and Robb, 1985.) These are:

(a) $E\left(Y_j^1 - Y_j^0 | D_j(\varphi_j) = 1, \varphi_j\right)$ — "the effect of treatment on the treated" for persons in regime j at policy intensity φ_j.

(b) $E\left(Y_j^1 - Y_j^0 | \varphi_j = \bar{\varphi}\right)$ — where $\varphi_j = \bar{\varphi}$ sets $N_1(\bar{\varphi}) = \bar{N}$. This is the effect of universal direct participation in program j compared to universal nonparticipation in j at level of program intensity $\bar{\varphi}$.

(c) $E\left(Y_j^1 - Y_j^0 | \varphi_j\right)$ — The effect of randomly selecting someone for direct treatment and forcing their compliance with this treatment compared to their position in the no participation state under policy j at program intensity level φ_j.

Parameter (ii) can be obtained from simple mean differences between the treated and the nontreated. No adjustment for selection bias is required. Parameter (i) can be obtained from knowledge of the net movement of persons into or out of direct participation in the program in response to the policy change, something usually not measured in micro policy evaluations (for discussions of this problem, see Moffitt, 1992 or Heckman, 1992). Parameter (iii) can be obtained from cost data. It should include full social costs of the program, including the welfare cost of raising public funds, although these are often ignored.

It is informative to place additional structure on this model. This leads to a representation of a criterion that is widely used in the literature on microeconomic program evaluation and also establishes a link with the discrete choice literature in econometrics. Assume a binary choice random utility framework like that used in the Roy model. Suppose that under policy regime j with program intensity level φ_j agents make choices to directly participate or not based on net utility and that policies affect participant utility through an additively-separable term, $k(\varphi_j)$, that is assumed scalar and differentiable. Net utility from participating in the program is $U_j = X + k(\varphi_j)$, where k is monotonic in φ_j and where the joint distributions of (Y_j^1, X) and (Y_j^0, X) are $F(y_j^1, X)$ and $F(y_j^0, X)$, respectively.[23] In the special case of the Roy model, $X = Y_j^1 - Y_j^0$ and $k = 0$. In general, $D_j(\varphi_j) = 1(U_j \geq 0) = 1(X \geq -k(\varphi_j))$, so

$$\begin{aligned} N_1(\varphi_j) &= \bar{N} \Pr(U_j \geq 0) = \bar{N} \textstyle\int_{-k(\varphi_j)}^{\infty} f(x)\, dx \\ N_0(\varphi_j) &= \bar{N} \Pr(U_j < 0) = \bar{N} \textstyle\int_{-\infty}^{-k(\varphi_j)} f(x)\, dx. \end{aligned}$$

The total output is

$$\bar{N} \int_{-\infty}^{\infty} y^1 \int_{-k(\varphi_j)}^{\infty} f\left(y^1, x|\varphi_j\right) dx dy^1 + \bar{N} \int_{-\infty}^{\infty} y^0 \int_{-\infty}^{-k(\varphi_j)} f\left(y^0, x|\varphi_j\right) dx dy^0 - c_j(\varphi_j).$$

Under standard conditions[24], we may differentiate under the integral sign to obtain the following expression for the marginal change in output with respect to a change in intensity parameters φ_j within policy regime j :

$$\begin{aligned} &M(\varphi_j) = \\ &\bar{N} k'(\varphi_j) f_x(-k(\varphi_j)) \left[E\left(Y_j^1|D(\varphi_j) = 1, X = -k(\varphi_j), \varphi_j\right) - E(Y_j^0|D(\varphi_j) = 0, X = -k(\varphi_j), \varphi_j)\right] \\ &+\bar{N}\left[\textstyle\int_{-\infty}^{\infty} y^1 \int_{-k(\varphi_j)}^{\infty} \frac{\partial f\left(y^1, x|\varphi_j\right)}{\partial \varphi_j} dx dy^1 + \int_{-\infty}^{\infty} y^0 \int_{-\infty}^{-k(\varphi_j)} \frac{\partial f\left(y^0, x|\varphi_j\right)}{\partial \varphi_j} dx dy^0\right] - c_j'(\varphi_j), \end{aligned}$$

where f_x, the marginal density of X, is evaluated at $X = -k(\varphi_j)$.

This model has a well-defined marginal entry condition: $X \geq -k(\varphi_j)$. The first set of terms corresponds to the gain arising from the movement of persons at the margin (the

[23] These are assumed to be absolutely continuous with respect to Lebegesue measure.

[24] See, *e.g.*, Royden (1968) for the required domination conditions.

term in brackets) weighted by the proportion of the population at the margin, $f_x(-k(\varphi_j))$, times the number of people in the population. This term is the net gain from switching from nonparticipant to participant status. The expression in brackets in the first term is a limit form of the "local average treatment effect" of Imbens and Angrist (1994). The second set of terms is the within-treatment-status change in output resulting from the change in the program intensity parameter. This term is ignored in many evaluation studies. It describes how people who do not switch their participation status are affected by the policy change. The third term is the direct marginal social cost of the policy change, which is rarely estimated. At a social planner's optimum, $M(\varphi_j) = 0$, provided standard second order conditions are satisfied. Marginal benefit should equal the marginal cost. Either a cost-based measure of marginal benefit or a benefit-based measure of cost can be used to evaluate the marginal gains or costs of the change in policy intensity.

Observe that the local average treatment effect is simply the effect of treatment on the treated for persons at the margin $(X = -k(\varphi_j))$:

$$\begin{aligned} & E\left(Y_j^1 | D_j(\varphi_j) = 1, X = -k\left(\varphi_j\right), \varphi_j\right) - E\left(Y_j^0 | D_j(\varphi_j) = 0, X = -k\left(\varphi_j\right), \varphi_j\right) \\ = \; & E\left(Y^1 - Y^0 | D(\varphi_j) = 1, X = -k\left(\varphi_j\right), \varphi_j\right). \end{aligned}$$

The proof of this result is immediate once it is recognized that the set $X = -k(\varphi_j)$ is the indifference set for this problem. Thus, the Imbens and Angrist parameter is a marginal version of the conventional "treatment on the treated" evaluation parameter for gross outcomes. This parameter is but one of the three ingredients required to produce an evaluation of social welfare under the cost-benefit criterion.

The conventional evaluation parameter "treatment on the treated"

$$E\left(Y_j^1 - Y_j^0 \, | D_j\left(\varphi_j\right) = 1, X, \varphi_j\right),$$

does not incorporate costs, does not correspond to a marginal change and includes the effect of intramarginal changes. This parameter is in general inappropriate for evaluating the effect of a policy change on GNP. However, under certain conditions which we now make precise, it is sometimes informative about the gross gain accruing to the economy from the existence of program j at level φ_j compared to the alternative of shutting it down and switching to policy "0". The social cost associated with policy "0" is $c_0(\varphi_0)$, which we assume is zero: $c_0(\varphi_0) = 0$.

The appropriate criterion for an all or nothing evaluation of a policy at level φ_j is

$$A(\varphi_j) = \left\{N_1(\varphi_j) E\left(Y_j^1 | D_j(\varphi_j) = 1, \varphi_j\right) + N_0(\varphi_j) E\left(Y_j^0 | D_j(\varphi_j) = 0, \varphi_j\right) - c_j(\varphi_j)\right\} - \bar{N} E(Y_0 | \varphi_0).$$

In the no policy regime, there is only one output Y_0 and everyone is in the "no program" state. If $A(\varphi_j) > 0$, total output is increased by establishing program j at level φ_j. In the special case where the outcome in the nonparticipation state under regime j, Y_j^0, is the same as the outcome in the no-program state (Y_0) both for participants and nonparticipants under regime j, we have

(23a) $$E(Y_j^0 | D_j(\varphi_j) = 0,\ \varphi_j) = E(Y_0 | D_j(\varphi_j) = 0,\ \varphi_0)$$

and

(23b) $$E(Y_j^0 | D_j(\varphi_j) = 1,\ \varphi_j) = E(Y_0 | D_j(\varphi_j) = 1,\ \varphi_0).$$

The right hand sides of both expressions describe hypothetical conditional expectations. The right hand side of (23a) is what the outcome in the no-program state would be for persons who do not directly participate in the program under policy j with parameters φ_j, *i.e.*, those for whom $D_j(\varphi_j) = 0$. The right hand side of (23b) is the corresponding expression for persons who would participate in the program under policy j with intensity parameters φ_j, *i.e.*, those for whom $D_j(\varphi_j) = 1$. These conditioning statements select out, respectively, non-participants and participants in policy regime j and compute the expected values of output in the policy "0" regime.

Assuming that the probability of participation in regime j under program intensity level φ_j does not depend on the value of φ_0 in the no-program state:

(A-3) $$\Pr(D_j = 1 | \varphi_j, \varphi_0) = \Pr(D_j = 1 | \varphi_j),$$

under assumption (A-1) we may use the law of iterated expectations to write

$E(Y^0 \,|\varphi_0) =$

$$E(Y_0 \,|D_j(\varphi_j) = 1, \varphi_0) \Pr(D_j(\varphi_j) = 1 \,|\varphi_j) + E(Y_0 \,|D_j(\varphi_j) = 0, \varphi_0) \Pr(D_j(\varphi_j) = 0 \,|\varphi_j).$$

From (23a) and (23b) and (A-3) we obtain

$E(Y^0 \,|\varphi_0) =$

$$E(Y_j^0 \,|D_j(\varphi_j) = 1, \varphi_j) \Pr(D_j(\varphi_j) = 1 \,|\varphi_j) + E(Y_j^0 \,|D_j(\varphi_j) = 0, \varphi_j) \Pr(D_j = 0 \,|\varphi_j).$$

Substituting for $E(Y_j^0 \,|\varphi_0)$ in the expression for $A(\varphi_j)$, we obtain

(24) $$A(\varphi_j) = N(\varphi_j)E(Y_j^1 - Y_j^0 \,|D_j(\varphi_j) = 1, \varphi_j) - c_j(\varphi_j),$$

which vindicates the use of the parameter "treatment on the treated" as an evaluation parameter in the case in which there are no general equilibrium effects in the sense of assumption (A-1). This important case is applicable to small-scale social programs with partial participation. For evaluating the effect of "fine-tuning" the intensity levels of existing policies, measure $M(\varphi_j)$ is more appropriate.

Empirical Evidence On the Importance of Adjusting For Direct Costs and the Welfare Costs of Taxation in Cost-Benefit Analysis

This subsection examines the effect of accounting for both direct costs and the welfare costs of raising government tax revenue in computing benefit-cost estimates for a prototypical government training program. Accounting for direct costs and the welfare costs of government revenue substantially reduces the estimated return to government training programs over what is conventionally reported. Our estimates of the difference between costs and benefits for the JTPA program appear in Table 6. Benefits are measured using the difference in mean earnings between the experimental treatment and control groups in the JTPA data which are described more fully in Section 4 and Appendix B. Direct costs represent the estimated difference in training costs between the treatment and control groups[25],

[25] Both the impact and cost estimates are drawn from the analysis in Orr, et al. (1995). The first row of Table 6 corresponds to the case they consider. The impact estimates in Orr, et al. (1995) differ somewhat from those presented in Bloom, et al. (1993) due to differences in sample composition and earnings measure. The remainder of our empirical evidence is based on the Bloom, et al. (1993) sample and earnings measure, which we prefer because it does not combine earnings information from different

and are assumed to occur within the first six months after random assignment. The first row for each demographic group presents the experimentally-estimated unadjusted benefits of participation over the 30 month post random-assignment period for which data are available. Each of the remaining rows of Table 6 presents estimates that net out the direct costs of training based on the assumptions stated in those rows about the duration of program benefits (30 months or 7 years), the interest rate used to discount the benefits (0.00 or 0.25 over six months), and the welfare cost of taxation ($0.0 or $0.50 per dollar of revenue).

Three main conclusions emerge from this analysis. First, netting out the direct costs of training is empirically important. For job training programs, costs are often large relative to the estimated benefits, as is clearly the case for female youth where benefits and costs are roughly equal. Second, accounting for the welfare costs of taxation has a substantial effect on the cost-benefit calculation. For adult females with benefits assumed to last 30 months, and assuming no discounting, netting out welfare costs of taxation equal to $0.50 per dollar changes the difference between costs and benefits from $532 to $-54. The estimates of the welfare cost of public funds presented in the literature vary over the range from $0.00 to $3.00 per dollar of taxes. See Browning (1987). However, the "consensus value" is less than $1.00 and typically in the range of $0.30-0.50. If the welfare cost of taxation rises as the amount of taxes raised increases, an issue arises about whether programs should be evaluated as if they were the first program (so that taxes increase from zero to the level needed to finance the program), the marginal program (so that taxes increase from existing levels by the cost of the program) or some intermediate value. Even more problematic is the case where the tax effects of groups of programs interact and programs are bundled in the legislative process. In this paper we use the marginal cost of funds given the current scale of government, which is appropriate because the scale the JTPA program is relatively small and we are evaluating it in isolation from other programs.

A third conclusion from Table 6 is that the estimated cost-benefit difference is sensitive

data sources. We use the Orr, et al. (1995) in Table 6 because cost estimates are not readily available for the Bloom, et al. (1993) sample.

to the assumed duration of benefits. The best evidence on the longevity of program benefits is that of Couch (1993). He shows that benefits to adult women from the National Supported Work program remain at the same level for at least seven years beyond random assignment so that our estimate for this group is conservative. There is also evidence from U.S. General Accounting Office (1996) that JTPA program benefits extend at least five years after random assignment for a subset of the experimental sample.

(4) Evidence on Impact Heterogeneity and the Value of Self-Assessments and Revealed Preference Information

This section of the paper addresses three questions. Question (1) is: "What is the empirical evidence on heterogeneity in program impacts among persons?" The conventional approach implicitly assumes impact homogeneity conditional on observables. This assumption greatly simplifies the task of evaluating the welfare state. Using data from an experimental evaluation of a prototypical job training program, we utilize many of the assumptions presented in Section (2) to bound or identify the joint distribution of outcomes conditional on $D = 1$. We find considerable evidence of heterogeneity of program impacts, so that conventional econometric methods do not take us very far in constructing the evaluation criteria discussed in Section 1. Use of experimental data enables us to avoid the self-selection problems that plague ordinary observational data, and simplifies our analysis.

Given our evidence on impact heterogeneity, we ask question (2): "How sensitive are the estimates of the proportion of people who gain from the program - what we have called the "voting criterion" - to alternative assumptions about the dependence between Y^0 and Y^1?" We find that the estimates are very sensitive to alternative assumptions. At the same time, for adult women, the estimated percentage that benefit from the program exceeds 50 percent in every case we consider but one, and is close to 100 percent in some cases.

Some of the estimates used to answer question (2) assume that Y^0 and Y^1 are positively dependent given $D = 1$. We established in Section 2 that under purposive selection based on outcomes in the treated and untreated states, such dependence among participants arises even if Y^1 and Y^0 are independent or negatively correlated in the population as a whole. An alternative to imposing a particular decision rule is to infer it from self-assessments of the program. These assessments are all that are required for a libertarian evaluation of the welfare state. We examine the implicit value placed on the program by addressing the following questions: (3a) "Are persons who applied to the program and were accepted into it but then randomized out of it placed in an inferior position relative to those accepted applicants who were not randomized out?" We measure *ex ante* rational regret using second-order stochastic dominance, which is an appropriate measure under the assumption that individuals are completely uncertain of both Y^1 and Y^0 before going into the program. We also consider *ex post* evaluations of participants by asking: (3b) "How "satisfied" are participants with their experience in the program?" Self-assessments of programs are widely used in evaluation research (see *e.g.*, Katz, et al., 1975), but the meaning to be placed on them is not clear. Do they reflect an evaluation of the experience of the program (its process) or an evaluation of the benefits of the program? Our evidence suggests that respondents report a net benefit inclusive of their costs of participating in the program. Groups for whom the program has a negative average impact as estimated by the "objective" experimental data express as much (or more) enthusiasm for the program as groups with positive average impacts. A third source of revealed preference evaluations uses the revealed choices of attriters from the program. Econometric models of self-selection since Heckman (1974a,b) have used revealed choice behavior to infer the evaluations people place on programs either by selecting into them or dropping out of them. The third part of the third question is thus (3c): "What implicit valuation of the program do attriters place on it?"

(A) Data

Our estimates are based on data from the recent experimental evaluation of the employment and training programs funded under the Title II-A of the U.S. Job Training Partnership Act (JTPA) (see Orr, et al. (1995)). This program provides classroom training, on-the-job training and job search assistance to the economically disadvantaged. We focus primarily, but not exclusively, on adult women (age 22 or older) for many, but not all, of our analyses. We also present selected results for other demographic groups: adult men (age 22 and older) and male and female out-of-school youth (ages 16-21). Our largest samples are for adult women. Given that many of the adult women in the program are welfare recipients, their experiences with training are of special interest given recent reforms in the U.S. welfare system. Appendix B describes the JTPA data in greater detail.

(B) Evidence on Impact Heterogeneity

This subsection presents evidence on variability in the response to training. We find strong evidence against homogeneity. However unless the dependence across outcomes in the treated and untreated states is very high, the estimated variability in program gains is implausibly large.

Suppose that the JTPA experiment satisfies (I-10). Suppose that there are N treated persons and N nontreated persons. Suppose that the outcomes are continuously distributed. Rank the individuals in each treatment category in the order of their outcome values from the highest to the lowest. Define as $Y^j_{(i)}$ as the i^{th} highest-ranked person in the j distribution. Ignoring ties, we obtain two data distributions:

$$\underline{\text{Treatment Outcome}}\text{:}F(y^1|D=1) \qquad \underline{\text{Non-Treatment Outcome}}\text{:}F(y^0|D=1)$$

$$\underset{\sim}{Y}^1 = \begin{pmatrix} Y^1_{(1)} \\ \vdots \\ Y^1_{(N)} \end{pmatrix} \qquad\qquad \underset{\sim}{Y}^0 = \begin{pmatrix} Y^0_{(1)} \\ \vdots \\ Y^0_{(N)} \end{pmatrix}$$

We know the marginal data distributions $F(y^1|D=1)$ and $F(y^0|D=1)$, but we do not know where person i in the treatment distribution would appear in the non-treatment distribution. These distributions can also be defined conditional on X. Corresponding to the ranking of the treatment outcome distribution, there are $N!$ possible patterns of outcomes in the associated non-treatment outcome distribution. By considering all possible permutations, we can form a collection of possible impact distributions, *i.e.,* alternative distributions of

$$\underset{\sim}{\Delta} = \underset{\sim}{Y}^1 - \Pi_\ell \underset{\sim}{Y}^0 \qquad \ell = 1, \ldots, N!$$

where Π_ℓ is a particular $N \times N$ permutation matrix of Y^0 in the set of all $N!$ permutations associating the ranks in the Y^1 distribution with the ranks in the Y^0 distribution and $\underset{\sim}{\Delta}$, $\underset{\sim}{Y}^1$ and $\underset{\sim}{Y}^0$ are $N \times 1$vectors of impacts, treated and untreated outcomes. By considering all possible permutations, we obtain all possible sortings of treatment, Y^1, and non-treatment, Y^0, outcomes using realized values from one distribution as counterfactuals for the other.

The "dummy endogenous variable" model assumes a constant treatment effect for all persons. This model admits only one permutation: $\Pi = I$ for each X. The best in one distribution is the best in the other distribution. In the common effect case, Y^1 and Y^0 differ by a constant for each person. A generalization of that model preserves perfect dependence in the ranks between the two distributions but does not require the impact to be the same at all quantiles of the base state distribution.

In place of ranks, we work with the percentiles of the Y^1 and Y^0 distributions, which have much better statistical properties. (See Heckman and Smith, 1993, Heckman, Smith and Clements, 1997). Equating percentiles across the two distributions, we form the pairs given in expression (13) and obtain the deterministic gain function given in (14). For the case of absolutely continuous distributions with positive density at y^0, the gain function (14) can be written as $\Delta(y^0) = F_1^{-1}(F_0(y^0|D=1)) - y^0$. We can test non-parametrically for the classical common effect model by determining if percentiles are uniformly shifted at all points of the distribution. We can form other pairings across percentiles by mapping percentiles from the Y^1 distribution into percentiles from the Y^0 distribution using the map $T : q_1 \to q_0$. The data are consistent with all admissible transformations including $q_0 = 100 - q_1$, where the best in one distribution is mapped into the worst in the other. They cannot reject any of these models or more general models where Π_ℓ is now a Markov transition matrix and we consider all possible Markov matrices.

Figure 1 presents empirical evidence on the question of the constancy of the gain effect across quantiles. It displays the estimate of $\Delta(y_0)$ for adult women assuming that the best persons in the "1" distribution are the best in the "0" distribution. More formally, it assumes that the permutation matrix $\Pi = I$. No conditioning is made so the full sample is utilized. Between the 25th and 85th percentiles the assumption of a constant impact is roughly correct. It is grossly at odds with the data at the highest and lowest percentiles.[26] Heckman, Smith and Clements (1997) and Heckman and Smith (1993) present a more extensive empirical analysis of this model for different conditioning sets and reach essentially the same conclusion.

Frechet Bounds

The Frechet bounds of expression (17) can also be applied to conditional (on $D = 1$) distributions. Both the lower and the upper Frechet bounds are proper probability distributions. At the upper bound bound, Y^1 is a non-decreasing function of Y^0. At the lower bound, Y^0 is a non-increasing function of Y^1. These bounds are not helpful in bounding the distribution of gains $\Delta = Y^1 - Y^0$, although they bound certain features of it. From a theorem of Cambanis, et al. (1976), if $k(Y^1, Y^0)$ is superadditive (or subadditive),[27] then extreme values of $E(k(Y^1, Y^0)|D = 1)$ are obtained from the upper and lower bounding distributions obtained from the experimental data.

Since $k(Y^1, Y^0) = Y^1Y^0$ is superadditive, the maximum attainable product-moment correlation $\rho_{Y^0Y^1}$ is obtained from the upper bound distribution while the minimum attainable product moment correlation is obtained at the lower bound distribution. Since $VAR(\Delta)$ is a subadditive function, it is possible to bound the variance of $\Delta(= VAR(Y^1) + VAR(Y^0) - 2\rho_{Y^0Y^1}[VAR(Y^1)VAR(Y^0)]^{\frac{1}{2}})$ and thus determine if the data are consistent with the common effect model where $Y^1 - Y^0 = \alpha$, a constant, which implies $VAR(\Delta) = 0$. Kendall's τ

[26] Standard errors for the quantiles are obtained using methods described in Csörgo (1993).

[27] A function $k(x, y)$ is superadditive if $x > x'$ and $y > y'$ implies that $k(x, y) + k(x', y') > k(x', y) + k(x, y')$. Subadditively reverses the inequality. Strict forms of these ideas convert weak inequalities into strong ones.

and Spearman's ρ also attain their extreme values at the bounding distributions.[28] However, the Frechet inequalities do not provide bounds on the quantiles of the $\Delta = (Y^1 - Y^0)$ distribution. Only the extreme high and extreme low quantile values are obtained from the Frechet bounds of the joint distribution. Table 7 presents the range of values of $\rho_{Y^1Y^0}$, Kendall's τ, Spearman's ρ and $[VAR(\Delta)]^{\frac{1}{2}}$ for the JTPA data for adult women. The ranges are rather wide, but it is interesting to observe that the Frechet bounds rule out the common effect model, as $VAR(\Delta)$ is bounded away from zero.[29] They clearly do not rule out the deterministic case of perfect correlation in the ranks across outcome distributions as long as Δ is not a constant.

Sensitivity to Alternative Assumptions About Dependence Across the Distributions

Using the sample data, we can pair percentiles of the Y^1 and Y^0 distributions for any choice of rank correlation τ between -1.0 and 1.0. The case of τ = 1.0 corresponds to the case of perfect positive dependence, where $\Pi = I$ and $q_1 = q_0$. The case where τ = -1.0 corresponds to the case of perfect negative dependence, where $q_1 = 100 - q_0$. The first and last rows of Table 8 display estimates of quantiles of the impact distribution and other features of the impact distribution for these two cases.

Heckman, Smith and Clements (1997) show how to obtain random samples of permutations conditional on values of τ between 1.0 and -1.0. We display two sets of estimates from their work. The first set assumes positive but not perfect dependence between the percentiles of Y^1 and Y^0, with $\tau = 0.95$. Estimates based on a random sample of 50 percentile permutations with this value of τ appear in the second column of Table 8. These results show that even a modest departure from perfect positive dependence substantially widens the distribution of impacts. More striking still are the results in the third column

[28]Tchen (1980).

[29]Heckman, Smith and Clements (1997) conduct a Monte Carlo analysis of the standard errors of the standard deviation of Δ. They find that these standard errors are not reliable guide to inferences regarding the null hypothesis that the true impact standard deviation is zero, using inference based on asymptotic normality of the test statistics. However, Monte Carlo estimation of the sampling distribution under the null that $Var(\Delta) = 0$ indicates that the null can be rejected in these data at the 0.0001 level.

of Table 8, which correspond to the case where $\tau = 0.0$. This value of τ is implied by independence between the percentiles of Y^1 and Y^0. Here (as in the case with τ = -1.0) the distribution of estimated impacts is implausibly wide with large positive values in each distribution often matched with zero or small positive values in the other. However, the conclusion that a majority of adult female participants benefit from the program is robust to the choice of τ.[30]

Assuming the Gain Is Independent of the Base

Another source of identifying information for the joint distribution of outcomes and the distribution of impacts postulates that the gain, Δ, is independent of the base Y^0, so that $Y^0 \perp\!\!\!\perp \Delta | D = 1$. Letting $R = 1$ if a person who applies and is provisionally accepted into the program is randomized into the program, and $R = 0$ if a provisionally accepted applicant is randomized out, $Y = Y^0 + R\Delta$, and $R\Delta \perp\!\!\!\perp Y^0$. Throughout we condition on $D = 1$. This identifying condition would be satisfied if Y^0 is known but the gain, Δ, cannot be forecast at the time decisions are made about program participation. This case is extensively discussed in Heckman and Robb (1985, p.181), and produces a model that is intermediate between the common-effect model and the variable-impact model when the impact is anticipated by agents.

Setting $Y^0 = X\beta + U_0$, we obtain a conventional random coefficient model for a regression: $Y = RY^1 + (1 - R)Y^0 = X\beta + R\Delta + U_0$. Using a components of variance model we may write $E(\Delta) = \overline{\Delta}, \varepsilon = \Delta - \overline{\Delta}$ to obtain

$$Y = X\beta + R\overline{\Delta} + \varepsilon R + U \qquad E(\varepsilon R + U | X, R) = 0.$$

Following the analysis presented in Section 2, we estimate the variance of ε.

The first row of Table 9 presents estimates of the random coefficient based on the identifying assumption $\Delta \perp\!\!\!\perp Y_0 | D = 1$. The evidence supports the hypothesis that

[30] Heckman, Smith and Clements, (1997) present methods for allowing for mass points of zero earnings in the population, and some evidence derived from such methods. Their qualitative conclusions on variability are similar to ours.

$VAR(\Delta) > 0$, suggesting that a more elaborate approach to estimating the distribution of Δ based on deconvolution is likely to be fruitful. If we maintain normality of Y^1 and Y^0 (given $D = 1$ and X), the distribution of Δ is normal with mean $\overline{\Delta}$ and variance $VAR(\Delta)$ and deconvolution is easy to perform. Under this assumption, we can estimate the voting criterion and determine the estimated proportion of people who benefit from the program.

More generally, it is not necessary to assume that the distribution of Δ is normal. We use the deconvolution procedure discussed in Heckman, Smith and Clements (1997), to estimate the distribution of impacts nonparametrically. Table 9 presents parameters calculated from this distribution. The evidence suggests that under this assumption, about 43% of adult women were harmed by participating in the program. The estimated density is presented in Figure 2 and is clearly non-normal. Nonetheless, the estimated variance of the nonparametric gain distribution matches the variance for the gain distribution obtained from the random coefficient model within the range of the sampling error of the two estimates. The estimates of the proportion who benefit are in close agreement across the two models when normality is imposed on the random coefficient model. The fact that we obtain a positive density indicates that the assumption $Y_0 \perp\!\!\!\perp \Delta \,|D = 1$ is consistent with the data for women and provides support for the hypothesis that agents do not select into the program based on Δ.

(C) Evidence from Participant Behavior

Testing For Ex Ante Stochastic Rationality of Participants

If individuals choose whether or not to participate in the program based on the gross gains from the program, if they possess a common, but unknown, concave utility function, and if they know the marginal distribution of outcomes in the participation and non-participation states, then second-order stochastic dominance should order the distributions of outcomes for persons who sought to go into the program. For non-negative y^1, y^0 this form of rationality implies

$$(25) \qquad \int_0^{\alpha} F_{11}(y^1|D=1)dy^1 < \int_0^{\alpha} F_0(y^0|D=1)dy^0 \text{ for all } \alpha \in R_+$$

Draws from the Y^1 distribution produce higher expected utility than draws from the Y^0 distribution among participants. The difference between the two integrals is a measure of regret among persons randomized out from the program and forced into the no-treatment state. This condition may fail for many reasons: persons may possess more information about their potential outcomes than just the marginal distributions; persons may have different utility functions; and persons may participate in the program on a principle other than expected utility formulated in terms of gross outcomes.

We test condition (25) by comparing the integrals of the empirical CDFs of the control and treatment group earnings distributions for various values of α. Table 10 displays the results of tests of the null hypothesis of equality of the integrated distributions in (25) for adult males and females and male and female youth using self-reported earnings in the eighteen months after random assignment. The table displays test results for $\alpha \epsilon \{\$2500, \$5000, \$10,000, \ \$15,000, \$20,000, \$25,000\}$. Standard errors are obtained by bootstrapping. For adult males, the integrated CDF of earnings for the control group exceeds that for the treatment group at every point, with a p-value below 0.05 for $\alpha < \$16,500$, and below 0.10 for $\alpha < \$22,500$, which includes most of the supports of the two earnings distributions. The data for adult females provide strong evidence of rational behavior in the sense of (25), passing the test at the five percent level or better for every value of α. This evidence suggests that personal objectives and program objectives are aligned for adult women. Results for youth are mixed. For male youth, for whom the mean experimental impact is significantly negative, the difference in integrated CDF's is negative for most values of α, though not statistically significant. For female youth, the difference switches signs around $\alpha = \$11,000$, but is never close to statistical significance.

Evidence from Self-Assessments of Program Participants

Self-assessments of program participants represent an alternative to comparisons of observed outcomes as a measure of program impact. Unlike the *ex ante* measures based on second-order stochastic dominance, these measures are statements about *ex post* expectations. There is no reason why the two measures should agree if people revise their assessments based on what they learn about a program by participating in it. In this section, we consider the strengths and limitations of self-reported assessments of satisfaction with the program as an evaluation criterion, and report on self-evaluations by participants in the JTPA experimental treatment group. We also consider what can be learned from self-assessment data regarding the heterogeneity of individual treatment effects and the rationality of program participants.

Using participant assessments to evaluate a program has two main advantages relative to the approaches already discussed. First, participants have information not available to external program evaluators. They typically know more about certain components of the cost of program participation than do evaluators. Most evaluations, including the National JTPA Study, do not even attempt to value participant time, transportation, child care or other costs in evaluating program effectiveness, unless they are paid by the program through subsidies. Participants are likely to include such information in arriving at their self-assessments of the program. Second, participant evaluations provide information about the values placed on outcomes by participants relative to their perceived cost. They have the potential of providing a more inclusive measure of the program's effects than would be obtained from looking only at gross outcomes–one that includes "client satisfaction". To some parties in the welfare state, "customer satisfaction" is an important aspect of a program.

However, participant self-assessments may not be informative on the outcomes of interest to other parties in the welfare state. In evaluations of medical interventions, for example, treatment effects may not be observed by participants or may be difficult for them to assess compared to what observing scientists might report. Participant assessments of the counterfactual state may be faulty because their judgements are based on inputs or on outcome levels rather than gains over alternative levels. Persons who chose to go into the program may rationalize their participation in it in responding to questions. In addition, self-assessments, like all utility-based measures, are difficult to compare across individuals.

The top panel of Table 11 reports JTPA participant responses to a question about whether or not the program made them better off.[31] Assuming people answer honestly, and are reporting a gross impact, the self-assessment data clearly contradict the hypothesis of impact homogeneity. For all four demographic groups, 65 to 70 percent of self-reported participants give a positive self-assessment, not the 100% or 0% predicted if impacts were homogeneous. However, if respondents are reporting a perceived net impact, the evidence reported in Table 11 does *not* necessarily contradict an assumption of gross impact homogeneity if there is heterogeneity in costs across participants. The entries in the third row of Table 11 reveal that the fractions reporting a positive impact are far lower than those obtained from all of the analyses using outcome data. This evidence is consistent with one of two hypotheses: (a) that respondents are reporting net outcomes and that costs borne by participants are a substantial fraction of gross outcomes or (b) that self-assessments are inaccurate.

The evidence suggests that the self-assessments are at least partly based on inputs received rather than on outputs produced by the program. The lower panel of Table 11 shows the fraction of persons receiving each type of training whose self-assessment of the program was positive. The fraction increases with the level of treatment intensity for all four demographic groups. Expensive and more intensive services such as classroom training in occupational skills (CT-OS) and on-the-job training at a private firm (OJT) elicit a higher

[31] The exact wording of the survey question is "Do you think that the training or other assistance you got from the program helped you get a job or perform better on the job?". The question is asked only of treatment group members who report receiving JTPA services.

proportion of positive self-assessments than do less expensive services such as job search assistance (JSA) or basic education. However, the experimental impact estimates presented in Bloom, et al. (1993) reveal that treatment effectiveness and treatment intensity are not positively related. For example, for female youth classroom training in occupational skills has a more negative mean impact than the less expensive services in the "other" treatment stream. This evidence suggests that participants may have difficultly correctly constructing what would have happened to them in the absence of treatment, and so rely in part on treatment intensity or program inputs as a proxy for treatment impact.

Finally, for adult women we consider how well the self-assessment data match up with the analyses considered in earlier sections. The self-assessment data are not consistent with the assumption of perfect positive dependence in outcomes across the two states. As shown in Figure 1, for adult women the JTPA data indicate that perfect positive dependence in outcomes between the treated and untreated states implies a strictly positive impact of the program for about 85 percent of participants - all except those with zero earnings in both states. This value far exceeds the overall self-reported effectiveness rate of 44 percent reported in row 3 of Table 11. The 44 percent rate lies below that found even for the case of perfect negative dependence. Overall, the self-reported impact data appear to be too negative when compared to our analyses of the experimental earnings data. This evidence is consistent with participants reporting a net measure while the experimental "treatment effect" measures gross outcomes. The lower positive rating of the program from self assessment data than from gross outcome data is all the more striking when it is realized that the self-assessments are only recorded for people who report receiving training while the gross outcome data for participants include those who leave the program and the attriters have lower earnings than the non-attriters.

Evidence from Program Dropouts

As a result of the relatively early placement of random assignment in the JTPA participation process, many treatment group members never enroll in the JTPA program and so do not receive JTPA services.[32] In this section, we investigate what the information on dropout behavior reveals about treatment heterogeneity and participant rationality. A key limitation in doing this is that the enrollment decision depends not just on agent choices but is a joint decision of the potential participant and of JTPA staff members. The JTPA performance standards system, which rewards individual training centers based on the labor market outcomes of their enrollees, provides both a mechanism and incentive for manipulation of the enrollment decision in order to increase center performance.[33] Because we have no data on the preferences of bureaucrats, we ignore this problem and assume that the decision we observe is solely that of the potential participant.

If anticipated discounted net impacts are the same across all persons, then everyone either participates in the program or drops out of it. The substantial dropout rates reported in the first column of Table 12 for all four demographic groups provide evidence that anticipated discounted impacts are heterogenous.

Next consider the implications of these data for participant rationality. Assume a common discount rate and constant returns per period from the program. Suppose that persons apply and are accepted into the program if $E(\Delta|I) > 0$, where I is the information available at application. Suppose further that Δ is revealed at the time of acceptance into the program, that persons drop out whenever $\Delta \leq 0$, and that $\Delta \perp\!\!\!\perp Y^0$ (this is identifying assumption (I-5)). If persons entering the program cannot forecast Δ, then letting $e = 1$ if a person enrolls in the program and $e = 0$ if the person drops out, $E(Y^0|e = 1, D = 1) = E(Y^0|e = 0, D = 1)$ and $E(Y^1|e = 1, D = 1) = E(Y^0|e = 0, D = 1) + E(\Delta|\Delta > 0, D = 1)$.

Table 12 presents the mean earnings of JTPA enrollees and dropouts in the 18 months after random assignment for the four demographic groups, along with the experimental impact estimates and the implied differences in Y^0 between the two groups. For adult females and for female youth, the data are consistent with this model, since the difference between the mean earnings of enrollees and dropouts is not statistically distinguishable

[32] Heckman and Smith (1993, 1995) discuss this phenomenon.

[33] This is discussed in Heckman, Smith and Taber (1997).

from the experimental impact estimate. However, the data for adult males and male youth are not consistent with this model.

If, however, we relax the assumption of independence between Δ and Y^0 we can rationalize the male data. Suppose that $\Delta = \Delta(Y^0)$. If $\Delta(Y^0)$ an increasing function of Y^0, this implies that
$E(Y^1|e=1, D=1) - E(\Delta|\Delta > 0, D=1) = E(Y^0|e=1, D=1) > E(Y^0|e=0, D=1)$, which is consistent with the patterns in Table 12 for adult males and for male youth.

Another model assumes that the true treatment effect is revealed after random assignment and the net response varies over time. In this case, a person who values only the outcomes from the program will remain in it if

$$\sum_{t=1}^{T} \delta^{t-1} \Delta_t > 0,$$

where Δ_t is the outcome in the t^{th} period after random assignment and δ is a discount rate. If the inequality is reversed, or becomes an equality, the person drop out.

The implications of this model depend on the temporal pattern of the Δ_t's. For example in classroom training, where the trainee forgoes earnings initially in order to invest in human capital, we would expect $\Delta_{t'} < 0$ for $t' \leq t$, and $\Delta_{t'} > 0$ for $t' > t$, where $t' \leq t$ are periods of human capital accumulation. In the case of a constant Δ, there would be perfect sorting by discount rate into the dropout and enrollee categories. Persons with low δ's would drop out while those with high δ's would complete the training.

We calculate the interest rate r (where $\delta = \frac{1}{1+r}$) required to equate the discounted present value of mean earnings in the dropout and enrollment states for persons in the classroom training treatment stream under two sets of assumptions about the time pattern of impacts more than 18 months after random assignment. The top panel of Table 13 shows the interest rate necessary to equalize the present value of dropout and enrollee earnings if the impact falls to zero after 18 months. That these estimated rates are sometimes negative reflects the fact that the returns to training for some groups are insufficient to balance out the earnings loss in the initial period unless there is negative time preference. The lower panel of Table 13 shows the interest rate necessary to equate the present value of dropout and enrollee earnings under the assumption that the impact in the final six month period persists through seven years.

Potential trainees exhibit high rates of time preference. Discount rates of this magnitude are reported by Thaler (1992). Such high rates are consistent with the view that the poor, who are the primary target of the JTPA program, are poor because they discount the future heavily.

(D) Summary of the Evidence on Impact Heterogeneity and Its Consequences

Table 14 presents a summary of the main findings of this section. (1) Under a variety of assumptions, we find evidence of heterogeneity in net impacts, Δ. (2) The analysis of self-assessments suggests that respondents are reporting different impacts from the "objective" impacts determined from experimental data. This is a further source of heterogeneity and a source of disparity across studies. (3) Departures from high levels of positive dependence between Y^0 and Y^1 produce absurd ranges of impacts on gross outcomes. (The implicit correlations between Y^0 and Y^1 produced under different identifying assumptions are given in the last column of the table). (4) The range of the estimated proportion of people benefiting from the program in the sense of gross outcomes (the "voting criterion") varies widely under different assumptions about the dependence in outcomes. The data from the self-report and attrition studies show a lower proportion benefiting - a phenomenon consistent with the hypothesis that net returns and not gross returns are being reported by participants.

(5) Summary

In his Nobel lecture (1970), Ragnar Frisch recognized the diversity of preferences regarding the outcomes of public policies that characterize participants in welfare states. This diversity in values gives rise to a multiplicity of criteria for evaluating policies. This paper has considered these criteria and presents formal analysis of the information required to evaluate public policies under different criteria. We present the approximations required to go from microeconomic evaluations to conclusions about the general equilibrium outcomes of alternative policies. We provide conditions under which conventional econometric analyses of "treatment effects" provide part of the information required to conduct general equilibrium cost-benefit analyses. We note that personal evaluations of policies may not coincide with the evaluations useful in the political arena of the welfare state and present methods to reveal private or "subjective" evaluations to supplement and complement the "objective" evaluations..

To implement many of the criteria used to evaluate the welfare state requires information on the joint distribution of outcomes across policies. Traditional cost-benefit analysis avoids this problem by assuming that a background social welfare function automatically solves all of the distributional problems of the welfare state. In this case, which is assumed in much of the micro-econometric evaluation literature, simple per capita measures of economic efficiency based on the change in aggregate output attributable to a policy suffice to evaluate the welfare state. However, even in this case we note that estimators widely used in the econometric evaluation literature do not provide the ingredients required for a comprehensive cost-benefit analysis. In an empirical analysis, we demonstrate that when conventional estimators are modified to account for direct costs and the welfare costs of taxation, they produce very different inferences about program impacts than are produced using standard econometric methods. We present conditions under which standard econometric estimators provide reliable answers to well-posed general equilibrium evaluation questions.

Homogeneity in the response to a policy across persons with the same observed characteristics is the central implicit identifying assumption that underlies most widely-used econometric policy evaluation methods. The assumption of response homogeneity greatly simplifies the evaluation problem. Part of the conflict in the estimates produced from different evaluation criteria arises from heterogeneity in impacts of the same program across persons. We present evidence from a major social experiment that heterogeneity in response to treatment is an empirically important phenomenon.

An evaluation strategy that properly accounts for individual heterogeneity requires more information than traditional econometric evaluation methods. We demonstrate how information about participant self-selection choices and program participation rules aids in identifying the distributions of outcomes across policies and also provides information on personal valuations of program outcomes. We discuss how social experiments and different types of micro data can be used to identify the criteria considered in this paper and how they can be supplemented with additional behavioral and statistical assumptions to construct all of the criteria. Unless special individual decision rules characterize program participation, these sources of data do not resolve the fundamental evaluation problem that persons cannot occupy mutually outcome states at the same time.

We apply some of the methods developed in this paper to data from a major job training program. For adult females, we conclude that the program benefited most participants according to the "objective" evaluation criteria based on gross outcomes, but did not benefit a majority of participants according to self-assessments or the revealed preference behavior of attriters from the program. The disagreement among the alternative criteria highlights the need for providing information about all of them satisfy the different parties in the welfare state.

References

[1] Ashenfelter, O. (1978): "Estimating The Effect of Training Programs on Earnings," **Review of Economics and Statistics**, 60(1), 47-57.

[2] Barros, R. (1987): "Two Essays on The Nonparametric Estimation of Economic Models with Selectivity Using Choice-Based Samples," Ph.D. dissertation, Department of Economics, University of Chicago.

[3] Baumol, W. and Quandt, R. (1966): "The Demand For Abstract Transport Modes: Theory and Measurement," **Journal of Regional Science**, 6, 13-26.

[4] Bloom, H., L. Orr, G. Cave, S. Bell, and F. Doolittle (1993): *The National JTPA Study: Title II-A Impacts on Earnings and Employment at 18 Months.* Bethesda, MD: Abt Associates.

[5] Boadway, R. and Bruce, N. (1984): *Welfare Economics.* Oxford, England: Basil Blackwell.

[6] Browning, E.K. (1987): "On The Marginal Welfare Cost of Taxation," **American Economic Review**, 77(1), 11-23.

[7] Browning, M., Hansen, L, Heckman, J. and Taber, C. (1997): "Estimating Dynamic General Equilibrium Models," forthcoming in *Handbook of Macroeconomics*, ed. by J. Taylor and M. Woodford. Amsterdam: North Holland.

[8] Cambanis, S., G. Simons, and W. Stout (1976): "Inequalities for $E(k(X,Y))$ When the Marginals Are Fixed," **Zeitschrift Fur Wahrscheinlichskeitstheorie**, 36, 285-294.

[9] Cameron, S. and J. Heckman (1991): "A Discrete Factor Structure Model For Dynamic Discrete Choice With Continuous Outcomes," unpublished manuscript, University of Chicago.

[10] ____ (1991): "Dynamic Models For Panel Data," unpublished manuscript, University of Chicago.

[11] Chipman, J. and J. Moore (1976): "Why An Increase in GNP Need Not Imply An Improvement in Potential Welfare," **Kyklos**, 29, 391-418.

[12] Cosslett, S. (1983): "Distribution-Free Maximum Likelihood Estimators of the Binary Choice Model," **Econometrica**, 51, 765-872.

[13] Couch, K. (1993): "New Evidence on the Long-Term Effects of Employment Training Programs," **Journal of Labor Economics**, 10, 380-388.

[14] Csörgo (1993): *Quantile Processes with Statistical Applications.* Philadelphia: Society for Industrial and Applied Mathematics.

[15] Domencich, T. and D. McFadden (1975): *Urban Travel Demand: A Behavioral Analysis.* North Holland: Amsterdam.

[16] Frechet, M. (1951): "Sur Les Tableux de Correlation Dont Les Marges Sont Donnes," **Ann. University Lyon**: **Sect. A**, 14, 53-77.

[17] Frisch, R. (1934). "Circulation Planning: Proposal for a National Organization of a Commodity and Service Exchange," **Econometrica**, 2, 274-290.

[18] ----- (1970): "From Utopian Theory to Practical Applications: The Case of Econometrics," **Les Prix Nobel en 1969**. Stockholm.

[19] Harberger, A. (1971): "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay," **Journal of Economic Literature**, 9, 785-797.

[20] Harsanyi, J. (1955): "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility," **Journal of Political Economy**, 63, 309-321.

[21] ----- (1975): "Can the Maximun Principle Serve as a Basis for Morality? A Critique of John Rawls' Theory," **American Political Science Review**, 69(2), 594-606.

[22] Heckman, J. (1974a): "The Effect of Child Care Programs on Women's Work Effort," **Journal of Political Economy,** 82(2), 5136-5163. Reprinted in *Economics of the Family: Marriage, Children, and Human Capital,* ed. by T.W. Schultz. Chicago: University of Chicago Press.

[23] ----- (1974b): "Shadow Prices, Market Wages and Labor Supply," **Econometrica,** 42(4) 679-94.

[24] ----- (1978): "Dummy Endogenous Variables In A Simultaneous Equation System," **Econometrica**, 46(4), 931-959.

[25] ----- (1990a): "Varieties of Selection Bias," **American Economic Review**, 80(2), 313-318.

[26] —— (1990b): "Alternative Approaches To The Evaluation of Social Programs: Econometric and Experimental Methods," Barcelona Lecture, World Congress of the Econometric Society.

[27] —— (1992): "Randomization and Social Program Evaluation," in *Evaluating Welfare and Training Programs*, ed. by C. Manski and I. Garfinkel. Boston: Harvard University Press,201-230.

[28] —— (1993): "The Case for Simple Estimators," Mimeo, University of Chicago.

[29] —— (1996): "Randomization As An Instrumental Variable," **Review of Economics and Statistics**, 77(2), 336-341.

[30] —— (1997a): "Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely-Used Estimator Used in Making Program Evaluations," forthcoming in **Journal of Human Resources**.

[31] —— (1997b): "Constructing Econometric Counterfactuals Under Different Assumptions," Mimeo, University of Chicago.

[32] Heckman, J., N. Hohmann, M. Khoo and J. Smith (1997): "Did We Learn the Right Lesson from the National JTPA Study? Substitution Bias in Social Experiments," Mimeo, University of Chicago.

[33] Heckman, J. and B. Honoré (1990): "The Empirical Content of the Roy Model," **Econometrica**, 58(5), 1121-1149.

[34] Heckman, J., H. Ichimura, J. Smith and P. Todd. (1994): "Characterizing Selection Bias Using Experimental Data," Mimeo, University of Chicago, under revision at **Econometrica**.

[35] —— (1996): "Sources of Selection Bias in Evaluating Programs: An Interpretation of Conventional Measures and Evidence on The Effectiveness of Matching As A Program Evaluation Method," **Proceedings of The National Academy of Sciences**, 93(23), 13416-13420.

[36] Heckman, J., H. Ichimura and P. Todd (1997a): "Matching As An Econometric Evaluation Estimator: Evidence on Its Performance Applied To The JTPA Program, Part I. Theory and Methods," forthcoming in **Review of Economic Studies**.

[37] ——— (1997b): "Matching As An Econometric Estimator: Evidence on Its Performance Applied to the JTPA Program, Part II. Empirical Evidence," forthcoming in **Review of Economic Studies**.

[38] Heckman, J. and R. Robb (1985): "Alternative Methods For Evaluating The Impact of Interventions," in *Longitudinal Analysis of Labor Market Data,* ed. by J. Heckman and B. Singer. New York: Cambridge University Press, 156-245.

[39] Heckman, J., R. Robb, and J. Walker (1990): "Testing the Mixture of Exponentials Hypothesis and Estimating the Mixing Distribution by the Method of Moments," **Journal of the American Statistical Association**, 85, 582-589.

[40] Heckman, J. and J. Smith (1993): "Assessing The Case For Randomized Evaluation of Social Programs," in *Measuring Labour Market Measures: Evaluating the Effects of Active Labour Market Policies*, ed. by K. Jensen and P. K. Madsen. Copenhagen: Danish Ministry of Labor, 35-96.

[41] ——— (1995): "Assessing The Case For Social Experiments," **Journal of Economic Perspectives**, 9, 85-110.

[42] ——— (1996): "Experimental and Nonexperimental Evaluation," in *International Handbook of Labor Market Policy and Evaluation*, ed. by G. Schmidt, J.O' Reilly and K. Schömann. Cheltenham, U.K: Elgar Publishers.

[43] Heckman, J., J. Smith and N. Clements (1997): "Making The Most Out of Program Evaluations and Social Experiments: Accounting for Heterogeneity in Program Impacts," forthcoming in **Review of Economic Studies.**

[44] Heckman, J., J. Smith, and C. Taber (1997): "Accounting For Dropouts in Evaluations of Social Programs," forthcoming in **Review of Economics and Statistics**.

[45] Heckman, J. and C. Taber (1994): "Econometric Mixture Models and More General Models For Unobservables in Duration Analysis," Mimeo, University of Chicago.

[46] Imbens, G. and J. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," **Econometrica**, 62, 467-471.

[47] Katz, D., B. Gutek, R. Kahn and E. Barton (1975): *Bureaucratic Encounters: A Pilot Study in the Evaluation of Government Services.* Ann Arbor, MI: Institute for Social Research.

[48] Kendall, M. G. and Stuart, A. (1977): *The Advanced Theory of Statistics, Vol 1, Fourth Edition.* London: Griffen.

[49] Knight, F. (1921): *Risk, Uncertainty and Profit.* New York: Houghton Mifflin Company.

[50] Laffont, J. J. (1989): *Fundamentals of Public Economics.* Cambridge, MA: MIT Press.

[51] LaLonde, R. (1986): "Evaluating The Econometric Evaluation of Training Programs with Experimental Data," **American Economic Review**, 76(4), 604-620.

[52] Lancaster, K. (1971): *Consumer Demand: A New Approach.* New York: Columbia University.

[53] Lucas, R. and T. Sargent (1981): "Introduction," in *Essays on Rational Expectations and Econometric Practice.* Minneapolis: University of Minnesota Press, xi-xl.

[54] Mardia, K.V. (1970): *Families of Bivariate Distributions.* London: Griffen.

[55] Marschak, J. (1953): "Economic Measurements For Policy and Prediction," in *Studies in Econometric Method,* ed. by W. Hood and T. Koopmans. New York: John Wiley, 1-26.

[56] Matzkin, R. (1990): "Least Concavity and the Distribution-Free Estimation of Nonparametric Concave Functions," Cowles Discussion Paper.

[57] ——— (1992): "Nonparametric and Distribution-Free Estimation of the Threshold Crossing and Binary Choice Models," **Econometrica**, 60, 239-270.

[58] Moffitt, R. (1992): "Evaluation of Program Entry Effects," in *Evaluating Welfare and Training Programs,* ed. by C. Manski and I. Garfinkel. Boston: Harvard University Press, 231-252.

[59] Moulin, H. (1983): *The Strategy of Social Choice.* Amsterdam: North Holland.

[60] Orr, L., H. Bloom, S. Bell, W. Lin, G. Cave, F. Doolittle (1995): *The National JTPA Study: Impacts, Benefits and Costs of Title II-A.* Bethesda, MD: Abt Associates.

[61] Quandt, Richard (1972): "Methods For Estimating Switching Regressions," **Journal of the American Statistical Association**, 67(338), 306-310.

[62] ——— (1988): *The Econometrics of Disequilibrium.* Oxford, Blackwell.

[63] Rawls, J. (1971): *A Theory of Justice.* Cambridge, MA: Harvard University Press.

[64] Ridder, G. (1990): "Attrition in Multiwave Panel Data," in *Panel Data and Labor Market Studies,* ed. by J. Hartog, G. Ridder and J. Theeuwes. Amsterdam: North Holland, 45-67.

[65] Roy, A. (1951): "Some Thoughts on the Distribution of Earnings," **Oxford Economic Papers**, 3, 135-146.

[66] Royden, H. L. (1968): *Real Analysis, Second Edition.* New York: MacMillan Press.

[67] Sen, Amartya (1973): *On Economic Inequality.* Oxford: Clarendon Press.

[68] ______ (1979): "Strategies and Revelation: Informational Constraints in Public Decisions," in *Aggregation and Revelation of Preferences*, ed. by J. J. Laffont. Amsterdam: North Holland.

[69] Smith, J. (1997): "The JTPA Selection Process: A Descriptive Analysis," in *Performance Standards in a Federal Bureaucracy: Analytical Essays on the JTPA Performance Standards System*, ed. by J. Heckman. Kalamazoo, MI: W.E. Upjohn Institute.

[70] Tchen, A. (1980): "Inequalities for Distributions With Given Marginals," **Annals of Probability**, 8, 814-827.

[71] Thaler, R. (1992): *The Winner's Curse.* Princeton: Princeton University Press.

[72] Tinbergen, J. (1956): *Economic Policy: Principles and Design.* Amsterdam: North Holland.

[73] Tong, Y. L. (1980): *Probability Inequalities in Multivariate Distributions.* New York: Academic Press.

[74] U.S. General Accounting Office (1996): "Job Training Partnership Act: Long-Term Earnings and Employment Outcomes." GAO/HEHS-96-40.

[75] Vickrey, W. (1945): "Measuring Marginal Utility By Reactions To Risk," **Econometrica,** 13, 319-333.

Table 1A

Population Data Requirements To Implement Criterion
General Population (Compulsory Programs)

Program j compared to program k

	Cost Benefit	Benthamite Criterion	General Social Welfare Function with Interdependent Preferences	Selfish Voting
Criterion	$E(Y_j) - E(Y_k) \geq 0$	$E(U(Y_j,\theta))-E(U(Y_k,\theta))\geq 0$ $E(U(Y_\ell,\theta)) = \int U(y_\ell,\theta)dF(y_\ell,\theta)$ $\ell = j,k$	$W(j) > W(k)$ $W(\ell) = W(U_1(Y_{1\ell},...,Y_{N\ell}),...U_N(Y_{1\ell},...,Y_{N\ell}))$ $\ell = j,k$**	$\int 1(U(y_j,\theta)\geq U(y_k,\theta)) dF(y_j,y_k,\theta) \geq 0$
Require	Population Means $E(Y_j)$, $E(Y_k)$	$U(Y_\ell,\theta)$ and distribution of (Y_ℓ,θ) $F(y_\ell,\theta)$ $\ell = j,k$	Need each $U_i(Y_{1\ell},...,Y_{N\ell})$ for all i. Need outcomes for each person*	Need $U(Y,\theta)$, $F(y_j,y_k,\theta)$***
Estimable on Aggregate Time Series Data?	Yes, if data exist on aggregate economy in both regimes and can eliminate trend	No, unless θ the same for everyone (homogeneity); $U(Y_\ell,\theta)$ known and the moment $\int U(y_\ell,\theta)dF(y_\ell)$ known or estimable $\ell = j,k$	No, except in the special cases previously considered.	No

*In special cases, summary statistics of the distribution of Y may suffice.
**This includes the special case where individual utility depends only on individual consumption.
***For altruistic voting, U depends on $Y_{1j},....,Y_{Nj}$ or various sub aggregators.

Table 1B

Population Data Requirements To Implement Criterion
(Voluntary Programs: Conditional on $D_j = 1$)

Program j compared to program k

	Cost Benefit	Benthamite Criterion	General Social Welfare function with Interdependent Preferences*	Selfish Voting
Criterion	$E(Y_j \mid D_j=1)-E(Y_k \mid D_j=1)$ What j participants gain over state k	$E(U(Y_j,\theta) \mid D_j=1)) - E(U(Y_k,\theta) \mid D_j=1)) \geq 0$ where $E(U(Y_\ell,\theta) \mid D_j=1) = \int U(y_\ell,\theta) dF(y_\ell,\theta \mid D_j=1)$ $\ell = j,k$	$W(j) > W(k)$ $W(\ell) = W(U_1(Y_{1\ell},\ldots,Y_{N\ell}),\ldots, U_N(Y_{1\ell},\ldots,Y_{N\ell}))$ $\ell = j,k$	$\int 1(U(y_j,\theta) \geq U(y_k,\theta))\, dF(y_j,y_k,\theta \mid D_j=1) \geq 0$
Require	Population Conditional Means $E(Y_j \mid D_j=1)$, $E(Y_k \mid D_j=1)$	$U(Y_\ell,\theta)$ for $D_j=1$, $\ell = j,k$	Need each $U_i(Y_{1\ell},\ldots,Y_{N\ell})$ for all i for whom D=1 Need identity of outcomes for each person*	Need $U(Y,\theta)$, $F(y_j,y_k,\theta) \mid D_j=1)$
Estimable on Aggregate Time Series Data?	Yes, if aggregate data for participants exist in both regimes, can eliminate trend.	No, unless θ the same for everyone (homogeneity); $U(Y_\ell,\theta)$ known and the moment $\int U(y_\ell,\theta) dF(y_\ell \mid D_j=1)$ is known or estimable $\ell = j,k$	No, except in the special cases previously considered.	No

*This criterion is not well defined when restricted to subsets of the population. If only the utility of voluntary participants is considered, some position about the utility of nonparticipants must be taken, and the feedback between participants and nonparticipants must be explicitly modeled. When individual utility only depends on individual consumption, the criterion is well defined.

Table 2

What Cross Section Data Nonparametrically Identify From
$F(y^1 \mid D=1,X)$ and $F(y^0 \mid D=0,X)$
$Pr(D=1 \mid X)$
Under Different Assumptions

Assumption	Recovers	Behavioral Assumption	Criteria Recovered
(1) $E(Y^0 \mid D=1,X) = E(Y^0 \mid X) = E(Y^0 \mid D=0,X)$ (Conditional Mean Independence)	$E(Y^1 - Y^0 \mid D=1,X) = E(Y^1 - Y^0 \mid X)$	Y^0 does not determine participation conditional on X	Cost-Benefit (For D = 1 and total population)
(2): $Y^0 \perp\!\!\!\perp D \mid X$ (Conditional Independence)	(1) plus $F(y^0 \mid D = 1,X) = F(y^0 \mid X)$ $= F(y^0 \mid D = 0,X)$	Y^0 does not determine participation conditional on X	Cost Benefit (For D = 1 and total population) Benthamite (for D = 1)
(3): $(Y^0,Y^1) \perp\!\!\!\perp D \mid X$ (Conditional Independence)	(2) plus $F(y^1 \mid D = 1,X) = F(y^1 \mid X)$ $= F(y^1 \mid D = 0,X)$	(Y^0, Y^1) do not determine participation conditional on X	Cost Benefit and Benthamite (For D = 1 and total population)
(4): (3) plus $\Delta \perp\!\!\!\perp Y_0 \mid X$, $\Delta = Y_1 - Y_0$	(3) plus $F(y^0 \mid X)$	(Y^0, Y^1) do not determine participation conditional on X	All criteria
(5): (3) plus quantile assignment rule $q_1 = \varphi(q_0)$ (φ may be deterministic or random)	Same as (4)	(Y^0, Y^1) do not determine participation conditional on X	All criteria
(6) $E(U_0 \mid X,Z) = 0$ and existence of an exclusion restriction (Z doesn't appear in outcome equation) and $E(U_1-U_0 \mid X,Z,D=1)=E(U_1-U_0 \mid X,D=1)$*	Same as (1)	$\Delta=Y^1 - Y^0$ doesn't enter agents decision rule and exclusion restriction (Z does not appear in outcome equation)	Same as 1
(7) $D = 1(Y^1 > Y^0)$; separability of outcome equations and support condition in Theorem A-1	Same as (4)	Participation solely in terms of (Y^0,Y^1,X). (Model unobservables are U^0,U^1,X)	All criteria
(8) D determined by more general decision rule (Separability, independence and support conditions in Theorem A-2)	$F(y^1,D \mid X)$, $F(y^0,D \mid X)$ for D=1 and D=0.	Y^0,Y^1,X and other factors may determine participation	Cost Benefit and Benthamite (for D = 1 and total population)

*If the IV assumption is interpreted to be $(U^0, U^1) \perp\!\!\!\perp D \mid X,Z$, or $(Y_0, Y_1) \perp\!\!\!\perp D \mid X,Z$ IV is just matching conditional on X,Z and line 3 applies for this conditioning set.

Table 3

Randomization At Entry

Assumption	Behavioral Assumption	Recovers	Criteria Recovered Without Further Assumptions
No Randomization or Substitution Bias	Same as column 1	$F(y^1 \mid D=1,X)$, $F(y^0 \mid D=1,X)$	Cost-Benefit and Benthamite Criteria (For participants (D = 1))
		Randomization at Eligibility	
No Randomization or Substitution Bias	Same as column 1	$F(y^0 \mid X)$, $F(y^1 \mid D=1,X)$ $F(y^0 \mid D=1,X)$, $Pr(D=1 \mid X)$	Cost Benefit and Benthamite Criteria (For participants (D = 1))

Table 4

Panel Data Main Cases For Distributions

Assumption	Behavioral Assumption	Recovers	Criteria Recovered
(1): $F(y_{t'}^{0}, y_{t}^{1} \mid X) = F(y_{t}^{0}, y_{t}^{1} \mid X)$ $y_{t'}^{0} = y_{t}^{0}$ *and*	Stationarity or Sufficient Information To Adjust To Stationarity	Nothing by Itself (Has no empirical counterpart)	None
$F(y_{t}^{0}, y_{t'}^{1} \mid X) = F(y_{t}^{0}, y_{t}^{1} \mid X)$ $y_{t}^{1} = y_{t'}^{1}$	Stationarity plus reversibility	Nothing by Itself (Has no empirical counterpart)	None
(2):(1) plus $F(y_{t'}^{0}, y_{t}^{1} \mid D_{t'}=0, D_{t}=1, X) = F(y_{t}^{0}, y_{t}^{1} \mid D_{t}=1, X)$ $y_{t'}^{0} = y_{t}^{0}$	Among participants in t, nonparticipants in t′ not different from participants in t′. Consistent with both irreversibility and reversibility of program	$F(y_{t}^{0}, y_{t}^{1} \mid D_{t}=1, X)^{*}$	All criteria conditional on $D_t = 1$
(3): (2) plus $F(y_{t}^{0}, y_{t'}^{1} \mid D_{t'}=1, D_{t}=0, X) = F(y_{t}^{0}, y_{t}^{1} \mid D_{t}=0, X)$ $y_{t'}^{1} = y_{t}^{1}$	Among nonparticipants in t, participants in t′ not different from nonparticipants in t′. Consistent with reversibility of program for general t.	$F(y_{t}^{0}, y_{t}^{1} \mid X)$	All criteria

*We can identify the full distribution using the Theorem (A-3) proved in Appendix A. Thus if there is a limit set, $\Pr(D=1 \mid X)=0$ for $X \varepsilon S$, where S is a limit set we can identify $F(y_{t}^{0},(y_{t}^{1} \mid X)$ in that limit set. If there are exclusion restrictions, we can identify this distribution everywhere with additional structure given in Theorem (A-3).

Table 5

Panel Data and Repeated Cross Sections
Means and Marginal Distributions

Assumption	Behavioral Assumption	Recovers	Criteria Recovered
1(a) $E(Y^0_{t'} \mid X) = E(Y^0_t \mid X)$	Stationarity in Means or Sufficient Information to Adjust to Stationarity Holds in Irreversible Case.	Nothing By Itself	None
1b) $E(Y^1_{t'} \mid X) = E(Y^1_t \mid X)$	Same plus Holds in Reversible Case.	Nothing by Itself	None
2(a): (1a) plus $E(Y^0_{t'} \mid D_{t'}=0, D_t=1, X) = E(Y^0_t \mid D_t=1, X)$	In mean, among participants in t, nonparticipants in t′ not different from participants in t′. Holds In Irreversible Case.	$E(Y^0_t \mid D_t = 1, X)$	Cost benefit for $D_t=1$
2(b):(1b) plus $E(Y^1_{t'} \mid D_{t'}=1, D_t=0, X) = E(Y^1_t \mid D_t=0, X)$	In mean, among nonparticipants in t, participants in t′ not different from nonparticipants in t′. Holds in Irreversible Case.	$E(Y^1_t \mid D_t = 0, X)$	Cost benefit for $D_t = 0$
(3)(a) $F(y^0_{t'} \mid X) = F(y^0_t \mid X)$	Stationarity in Marginal Distribution or Sufficient Information to Adjust to Stationarity. Holds in Reversible Case.	Nothing By Itself	None
(3)(b) $F(y^1_{t'} \mid X) = F(y^1_t \mid X)$	Holds in Irreversible Case.	Nothing by Itself	None
(4)(a):3(a) plus $F(y^0_{t'} \mid D_{t'}=0, D_t=1, X) = F(y^0_t \mid D_t = 1, X)$	In distribution, among participants in t, nonparticipants in t′ not different from participants in t′. Holds Irreversible Case.	$F(y^0_t \mid D_t = 1, X)$	Cost benefit plus Benthamite criterion for $D_t=1$
(4)(b):3(b) plus $F(y^1_{t'} \mid D_{t'}=1, D_t=0, X) = F(y^1_t \mid D_t = 0, X)$	In distribution, among nonparticipants in t, participants in t′ not different from participants in t′. Holds in Reversible Case.	$F(y^1_t \mid D_t = 0, X)$	Cost benefit plus Benthamite criterion for $D_t=0$

TABLE 6

BENEFIT MINUS COST ESTIMATES FOR JTPA UNDER ALTERNATIVE ASSUMPTIONS REGARDING BENEFIT PERSISTENCE, DISCOUNTING AND WELFARE COSTS OF TAXATION

National JTPA Study 30 Month Impact Sample

Benefit Duration	Direct Costs Included?	6 month Interest Rate	Welfare Cost of Taxes	Adult Males	Adult Females	Male Youth	Female Youth
30 months	No	0.000	0.00	1354	1703	-967	136
30 months	Yes	0.000	0.00	523	532	-2922	-1180
30 months	Yes	0.000	0.50	108	-54	-3900	-1838
30 months	Yes	0.025	0.00	433	432	-2859	-1195
30 months	Yes	0.025	0.50	17	-154	-3836	-1853
7 years	No	0.000	0.00	5206	5515	-3843	865
7 years	Yes	0.000	0.00	4375	4344	-5798	-451
7 years	Yes	0.000	0.50	3960	3758	-6775	-1109
7 years	Yes	0.025	0.00	3523	3490	-5166	-610
7 years	Yes	0.025	0.50	3108	2905	-6143	-1268

1. Benefit duration indicates how long the estimated benefits from JTPA are assumed to persist. Actual estimates are used for the first 30 months. For the seven year duration case the average of the benefits in months 18-24 and 25-30 is used for the benefits in each future period.
2. Welfare cost of taxes indicates the additional cost in terms of lost output due to each additional dollar of taxes raised. The value of 0.50 lies in the range suggested by Browning (1987).
3. Estimates are constructed by breaking up the time after random assignment into six month periods. All costs are assumed to paid in the first sixth month period, while benefits are received in each six month period and discounted by the amount indicated for each row of the table.

TABLE 7

CHARACTERISTICS OF THE DISTRIBUTION OF IMPACTS ON EARNINGS IN THE 18 MONTHS AFTER RANDOM ASSIGNMENT AT THE FRECHET BOUNDS

National JTPA Study 18 Month Impact Sample
Adult Females

Statistic	Lower Bound	Upper Bound
Impact Standard Deviation	14968.76	674.50
	(211.08)	(137.53)
Outcome Correlation	-0.760	0.998
	(0.013)	(0.001)
Spearman's ρ	-0.9776	0.9867
	(0.0016)	(0.0013)

1. These estimates differ slightly from those reported for $\tau = 1.0$ and $\tau = -1.0$ in Table 8 because they were obtained using the empirical CDFs calculated at 100 dollar earnings intervals rather than using the percentiles of the two CDFs. See Heckman, Smith and Clements (1997) for details.
2. Bootstrap standard errors in parentheses.

TABLE 8

ESTIMATED PARAMETERS OF THE IMPACT DISTRIBUTION
PERFECT POSITIVE DEPENDENCE, POSITIVE DEPENDENCE WITH τ = 0.95, INDEPENDENCE AND PERFECT NEGATIVE DEPENDENCE CASES

National JTPA Study 18 Month Impact Sample
Adult Females

Statistic	Perfect Positive Dependence (τ = 1.0)	Positive Dependence with τ = 0.95	Independence of Y^1 and Y^0 (τ = 0.0)	Perfect Negative Dependence (τ = -1.0)
5th Percentile	0.00	0.00	-18098.50	-22350.00
	(47.50)	(360.18)	(630.73)	(547.17)
25th Percentile	572.00	125.50	-6043.00	-11755.00
	(232.90)	(124.60)	(300.47)	(411.83)
50th Percentile	864.00	616.00	0.00	580.00
	(269.26)	(280.19)	(163.17)	(389.51)
75th Percentile	966.00	867.00	7388.50	12791.00
	(305.74)	(272.60)	(263.25)	(253.18)
95th Percentile	2003.00	1415.50	19413.25	23351.00
	(543.03)	(391.51)	(423.63)	(341.41)
Percent Positive	100.00	96.00	54.00	52.00
	(1.60)	(3.88)	(1.11)	(0.81)
Impact Std Dev	1857.75	6005.96	12879.21	16432.43
	(480.17)	(776.14)	(259.24)	(265.88)
Outcome Correlation	0.9903	0.7885	-0.0147	-0.6592
	(0.0048)	(0.0402)	(0.0106)	(0.0184)

1. The values in this table are calculated using percentiles of the two distributions. The perfect positive dependence case matches the top percentile in the Y^1 distribution with the top percentile in the Y^0 distribution, the second percentile of the Y^1 distribution with the second of the Y^0 distribution and so on. The perfect negative dependence case matches the percentiles in reverse order, so that the lowest percentile of the Y^0 distribution is matched with the highest percentile of the Y^1 distribution and so on. The two intermediate cases match the percentiles of the Y^1 distribution with percentiles of a permutation of the Y^0 distribution such that the rank correlation of the matched percentiles has the value indicated.

2. The perfect positive and perfect negative dependence cases are based on the single permutation having this characteristic in the sample. The values reported for the intermediate cases represent means of random samples of 50 permutations with the indicated value of τ.

3. For each case, the difference between each percentile of the Y^1 distribution and the associated percentile of the Y^0 distribution is the impact for that percentile. Taken together, the percentile impacts form the distribution of impacts. It is the percentiles of these impact distributions that are reported in the upper portion of the table. The impact standard deviation, outcome correlation, and the percent positive are calculated using the percentile impacts. The impact standard deviation is the standard deviation of the percentile differences. The outcome correlation is the correlation of the matched percentiles from the two distributions. The percent positive is the percent of the percentile impacts greater than or equal to zero.

4. Bootstrap standard errors in parentheses.

TABLE 9

RANDOM COEFFICIENT AND DECONVOLUTION ESTIMATES IMPACT ON EARNINGS IN THE 18 MONTHS AFTER RANDOM ASSIGNMENT

National JTPA Study 18 Month Impact Sample
Adult Females

Analysis	Estimated Mean Impact	Estimated Impact Std Dev	Estimated Percent Positive
Random coefficient model	601.74 (201.63)	2271.00 (1812.90)	60.45
Deconvolution	614.00	1675.00	56.35

1. Estimated standard errors appear in parentheses where available.
2. Random coefficient model includes race/ethnicity, schooling and site indicators. Only the treatment coefficient is treated as random.
3. The estimated impact variance for the random coefficient model is obtained from a regression of the squared residuals from the corresponding fixed coefficient model on the treatment indicator.
4. The estimated percent positive for the random coefficient model assumes that Δ is normally distributed.
5. Mean impact, impact standard deviation and the fraction of positive impacts for the deconvolution case are obtained from the smoothed density. Values for the unsmoothed density differ only slightly from those reported here.

TABLE 10

TESTS OF SECOND ORDER STOCHASTIC DOMINANCE OF EXPERIMENTAL TREATMENT GROUP OVER EXPERIMENTAL CONTROL GROUP EARNINGS IN THE 18 MONTHS AFTER RANDOM ASSIGNMENT

National JTPA Study 18 Month Impact Sample

Earnings Value (α)	Adult Males	Adult Females	Male Youth	Female Youth
2,500	0.8836 (0.3162) [0.0052]	1.0296 (0.2978) [0.0005]	-0.3357 (0.4250) [0.4296]	0.6674 (0.5094) [0.1901]
5,000	1.8067 (0.6582) [0.0061]	1.9343 (0.5955) [0.0012]	-1.0482 (0.9344) [0.2620]	0.7137 (1.0022) [0.4764]
7,500	2.3903 (0.9983) [0.0166]	2.7811 (0.8933) [0.0019]	-1.8742 (1.4610) [0.1995]	0.4428 (1.4507) [0.7602]
10,000	2.9839 (1.3334) [0.0252]	3.7315 (1.1504) [0.0012]	-2.8489 (1.9790) [0.1500]	0.1308 (1.8486) [0.9436]
15,000	4.0191 (1.9826) [0.0435]	5.2659 (1.5768) [0.0008]	-4.0631 (2.8333) [0.1516]	-0.2717 (2.4032) [0.9100]
20,000	4.4428 (2.5434) [0.0807]	6.2660 (1.8551) [0.0007]	-5.8554 (3.5386) [0.0980]	-0.4484 (2.1750) [0.8688]
25,000	4.6171 (2.9192) [0.1137]	7.0279 (1.9980) [0.0004]	-6.3804 (4.0905) [0.1188]	-0.4503 (2.8641) [0.8751]

1. The first value in each cell is the difference (control minus treatment) in the integrated CDFs of 18-month earnings. A positive difference indicates that the treatment group distribution second order stochastically dominates the control group distribution. In calculating the integrals, each 100 dollar increment in earnings is normalized to have length one.
2. Bootstrap standard errors appear in parentheses. Bootstrap standard errors constructed using 100 samples equal in size to the original sample.
3. P-values appear in square brackets.

TABLE 11

SELF-ASSESSMENTS OF JTPA IMPACT
EXPERIMENTAL TREATMENT GROUP

National JTPA Study 18 Month Impact Sample

	Adult Males	**Adult Females**	**Male Youth**	**Female Youth**
Full Sample Percentages				
Percent who self-report participating:	61.63	68.10	62.62	66.29
	(0.81)	(0.68)	(1.29)	(1.09)
Percent of self-reported participants with a positive self-assessment:	62.46	65.21	67.16	71.73
	(1.04)	(0.85)	(1.59)	(1.29)
Overall percent with positive self-assessments:	38.49	44.41	42.06	47.55
	(0.81)	(0.73)	(1.32)	(1.16)
Percent of Self-Reported Participants with a Positive Self-Assessment by Primary Treatment Received				
None (dropouts)	48.89	51.44	58.90	61.56
	(2.07)	(1.85)	(3.33)	(2.79)
Classroom training in occupational skills	74.10	73.47	72.73	75.28
	(2.15)	(1.36)	(3.60)	(2.30)
On-the-job training at private firm	75.13	78.90	71.00	75.00
	(2.18)	(2.14)	(4.56)	(4.04)
Job search assistance	59.57	59.80	68.09	68.94
	(2.27)	(2.18)	(3.94)	(4.04)
Basic education	62.96	56.55	70.97	78.44
	(4.67)	(3.84)	(4.09)	(3.19)
Work experience	66.67	68.75	82.76	73.17
	(9.83)	(5.84)	(7.14)	(7.01)
Other	58.47	66.40	62.50	77.98
	(3.65)	(2.98)	(4.77)	(3.99)

1. Reported proportions are based on responses to the question "Do you think that the training or other assistance you got from the program helped you get a job or perform better on the job?". This question was asked only of self-reported participants within the treatment group. The overall fraction of positive self-assessments assumes that self-reported non-participants would have provided a negative self-assessment.

2. The primary treatment is the one in which the trainee participated for the most hours according to the administrative records of the JTPA sites. Most trainees received only one service, few received more than two. See Smith (1997) for a detailed discussion. Note that for some self-reported participants the JTPA administrative records indicate that no services were received.

3. Estimated standard errors in parentheses.

TABLE 12
COMPARISONS OF POST-RANDOM ASSIGNMENT EARNINGS OF TREATMENT GROUP ENROLLEES AND DROPOUTS

National JTPA Study 18 Month Impact Sample

	Percent Dropping Out	Mean Earnings of Enrollees	Mean Earnings of Dropouts	P-value	Mean Earnings Difference	Experimental Impact Estimate
Adult males	40.92	13638	12181	0.0010	1457	572
	(0.72)	(261)	(355)		(440)	(371)
Adult females	37.39	8428	7947	0.0975	480	793
	(0.66)	(156)	(244)		(290)	(223)
Male youth	34.83	10275	9442	0.1231	832	-801
	(1.14)	(310)	(441)		(539)	(475)
Female youth	37.02	6108	6251	0.6844	-143	-45
	(1.04)	(197)	(291)		(352)	(298)

1. P-values are from t-tests of equality of means assuming unequal variances in the two groups.
2. Mean earnings difference indicates the difference in mean earnings between the enrollees and the dropouts in the experimental treatment group.
3. Experimental impact estimates differ from those in Bloom, et al. (1993) because they are not regression-adjusted and because imputed values for adult female respondents were not used.
4. Estimated standard errors in parentheses.

TABLE 13

SIX MONTH INTEREST RATES AND DISCOUNT FACTORS THAT EQUALIZE THE DISCOUNTED PRESENT VALUE OF MEAN EARNINGS OF ENROLLEES AND DROPOUTS IN THE CT-OS TREATMENT STREAM UNDER TWO ASSUMPTIONS ABOUT BENEFIT DURATION

National JTPA Study 18 Month Impact Sample

Parameter	Adult Males	Adult Females	Male Youth	Female Youth
Benefits Fall to Zero After 18 Months				
Equalizing r	-0.0627	-0.7720	0.2082	-0.2677
Equalizing δ	1.0669	4.3858	0.8277	1.3656
Benefits Continue for 7 Years				
Equalizing r	1.1250	0.8720	1.4520	0.9580
Equalizing δ	0.4706	0.5342	0.4078	0.5107

1. The CT-OS treatment stream refers to persons recommended for Classroom Training in Occupational Skills prior to random assignment in the JTPA experiment. This group comprises roughly one third of the full experimental sample.
2. Negative values for the interest rate indicate that positive future preference is required to rationalize choosing the enrollee earnings stream over the dropout earnings stream.
3. Earnings streams are broken up into six month pieces in calculating the discount rates.

TABLE 14

Summary of Empirical Evidence of Impact Heterogeneity, the Voting Criterion and the Dependence Between Y^1 and Y^0
National JTPA Study 18 Month Experimental Impact Sample

	Description of Analysis	Evidence of Heterogeneity?	Standard Deviation of Impacts	Evidence on Voting Criterion	Dependence Between Y^1 and Y^0
Frechet Bounds	Statistical bounds on the joint distribution of outcomes, $F(y^0, y^1 \mid D=1)$, and on super- and sub-additive functions of the joint distribution.[1] See equation (17) in the text.	Yes, the impact standard deviation is bounded away from zero.[2]	Bounded between $675 and $14969.[2]	The bounds do not apply to the indicator function $1(Y^1 \geq Y^0)$ as this function is not super- or sub-additive. Thus, the voting criterion cannot be bounded.[2]	Product-moment correlation ρ between Y^1 and Y^0 bounded between -0.760 and 0.998. Kendall's rank correlation τ between Y^1 and Y^0 bounded between -0.9997 and 0.9867.[2]
Perfect Positive Percentile Dependence	Assumes $q_1 = q_0$ where q_1 is a percentile of Y^1 and q_0 is a percentile of Y^0. That is, the counterfactual for each percentile in the Y^1 distribution is the same percentile in the Y^0 distribution.	Yes, the impacts vary between $0 and $3250.[2]	$1857.[2]	100 percent of participants benefit or are indifferent.[2]	Product-moment correlation $\rho = 0.9903$ and Kendall's rank correlation is fixed at 1.00. Both are calculated using the percentiles of the two distributions.[2]
Perfect Negative Percentile Dependence	Assumes $q_1 = 100 - q_0$ where q_1 is a percentile of Y^1 given $D=1$ and q_0 is a percentile of Y^0 given $D=1$. That is, the counterfactual for each q^{th} percentile of the Y^1 distribution is the 100-q^{th} percentile in the Y^0 distribution in the population defined by $D=1$.	Yes, the impacts vary between -$48606 and $34102.[2]	$16432.[2]	52 percent positive.[2]	Product-moment correlation $\rho = -0.6592$ and Kendall's rank correlation is fixed at -1.00. Both are calculated using the percentiles of the two distributions.[2]
Positive Percentile Dependence with Rank Correlation $\tau = 0.95$	Assumes that the percentiles of Y^1 and Y^0 given $D=1$ have a rank correlation of 0.95. Estimates are based on a random sample of 50 such permutations.	Yes, the average minimum is -$14504 and the average maximum is $48544.[2]	Average standard deviation of $1857 (with standard deviation of $480).[2]	Average of 93 percent positive (with standard deviation of 3.88).[2]	Average product-moment correlation ρ of 0.7885 (with a standard deviation of 0.0402). Kendall's rank correlation τ fixed at 0.95. Both are calculated using the percentiles of the two distributions.[2]
Independence of Percentiles of Y^1 and Y^0, Which Implies a Percentile Rank Correlation τ of 0.0	Assumes that the percentiles of Y^1 and Y^0 given $D=1$ have a rank correlation τ of 0.0, which is implied by independence between them. Estimates are based on a random sample of 50 such permutations.	Yes, the average minimum is -$44175 while the average maximum is $60599.[2]	Average standard deviation of $12879 (with standard deviation of $259).[2]	Average of 54 percent positive (with standard deviation of 1.11).[2]	Average product-moment ρ of -0.0147 (with standard deviation of 0.0106). Kendall's rank correlation τ fixed at 0.0. Both are calculated using the percentiles of the two distributions.[2]

TABLE 14 (cont.)

Summary of Empirical Evidence of Impact Heterogeneity, the Voting Criterion and the Dependence Between Y^1 and Y^0
National JTPA Study 18 Month Experimental Impact Sample

	Description of Analysis	Evidence of Heterogeneity?	Standard Deviation of Impacts	Evidence on Voting Criterion	Dependence Between Y^1 and Y^0
Random Coefficient Model	Assumes that $\Delta \perp\!\!\!\perp Y_0 \mid D = 1$.	Yes, see Figure 2.[3]	Standard deviation is \$2271.[3]	If the random coefficient Δ is assumed to be normally distributed then 60.45 percent have positive impacts.[3]	The product-moment correlation $\rho = 0.9595$.[3]
Deconvolution	Assumes that $\Delta \perp\!\!\!\perp Y_0 \mid D = 1$.	Yes, see Figure 2.[3]	Standard deviation is \$1675.[3]	56.35 percent positive.[3]	The product-moment correlation $\rho = 0.9774$.[3]
Self-assessments	Ex post self-evaluations by participants were based on a survey question regarding whether or not the program provided a benefit.	Yes, some participants reported a benefit and others did not.	N.A.[4]	Varies from a low of 39.49 percent positive for adult men to a high of 47.55 percent positive for female youth.	N.A.[4]
Dropouts	Attrition decisions after application and acceptance into the program.	Yes. There are non-zero attrition rates and the evidence on the discount rates required to jusitfy a common coefficient model suggests that this model is false.	N.A.[4]	Dropping out ranges from a low of 34.83 percent for male youth to a high of 40.92 percent for adult males.	N.A.[4]

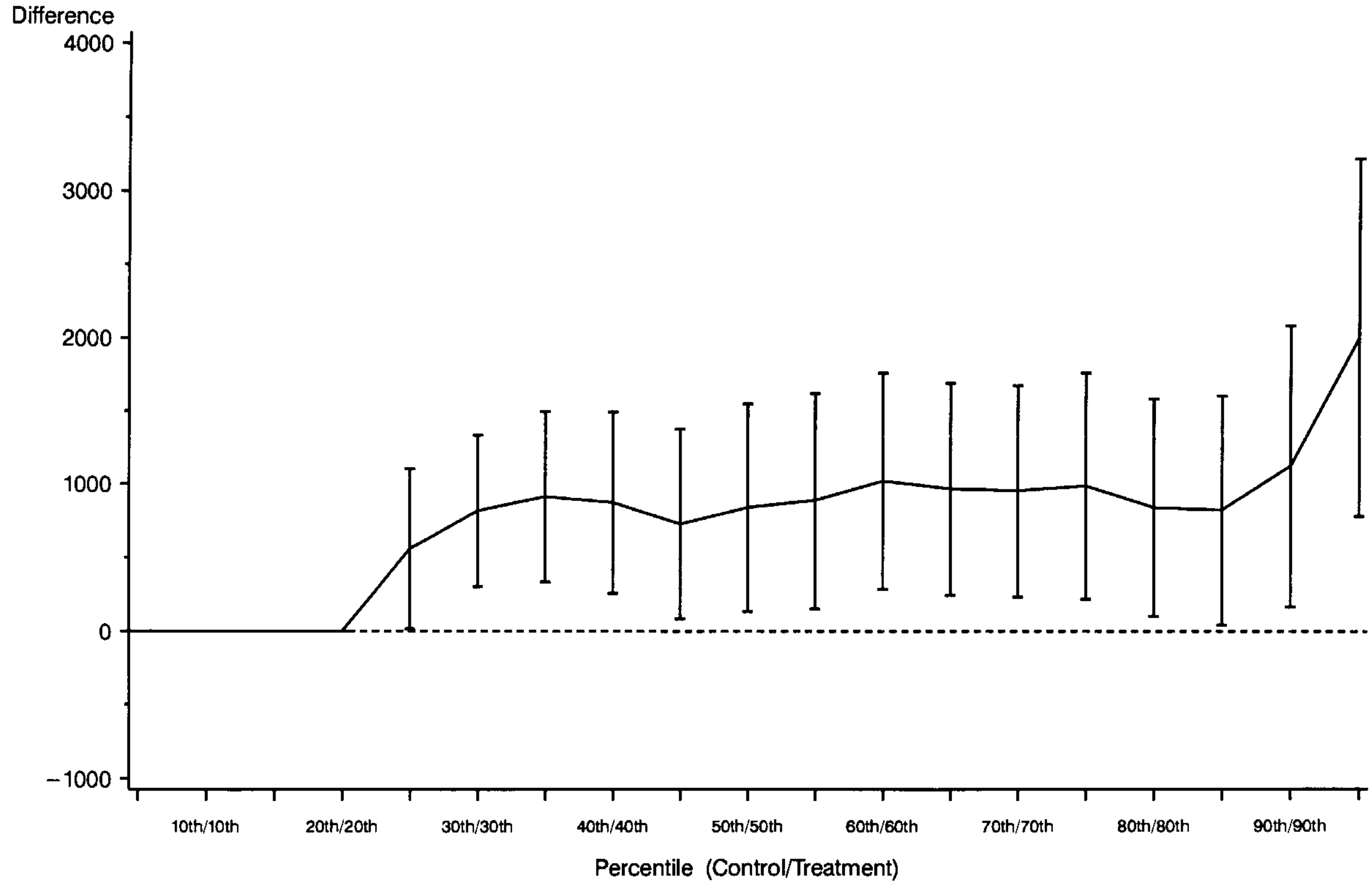
[1]A function $k(x,y)$ is superadditive if $x > x'$ and $y > y'$ implies that $k(x,y) + k(x',y') > k(x,y') + k(x',y)$. Subadditivity reverses the inequality.
[2]Results are for adult women only. Similar results are obtained for adult men and for male and female youth.
[3]Results are for adult women only. For the remaining demographic groups $Var\left(Y^1\right) < Var\left(Y^0\right)$ which indicates that neither the random coefficient model nor deconvolution is appropriate.
[4]N.A. = not applicable.

Figure 1

Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution

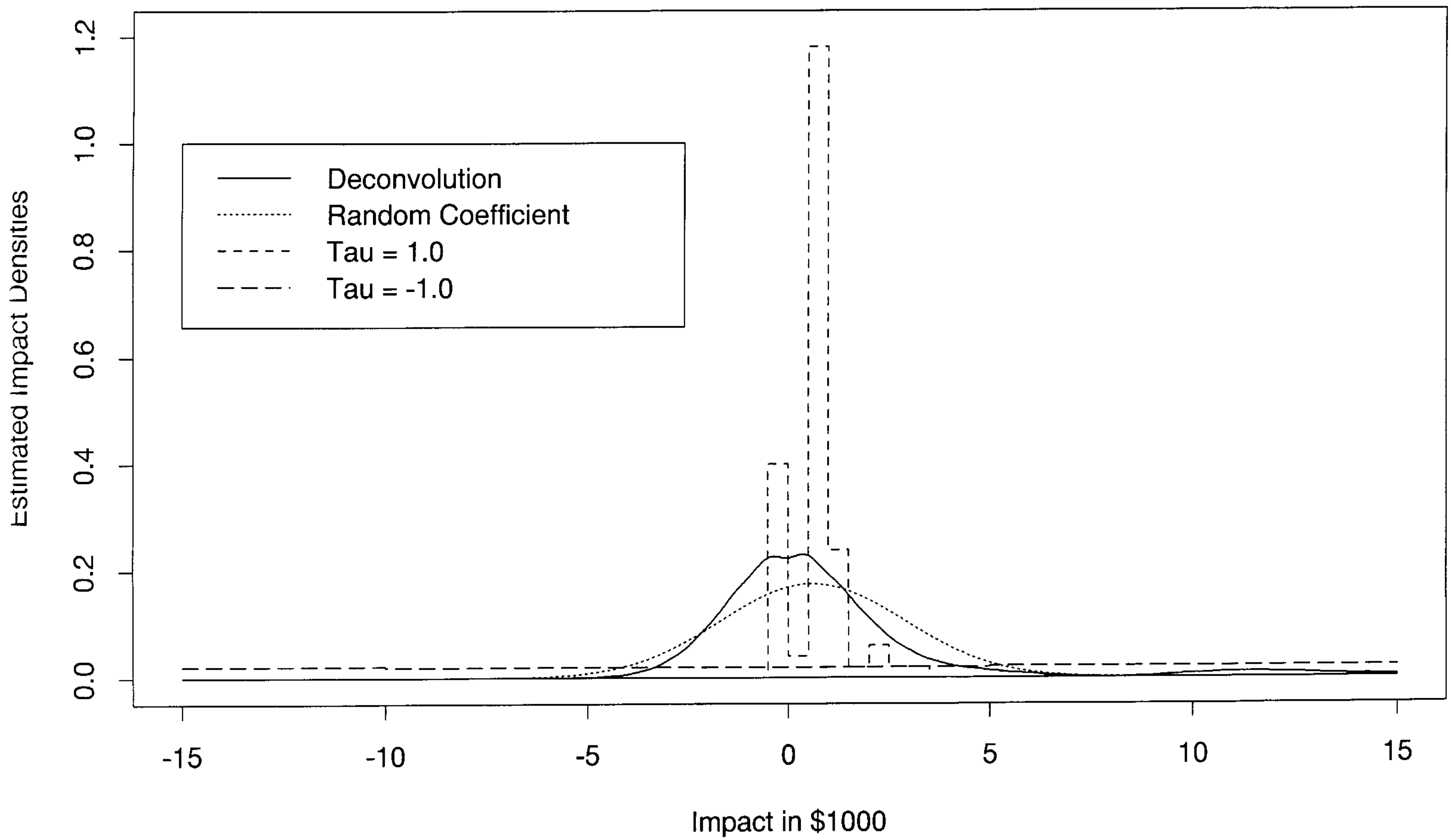
Perfect Positive Dependence Case

Adult Females



1. Sample consists of ABT's experimental 18–month study sample.
2. ABT imputed values were used in place of outlying values.
3. Standard errors for the quantiles are obtained using methods described in Csorgo (1993).

Figure 2 - Impact Densities Under Alternative Identifying Assumptions
Adult Females



Appendix A

Let outcomes Y^1 and Y^0 be written as functions of observed variables X and unobserved variables U_1 and U_0 respectively:

(A-1a) $$Y^1 = g_1(X_1, X_c) + U_1$$

(A-1b) $$Y^0 = g_0(X_0, X_c) + U_0,$$

where X_1 (a k_1-dimensional vector) and X_0 (a k_0-dimensional vector) are variables unique to g_1 and g_0 respectively and X_c (a k_c-dimensional vector) includes variables common to the two functions. The variables U_0 and U_1 are unobserved from the point of view of the econometrician. To simplify the analysis we delete program-specific subscripts.

The decision rule for program participation is given by

(A-1c) $$I = g_I(X_I, X_c) + U_I \text{ and } D = 1(I \geq 0),$$

where X_I (a k_I-dimensional vector) consists of measured variables some of which may appear in X_1 and X_0, and where U_I is unobserved by the econometrician. I is a latent index or net utility of the relevant decision maker. The joint distribution of (U_0, U_1, U_I) is denoted by $F(u_0, u_1, u_I)$. These variables are statistically independent of (X_0, X_1, X_I, X_c).

The Roy model is a special case of this framework in which selection into the program depends only on the gain from the program. In this case

$$g_I = g_1(X_1, X_c) - g_0(X_0, X_c) \text{ and } U_I = U_1 - U_0.$$

For convenience we write $F(u_0, u_1)$ as the joint distribution of (U_0, U_1).

The following theorem can be proved for the Roy model. It extends and clarifies a theorem in Heckman and Honoré (1990).

Theorem A-1: Let $Y^1 = g_1(X_1, X_c) + U_1$ and $Y^0 = g_0(X_0, X_c) + U_0$. Assume

(i) $(U_0, U_1) \perp\!\!\!\perp (X_0, X_1, X_c)$

(ii) $D = 1(Y^1 \geq Y^0)$

(iii) (U_0, U_1) absolutely continuous with $Support(U_0, U_1) = R_1 \times R_1$;

(iv) for each fixed X_c

$$g_0(X_0, X_c) : R_{k_0} \to R_1 \text{ for all } X_1,$$

$$g_1(X_1, X_c) : R_{k_1} \to R_1 \text{ for all } X_0,$$

and $$Support(g_0(X_0, X_c) \,|X_c, X_1) = R_1 \text{ for all } X_c, X_1,$$

$$Support(g_1(X_1, X_c)\,|X_c,\ X_0) = R_1 \text{ for all } X_c, X_0,$$
$$Support(X_0\,|X_1, X_c) = Support\ (X_0) = R_1 \text{ for all } X_1, X_c,$$
$$Support(X_1\,|X_0\ , X_c) = Support\ (X_1) = R_1 \text{ for all } X_0, X_c,$$

(v) The marginal distributions of U_0, U_1 have zero medians.

Then g_0, g_1 and $F(u_0, u_1)$ are nonparametrically identified from data on participation choices and outcomes.

Proof: By assumption, we know for all (X_0, X_1, X_c) in the support of (X_0, X_1, X_c) and for all y

(A) $\Pr(Y^1 \leq Y^0 \mid X_0, X_1, X_c) = \Pr(g_1(X_1, X_c) + U_1 \leq g_0(X_0, X_c) + U_0)$,

(B) $\Pr(Y^1 \leq y, Y^1 > Y^0 \mid X_0, X_1, X_c) =$
$\Pr(g_1(x_1, x_c) + U_1 \leq y, g_1(X_1, X_c) + U_1 > g_0(X_0, X_c) + U_0)$,

(C) $\Pr(Y^0 \leq y, Y^0 \geq Y^1) = \Pr(g_0(X_0, X_c) + U_0 \leq y, g_0(X_0, X_c) + U_0 > g_1(X_1, X_c) + U_1)$.

Fix X_c. Let $\bar{X}_1$ and $\bar{X}_0$ be in the support of X_1 and X_0, respectively. Using the information in (A), we can define sets of values (X_0,X_1) corresponding to contours of constant probability:

(A-2) $S(X_0, X_1 \mid X_c, \overline{p}) =$
$$\{(x_0, x_1) : \Pr(g_1(X_1, X_c) + U_1 > g_0(X_0, X_c) + U_0) = \Pr(g_1(\bar{X}_1, X_c) + U_1 > g_0(\bar{X}_0, X_c) + U_0) = \overline{p}\}$$
$$= \{(X_1, X_0) : g_1(X_1, X_c) + \ell = g_0(X_0, X_c)\}$$
for some unknown constant ℓ.

For any point in S we can use the information in (B) to write

(A-3) $\Pr(g_1(X_1, X_c) + U_1 \leq y_1, U_1 > U_0 + \ell)$

for all y_1. Varying X_1 over its full support from assumption (iv), we can find a compensating value X_0 within the set defined by (A-2) so that $\Pr(D = 1 \mid X_1, X_0, X_c) = \overline{p}$ is constant. This keeps fixed the second argument in (A-3). The variation in X_1 produces a set of (y, X_1) values for each value of X_c which identifies the function $g_1(X_1, X_c)$ over the support of X, up to an unknown constant. By similar reasoning, we can identify $g_0(X_0, X_c)$ up to an unknown constant using (C).

Tracing out (A-3) for all values g_1 and y identifies $F(u_1, u_0 - u_1)$ except for a location parameter. Using (C) we identify $F(u_0, u_1 - u_0)$. The location of U_0 and U_1 is determined by using the assumption that the medians of U_0, U_1 are zero using the marginals obtained by letting $g_0 \to -\infty$ (in B) and $g_1 \to -\infty$ (in C) respectively. (The information in B is actually all we need). With knowledge of the locations of U_0, U_1, we can determine the unknown additive constants absorbed in g_1 and g_0. By a standard transformation of variables, we obtain $F(u_0, u_1)$ from either B or C. Since X_c is arbitrary, this completes the proof because we can recover everything for all X_c. ■

The content of this theorem is that if there is sufficient variation in X_1, X_0 and X_c, and if we know that program participation is based solely on outcome maximization, no arbitrary parametric structure on the outcome equations or on the distribution of the unobservables generating outcomes needs to be imposed to recover the full distribution of outcomes using ordinary micro data.

Note that we can obtain $F(y^0 \mid D = 1, X)$ from $F(y^0 \mid D = 0, X)$, where X denotes the full set of conditioning variables. Using the law of iterated expectations,

$$F(y^0 \,|X\,) = F(y^0 \mid D = 1\,|X\,)\Pr(D = 1\,|X\,) + F(y^0 \mid D = 0\,|X\,)\Pr(D = 0\,|X\,).$$

From Theorem A-1, we can recover $F(y^0 \,|X\,)$. Since we know $\Pr(D = 1\,|X\,) = 1 - \Pr(D = 0\,|X\,)$, we can recover $F(y^0 \mid D = 1, X)$.

The assumptions made in Theorem A-1 about the supports of $X_1, X_0, g_1, g_0, U_1, U_0$ are made for convenience, in an effort to focus on main ideas. Nowhere is it literally required that any function or variable "go to infinity" as some authors have claimed.[1] A version of Theorem A-1 can easily be proved under the following alternative conditions.

$$Support(U_0) = Support(U_1) = Support(g_0) = Support(g_1)$$

where all of the supports are finite. Under these conditions, and assuming all of the other conditions hold, it is possible to retrace the argument of Theorem A-1 and produce essentially the same theorem, provided that $Support(X_0 \,|X_1, .X_c) = Support(X_0)$ and $Support(X_1 \,|X_0,\ X_c) = Support(X_1)$.

In more general cases where the supports of g_0, g_1, U_1 and U_0 do not coincide, or where there are restrictions on the supports of X_1 and X_c, a modified version of the theorem can be proved. It may be possible to construct a set of (X_1, X_0) values that satisfy a condition like (A-2) except now it is no longer necessarily true that we can vary x_1 over its full support within any isoprobability set (the set of values of (X_1, X_0) that set $\Pr(D = 1 \mid X_0, X_1, X_c) = \overline{p}$ for any $\overline{p}$). That is, for each $\overline{p}$, we are no longer guaranteed to be able to find a compensatory value of X_0 to assure that for each X_1, we can keep the probability fixed. Suppose $\underline{g} \leq g_1 \leq \bar{g}_1$ and $\underline{g_0} \leq g_0 \leq \bar{g}_0$. For each X_c and $\overline{p}$, the support

[1]See, *e.g.*, Imbens and Angrist (1994).

of $g_1 - g_0$ is $(\overline{g_1} - \underline{g_0}, \underline{g_1} - \overline{g_0} \mid X_c, \overline{p})$ provided $Support(X_0 \mid X_1, X_c) = Support(X_0)$ and $Support(X_1 \mid X_0, X_c) = Support(X_1)$. Only for subsets of the support of X_0 and X_1 can the argument below (A-2) in the proof be invoked. Because $g_0(X_0, X_c)$ and $g_1(X_1, X_c)$ are not necessarily identified over their full support, it follows that we are not necessarily guaranteed to be able to identify $F(u_0, u_1)$ over its full support. Moreover, in general we will only be able to identify the g_1 and g_0 functions up to unknown scale parameters. With these qualifications about the support, the conclusions of Theorem A-1, restated to include the restrictions on supports, remain intact.

The Roy model has an unusual structure because the participation rule and the outcome equations are tightly linked. As a consequence, we can recover the full joint distribution of $F(y^1, y^0 \mid X)$ and the decision rule knowing only conditional distributions (B) and (C) available from cross-section data. For more general decision rules such as (A-1c), which break the tight link between outcomes and participation decisions, it is not possible to use (B) and (C) to address those questions that can only be answered from the full joint distribution of (Y^1, Y^0). Even access to the data obtained from social experiments - $F(y^0|D = 1, X)$ - does not suffice to solve the fundamental evaluation problem that both Y^0 and Y^1 are never observed for the same person. However, a theorem analogous to Theorem A-1 can be proved that demonstrates that with sufficient variation in the X variables, it is possible to recover $F(y^0|D = 1, X)$ from non-experimental data. Before presenting a more general version of Theorem A-1, it is useful to review some recent results on the estimation of nonparametric and semiparametric discrete choice models that are required in the proof of the theorem.

We consider the nonparametric identification of decision rule (A-1c) under the assumption that

(A-4) $$(X_I, X_c) \perp\!\!\!\perp U_I.$$

The original proof is due to Cosslett (1983) who assumes $g_I = (X_I, X_c)\,\beta_I$. Matzkin (1990, 1992) considers the more general case that will be used here.[2] In the Roy model, (A-1c) was tightly linked to (A-1a) and (A-1b) and we observe Y^1 and Y^0 in censored samples. In the general case, nonparametric identification of g_I requires a separate argument.

From inspection of

$$\Pr(D = 1|X_I, X_c) = \Pr(g_I(X_I, X_c) + U_I \geq 0) = 1 - F_{U_I}(-g_I),$$

it is clear that without further restrictions on the set of candidate g_I functions, it will be impossible to identify a unique member of the set. For any alternative distribution function F^*, we can define g_I^* so that

$$F_{U_I}(-g_I) = F^*_{U_I}(-g_I^*),$$

[2] Heckman and Taber (1994) survey alternative approaches to identifiability in discrete choice and duration models.

with the result that (g_I, F_I) cannot be distinguished from (g_I^*, F^*).

Let G be the set of admissible functions. Matzkin (1990) shows that under the independence assumption, if $\bar{g}_I$ is a least-concave representation of $g_I \in G$, then $\left(\bar{g}_I, \bar{F}_I\right)$ is identified, where $\bar{F}_I$ is the associated distribution function.[3] Since concavity naturally arises in many economic settings of consumer and producer choice, her assumption is an attractive one. We record Matzkin's basic assumptions in addition to (A-4):

(M-1) g_I is concave;

(M-2) $Support(g_I(X_0, X_c)) \supset Support(U_I)$.

Theorem M1 (Matzkin 1): Under (M-1) and (M-2), $\bar{g}_I$ and the associated $\bar{F}_I$ are identified subject to a scale normalization for g_I.

Proof: See Matzkin (1990). ■

Matzkin (1992) also considers an alternative identifying assumption that can substitute for M-1 and M-2.

(M-3) There exists a subset $\bar{T}$ of the support of $X = (X_I, X_c)$ such that (i) for all $g_{1I}, g_{2I} \in G$, and all $x \in \bar{T}, g_{1I}(X) = g_{2I}(X)$ and (ii) for all t in the support of U_I, there exists $X \in \bar{T}$ such that $g_I(X) = t$.

In the estimation of production functions, there is a natural set of values $X = \underset{\sim}{0}$ where no input produces no output. Similarly, for cost functions, $c(0) = 0$ is a natural assumption. Matzkin develops a consistent estimator for g_I and F_I under additional assumptions.

Theorem M2 (Matzkin 2): Under (M-2) and (M-3), and with $X \perp\!\!\!\perp U_I, (g_I, F_I)$ is identified up to a scale normalization for g_I.

Proof: See Matzkin (1992). ■

We use Theorem M-1 or M-2 to claim that we can nonparametrically identify g_I and F_I over the support of the data. Obviously, if

$$0 < \underline{p} \leq \Pr(D = 1 | X_I, X_c) \leq \bar{p} < 1,$$

and the support of g_I is bounded and strictly contained in the support of U_I, we may be able to identify $F_I(U_I)$ only over a subset of its true support. We are now ready to prove

[3] A function $\bar{g}_I$ is a least-concave representation of concave function g_I if for any strictly increasing function h such that $h \ o \ g_I$ is concave, there exists a concave function t, such that $h \ o \ g_I = t \ o \ \bar{g}_I$. Since g_I is a monotonic transformation of $\bar{g}$, g_I and $\bar{g}$ must have the same isovalue sets.

Theorem A-2, which extends and clarifies a result in Heckman (1990a) that generalizes the proof of nonparametric identifiability of the Roy model.

Theorem A-2: Let (U_0, U_1, U_I) be median-zero, independently and identically distributed random variables with distribution $F(u_0, u_1, u_I)$. Assume structure (A-1a) - (A-1c) and knowledge of $F(y^0|D=0, X_0, X_I, X_c)$, $F(y^1|D=1, X_1, X_I, X_c)$ and $Pr(D=1|X_I, X_c)$. Assume:

(a-1) $(U_0, U_I) \perp\!\!\!\perp (X_0, X_I, X_c)$

or

(a-2) $(U_1, U_I) \perp\!\!\!\perp (X_1, X_I, X_c)$;

(b-1) (M-1) and (M-2)

or

(b-2) (M-3)

(c) $Support(U_I \times U_0) = R_1 \times R_1$
$Support(U_I \times U_1) = R_1 \times R_1$; and

(d) $g_I(X_I, X_c) : R_{k_I} \rightarrow R_1$ for all X_c
$g_0(X_0, X_c) : R_{k_0} \rightarrow R_1$ for all X_c
$g_1(X_1, X_c) : R_{k_1} \rightarrow R_1$ for all X_c
$Support(g_0(X_0, X_c)\,|X_c) = R_1$ for all X_c
$Support(g_1(X_1, X_c)\,|X_c) = R_1$ for all X_c
$Support(X_0\,|X_c) = Support(X_0)$
$Support(X_1\,|X_c) = Support(X_1)$.

Then:

(I) Under (a-1) or (a-2), (b-1) or (b-2), (c) and (d), F_I and g_I are identified.If (M-1) and (M-2) are used, g_I is understood to be the least-concave version of the original g_I.

(II) Under (a-1), (b-1) or (b-2), (c) and (d), $g_0(X_0, X_c)$ and $F(u_0, u_I)$ are identified over the supports of (X_0, X_c) and (U_0, U_I) respectively.

(III) Under (a-2), (b-1) or (b-2) and (c), $g_1(X_1, X_c)$ and $F(u_1, u_I)$ are identified over the supports of (X_1, X_c) and (U_1, U_I) respectively. ■

Proof: Claim (I) is established in the theorems by Matzkin previously summarized.[4] If we establish either the second or third claim, it is clear that the other claim can be proved by a similar argument. We consider only claim (II). Fix X_c. Observe that for X_0, X_I, X_c

[4]All that is required to prove claim (I) is $U_I \perp\!\!\!\perp (X_I, X_c)$ and (b-1) or (b-2) from Matzkin.

in the support of (X_0, X_I, X_c)

$$F\left(y^0|D=0, X_0, X_I, X_c\right) = \frac{\int_{-\infty}^{y_0-g_0(X_0,X_c)} \int_{-\infty}^{-[g_I(X_I,X_c)]} f\left(u_0,u_I\right) du_I du_0}{\int_{-\infty}^{-[g_I(X_I,X_c)]} f\left(u_I\right) du_I},$$

and further observe that we know the denominator on the right-hand side. Thus we know the left hand side of

$$F\left(y^0|D=0, X_0, X_I, X_c\right) \Pr\left(D=0|X_I, X_c\right) = \int_{-\infty}^{y^0-g_0(X_0,X_c)} \int_{-\infty}^{-[g_I(X_I,X_c)]} f\left(u_0,u_I\right) du_I du_0.$$

Under condition (d) for each X_c, we can vary X_0 freely and trace out $(y_0, g_0+\ell_0)$ for each $p = \bar{p} = \Pr(D=0\,|X_I\,, X_c)$. That is, we can vary X_0 as required to fix $F(y^0\,|D=0,\ X_0, X_I, X_c)$ at a given value when y^0 is varied and in this way we determine g_0 up to scale ℓ_0. Tracing out g_0 for all values of y_0 and X_0 for each value of p identifies $F\left(u_0|D=0, X_I, X_c\right)$ up to scale for u_0. Setting $\bar{p}=1$, (*i.e.*letting $g_I \to -\infty$) we obtain $F_0(u_0)$ by virtue of assumption (c). The location of U_0 is obtained from the assumption of a zero median. This pins down the constant ℓ_0 and g_0. With g_0 in hand, we can recover $F\left(u_0, -g_I\left(X_I, X_c\right)\right)$. Varying x_I over its full support, we can identify $F\left(u_0, u_I\right)$. As this is true for each value of X_c, we have established (II) and by similar reasoning we can establish (III). Thus we establish the theorem. ■

Observe that the theorem can be modified so that the variables in common between g_1 and g_I are different from the variables in common between g_0 and g_I. Furthermore, the supports of U_0, U_1 and the conditional supports of (X_0) and (X_1) do not have to be R_1. It is enough to have $Support(U_0) = Support(g_0(X_0, X_c)\,|X_c)$ and $Support(U_1) = Support(g_1(X_1, X_c)\,|X_c)$. Theorem A-2 has a simpler structure than Theorem A-1. A discussion similar to that conducted after Theorem A-1 regarding the support of the X applies.

First, if $0 < Support(p) < 1$, then it is not possible to trace out the full distribution of $F_I(u_I)$, nor is it possible to identify $F_0(u_0)$ using the limit $\bar{p}=1$. This could happen, for example, if the support of X_I is restricted for all X_c such that

(A-5) $$Support(g_I(X_I, X_c)) \subset Support(-U_I).$$

Under this restriction, it is not possible to trace out the full distribution of U_I. Alternatively, even if the support of X_I is R_{k_I} for all X_c, it is possible that (A-5) is satisfied by virtue of restrictions on the function g_I. Notice also that (A-5) might satisfied for some values of X_c but not for others. Similar remarks apply to g_0. Again, restrictions on the range of X_0 may prohibit recovery of $F(u_0|D=0, g(X_I, X))$ even if (A-5) does not hold. Thus it may happen that

(A-6) $$Support(g_0(X_0, X_c)) \subset Support(U_0 | g_I + U_I < 0),$$
which might arise because of restrictions on the support of X_0, or because of restrictions on g_0. Even if (A-6) does not hold for some X_c, it may hold for others.

Theorem A-2 is weaker than Theorem A-1. It implies that we can recover $F(u_0|D = 1, X_I, X_c)$, from the available cross section data provided its conditions are satisfied. To see this, recall that we can obtain $F(y^0|X)$ by letting $p \to 1$, which we are free to do since g_I can be varied independently of g_0 and the support of g_I is the whole real line. Since by hypothesis we know $F(y^0|D = 0, X)$, we can apply the identity

$$F\left(y^0|X\right) = F\left(y^0|D = 0, X\right) \Pr(D = 0|X) + F\left(y^0 \,|D = 1, X\right) \Pr(D = 1|X)$$

to solve for $F(y^0|D = 1, X)$ provided that $\Pr(D = 1|X) \neq 0$.

Using non-experimental data, we are in the same position as we would be in if we ran an experiment that satisfied assumption (I-10) in the text. In particular we can identify the mean impact of treatment on the treated, $E(Y^1 - Y^0|D = 1, X)$. In the general case covered by Theorem (A-2), social experiments do not solve the fundamental evaluation problem that we cannot observe the same person in both states simultaneously, and so cannot observe both components of (Y^0, Y^1).

Collecting all of the subscripted variables into a common vector X, under the conditions of Theorem (A-1) it is possible to generalize from the data recovered from a social experiment, $F(y^1|D = 1, X)$ and $F(y^0|D = 1, X)$, combined with data on non-participants, $F(y^0|D = 0, X)$, to recover the entire distribution $F(y^0, y^1|X)$ provided that assumption (I-10) is satisfied, and provided that there are no general equilibrium effects. Thus it is possible to answer all of the questions posed in Section 1 of the text if agents are income maximizers. In the case of the Roy model described by Theorem A-1, social experiments are not required to answer these questions because $F(y^0|D = 1, X)$ is redundant information.

An extension of Theorem A-2 is useful in identifying the full distribution of outcomes conditional on X in the panel data case when treatments are irreversible.

Theorem A-3: Let (U_0, U_1, U_I) be median zero, independently and identically distributed random variables with distribution $F(u_0, u_1, u_I)$. Assume outcome and decision structure (A-1a) - (A-1c) and knowledge of $F(y^0, y^1\,|D = 1, X_0, X_I, X_c)$ and $\Pr(D = 1\,|X_I\,, X_c)$. Assume

(a) $(U_0, U_1, U_I) \perp\!\!\!\perp (X_0, X_1, X_I, X_c)$;

(b-1) (M-1) and (M-2)

or

(b-2) (M-3);

(c) $Support(U_0, U_1, U_I) = R_1 \times R_1 \times R_1$;

(d) Partition X_I into (X_{II}, X_{I0}, X_{I1}), where X_{II} is a subset of X_I (with k_{II} variables) not in X_0, X_1 or X_c, X_{I0} are variables X_I in common with X_0 (k_{I0} in number) and X_{I1} are variables (k_{I1} in number) in common in X_I and X_1. Partition X_0 into (X_{00}, X_{01}, X_{I0}) where X_{00} is a subset of X_0 (with k_{00} variables) not in X_1 or X_I and X_{01} is the subset of variables (k_{01} in number) in common with X_0 and X_1. Partition X_1 into (X_{11}, X_{01}, X_{I1}) where X_{11} is the subset of k_{11} variables not in common with X_1 and X_I. In this notation, we assume

$$g_I(X_{II}, X_{I0}, X_{I1}, X_c) : R_{k_{II}} \to R_1 \text{ for all } X_{I0}, X_{I1}, X_c$$
$$g_0(X_{00}, X_{01}, X_{I0}, X_c) : R_{k_{00}} \to R_1 \text{ for all } X_{01}, X_{I0}, X_c$$
$$g_1(X_{11}, X_{01}, X_{I1}, X_c) : R_{k_{11}} \to R_1 \text{ for all } X_{01}, X_{I1}, X_c$$
$$Support(g_I(X_{II}, X_{I0}, X_{I1}, X_c) \,|X_{I0}, X_{I1}, X_c) = R_1 \text{ for all } X_{I0}, X_{I1}, X_c$$
$$Support(g_0(X_{00}, X_{01}, X_{I0}, X_c) \,|X_{01}, X_{I0}, X_c) = R_1 \text{ for all } X_{01}, X_{I0}, X_c$$
$$Support(g_1(X_{11}, X_{01}, X_{I1}, X_c) \,|X_{01}, X_{I1}, X_c) = R_1 \text{ for all } X_{01}, X_{I1}, X_c$$

and

$$Support(X_{II} \,|X_{I0}, X_{I1}, X_c) = Support(X_{II})$$
$$Support(X_{00} \,|X_{01}, X_{I0}, X_c) = Support(X_{00})$$
$$Support(X_{11} \,|X_{01}, X_{I1}, X_c) = Support(X_{11}).$$

(I). Under (a) and (b-1) or (b-2), (c) and (d), F_I is identified and g_I is identified where if (M-1) and (M-2) are used, g_I is understand to be the least-concave version of the original g_I. (This follows from Theorem M-1).

(II). Under (a), (b-1) or (b-2), (c) and (d) $g_1(X_1, X_c), g_0(X_0, X_c)$ and $F(u_0, u_1, u_I)$ are identified. ■

Proof: Claim (I) is established in the same way that claim (II) of Theorem (A-2) is established. All that is needed to establish this claim is that $U_I \perp\!\!\!\perp (X_I, X_c)$, (b-1) or (b-2) from Matzkin and $g_I(X_I, X_c) : R_{k_I} \to R_1$ for all X_c.

Claim (II) is established in essentially the same way that claim (II) of Theorem (A-2) is established. From assumption (c),

(A-7) $$F(y^0, y^1 \,|D = 1, X_0, X_I, X_c) \Pr(D = 1 \,|X_I, X_c) = \underset{u_0,u_1,u_I}{F} (y^0 - g_0(X_0, X_c), y^1 - g_1(X_1, X_c), -g_I(X_I, X_c)).$$

From condition (d), we can vary the components X_{00}, X_{11}, X_{II} freely. If we set X_{11} to the value such that $g_1 \to -\infty$ and if we set X_{II} to the value such that $g_I \to -\infty$, (so $\Pr(D = 1 \,|X) = 1$) we may trace out $g_0(X_{00}, ...)$ given the remaining arguments for each value of the conditioning arguments. (This limit operation "zeros out" the last two arguments of (A-7)). We may repeat this argument for all conditioning subsets and recover $g_0(X_{00}, X_{01}, X_{0I}, X_c)$ up to a constant. Tracing out g_0 for all values of y_0 identifies $F(u_0)$ up to scale. The scale is identified by the median zero assumption. By a parallel argument, but reversing the roles of X_{11} and X_{00}, and g_1 and g_0, we obtain g_1 and $F(u_1)$. Staying in the set where $\Pr(D = 1 \,|X_I,\ X_c) = 1$, we may construct $F(u_0, u_1) = F(y_0 - g_0, y_1 - g_1)$ by independently varying g_0 and g_1, which we are free to do as a consequence of assumption (d). More generally, we may repeat this argument for all values of $p = \Pr(U_I > -g_I)$. We may vary x_{II} to offset any changes induced in x_{00} and x_{11}. Thus we can identify

$$F(u_0, u_1, -g_I(X_{II}, X_{I0}, X_{I1}, X_c)) = F(y_0 - g_0, y_1 - g_1 \,|D = 1\ , X_I, X_c),$$

and hence we can identify the full joint distribution $F(u_0, u_1, u_I)$ by tracing out X_{00}, X_{11} and X_{II}. We can do this for all $X_{I0}, X_{I1}, X_{01}, X_c$ and hence the theorem is proved. ■

Observe that theorem does not require infinite supports. Thus it is enough to have

$$Support(U_I) = Support(g_I(X_{II}, X_{I0}, X_{I1}, X_c \,|X_{I0}, X_{I1}, X_c))$$
$$Support(U_0) = Support(g_0(X_{00}, X_{01}, X_{I0}, X_c \,|X_{01}\ , X_{I0}, X_c))$$
$$Support(U_1) = Support(g_1(X_{11}, X_{01}, X_{I1}, X_c \,|X_{01}, X_{I1}, X_c))\ .$$

Appendix B

The data analyzed in this paper were gathered as part of an experimental evaluation of the training programs financed under Title II-A of the Job Training Partnership Act (JTPA). The experiment was conducted at a sample of sixteen JTPA training centers around the country. Data were gathered on JTPA applicants randomly assigned to either a treatment group allowed access to JTPA training services or to a control group denied access to JTPA services for 18 months. Random assignment covered some or all of the period from November 1987 to September 1989 at each center. A total of 20,601 persons were randomly assigned.

Follow-up interviews were conducted with each person in the experimental sample during the period from 12-24 months after random assignment. This interview gathered information on employment, earnings, participation in government transfer programs, schooling, and training during the period after random assignment. The response rate for this survey was around 84 percent. The sample used here includes only those adult women who (1) had a follow-up interview scheduled at least 18 months after random assignment, (2) responded to the survey, and (3) had useable earnings information for the 18 months after random assignment.

The sample was chosen to match that used in the 18-month experimental impact study by Bloom, et al. (1993). As in that report, the earnings measure is the sum of self-reported earnings during the 18-months after random assignment. This earnings sum is constructed from survey questions about the length, hours per week, and rate of pay on each job held during this period. Outlying values for the earnings sum are replaced by imputed values as in the impact report. However, imputed earnings values used in the report for adult female non-respondents are not used.