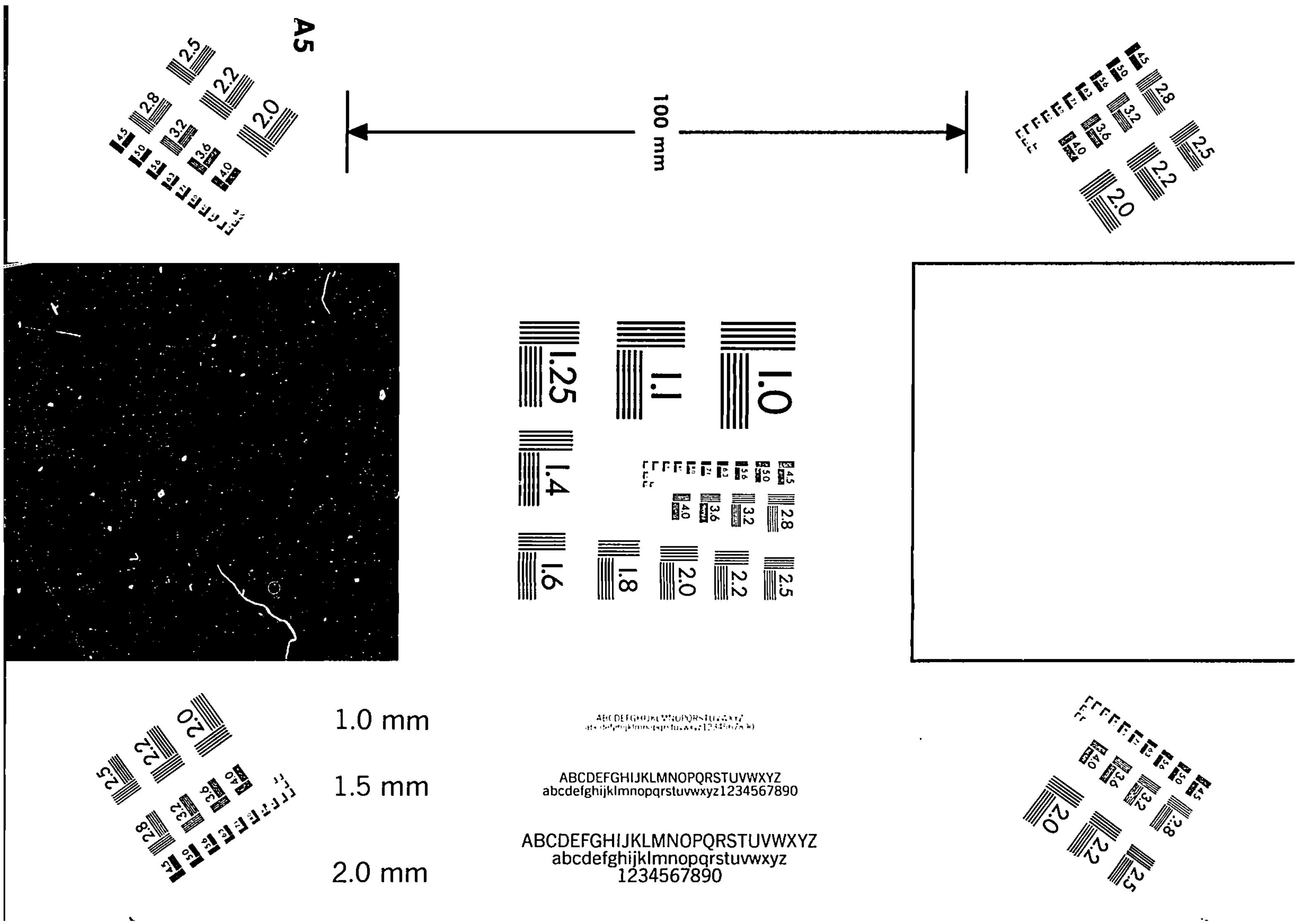

A5
100 mm
1.0 mm
1.5 mm
2.0 mm
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz1234567890
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890

DOCUMENT RESUME

ED 278 842 CE 046 390

AUTHOR	Johnson, Terry
TITLE	JTPA Evaluation at the State and Local Level. Volume V: A Guide for Net Impact Evaluations.
INSTITUTION	Washington State Dept. of Employment Security, Olympia.
SPONS AGENCY	International Business Machines Corp., Armonk, N.Y.; National Commission for Employment Policy (DOL), Washington, D.C.
PUB DATE	Mar 86
NOTE	134p.; A product of the JTPA Evaluation Design Project. Part of the appendix contains small print. For related evaluation materials, see CE 046 385-393.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Data Analysis; Data Collection; Educational Legislation; *Employment Programs; Evaluation Criteria; *Evaluation Methods; Federal Legislation; Literature Reviews; Local Issues; Outcomes of Education; *Program Evaluation; *Research Design; Research Methodology; *Statewide Planning
IDENTIFIERS	*Impact Studies; *Job Training Partnership Act 1982; Service Delivery Areas

ABSTRACT

This guide is intended to assist states and service delivery areas (SDAs) in addressing the new oversight responsibilities and opportunities stipulated by the Job Training Partnership Act (JTPA) with respect to net impact evaluations. The first chapter contrasts the objectives and provisions of JTPA with those of the Comprehensive Employment and Training Act (CETA). The literature review provided in the second chapter discusses previous studies that have analyzed the impacts of various employment and training programs and examines the implications of previous research for developing a state-level net impact model. The third chapter presents a conceptual framework for developing a state-level JTPA net impact mode. The conceptual framework, which builds on the results of the literature review, indicates the key outcome measures that should be examined, the subgroups of trainees for which impacts should be measured separately, the program activities (treatments) that should be examined, the types of economic and demographic characteristics that may affect outcomes, and the data sources that will be used to measure the aforementioned elements. A chapter on research design discusses various issues relating to the participant sample, use of a comparison group strategy, and sample sizes for the participant and comparison samples. The final chapter outlines a data analysis plan that addresses the issues of examining the adequacy of the comparision groups, estimating the net impacts of JTPA programs, obtaining net impact estimates for various subgroups, dealing with measurement error and other statistical issues, and making adjustments for potential data and design deficiencies. A sample JTPA application form is appended. (MN)

JTPA

EVALUATION
DESIGN PROJECT

JTPA Evaluation at the State and Local Level

Volume V: A Guide for Net Impact Evaluations

By Terry Johnson

March 1986

CONTEXT OF THIS VOLUME

This is one in a series of volumes produced by the JTPA EVALUATION DESIGN PROJECT.

PURPOSE AND PHILOSOPHY

The purpose of this project has been to develop a set of evaluation tools that are useful to states and local service delivery areas (SDAs) in judging the way their JTPA programs are being managed and the impact they are having. The intention has been to base these analytic and managerial tools on sound program concepts and research methods, and to design them such that the information obtained is of practical and direct use in improving JTPA policies and programs at the state and local level. This kind of information is also expected to make a unique contribution to national training policy and Federal oversight of JTPA.

It is hoped that these volumes will stimulate and support state and local evaluation efforts in JTPA, and promote more consistency than in previous programs with respect to the issues studied and the methods used to investigate them. An important goal is to encourage the generation of complementary information on program implementation and impact that is comparable across states and SDAs. Comprehensive, comparable information is essential to the development of a valid and reliable knowledge base for resolving problems and improving programs. It is also required for adjusting national training strategies to changing needs and priorities at the state and local level.

PRODUCTS

Consistent with this purpose and philosophy, the project has produced a set of materials to assist states and SDAs in evaluating their programs. These are to be useful in planning, designing and implementing evaluation activities. As an integrated collection, each set is developed to support comprehensive evaluations over the JTPA planning cycle.

The careful tailoring of these materials to state and local users is appropriate. JTPA represents a new employment and training policy shaped not only by the experience of managers and the perspectives of employers, but by scientific assessments of previous approaches for addressing unemployment, poverty and other barriers to economic security. In this context, the value of JTPA programs is also expected to be judged. In fact, the Act's assessment requirements are more explicit and sophisticated than those of any employment and training legislation to date. It clearly distinguishes between *monitoring* activities, whose purpose is to determine compliance (such as with performance standards) and *evaluation* activities, whose purpose is to determine how a program is being managed and implemented, and the kinds of effects it is having on recipients and relevant others. Equally significant, new constitutencies are expected to make these more rigorous assessments. States and SDAs now have this important responsibility. It is the first time in the history of employment and training programs that the Federal government's evaluation role has been significantly reduced.

This change affords states and local areas opportunities to influence public policy. It also requires them to assume new oversight responsibilities. Program evaluation is expected to become an integral part of the management of organizations administering, planning and delivering public training services. This is as it should be. The more information available at these levels, where changes in organizations can most readily be made, the more effective the management of JTPA programs. This project was undertaken in that context.

The evaluation tools produced by the project have been developed with a sensitivity to the differing needs, interests and resources of state and local users. They have been packaged into a single comprehensive and integrated set of volumes called *JTPA Evaluation at the State and Local Level*. The set contains planning and evaluation *guides* and *issue papers*. The following volumes are available in the set:

Volume	Author
I: Overview	Project Team
II: A General Planning Guide	Deborah Feldman
III: A Guide for Process Evaluations	David Grembowski
III Supplement: Some Process Issues at the State Level	David Grembowski
IV: A Guide for Gross Impact Evaluations	Carl Simpson
V: A Guide for Net Impact Evaluations	Terry Johnson
VI: An Implementation Manual for Net Impact Evaluations	Terry Johnson
VII: Issues Related to Net Impact Evaluations	
A. Issues in Evaluating Costs and Benefits	Ernst Stromsdorfer
B. The Debate Over Experimental vs. Quasi-Experimental Approaches	Ann Blalock
VIII: MIS Issues in Evaluating JTPA	David Grembowski

NOTE: Although each of the discrete products listed above is the responsibility of a single author, each seeks to incorporate the results of professional peer review, the many excellent recommendations of the advisory group, and the ideas and suggestions of the numerous practitioners interviewed in the process of developing these materials.

To further qualify these volumes, Volume III is accompanied by a supplement for state users. This is consistent with the significant differences between states and SDAs in the kinds of process issues that are most essential to study. The volume on net impact evaluations is sufficiently technical, because of the statistical methods involved, that a practical manual has been written to accompany it. This guide and manual tend to be more appropriate for states, since relatively large sample sizes are required for analysis. However, they are equally useful to larger SDAs and consortia of smaller SDAs which may want to jointly study the net impact of their programs. Regional evaluations, for example, can be very productive in providing management information relevant to regional labor markets. Although there is a separate issue paper on evaluating costs and benefits, this issue is also covered in the gross impact and net impact guides. In this respect, the user benefits from three related but different approaches to this important element of program evaluations. Also, the user should be aware that the Appendix of Volume II includes *A Report on a National/State Survey of Local JTPA Constituencies*. This survey was carried out by Bonnie Snedeker, with the assistance of Brian O'Sullivan, to provide additional input from practitioners to the development of the planning and process evaluation guides.

In conclusion, several expectations have directed the development of these volumes:

THE GUIDES

The General Planning Guide

This guide is to assist users in *planning, funding and developing an organizational capacity* to carry out process, gross outcome, and net impact evaluations and to *utilize their results*. Separate state and local versions are available.

The Evaluation Guides

These volumes are to have the following characteristics:

□ The guides are to *complement* one another.

- They are to provide information on program management and other characteristics of program implementation, which can:
  - —Describe the way in which administrative, managerial and service delivery policies and practices operate to affect outcomes, as a set of interventions separate from the program's services.
  - —Pinpoint the source, nature and extent of errors and biases for which adjustments must be made in gross and net impact evaluations.
  - —Help explain the results of gross and net impact evaluations.

- They are to provide information on aggregate gross outcomes, and outcomes differentiated by type of service and type of recipient, which can:
  - —Describe relationships between certain implementation modes and service strategies, and a broad array of client and employer outcomes.
  - —Help explain the results of net impact evaluations.
  - —Suggest the more important outcomes that should be studied in net impact evaluations.
  - —Help sort out those aspects of implementation that may be most critical to study in process evaluations.

- They are to provide information on net impact (the program's return on investment), which can:
  - —Closely estimate the effect of the program's services on clients.
  - —Suggest which services and client groups are most important to study in broader but less rigorous gross impact studies.
  - —Help identify the decision points in program implementation (particularly service delivery) which may be most important to study in process evaluations.

□ The guides are to enable the user to carry out *comprehensive assessments* of JTPA programs.

- They are to allow the user to acquire several different perspectives on the same program within a particular time period: on program implementation, on outcomes for clients and employers and on net impact.
- They are to permit the user to interrelate these different kinds of information to gain a wider understanding of what is happening in a program and why.

□ The guides are to describe approaches and methodologies as consistently as possible, to achieve *comparability*.

- They are to define variables and relationships as similarly as possible.
- They are to define research designs, and methods of data collection and analysis using as similar concepts as possible.

□ The guides are to *draw from past research* on employment and training programs, as well as seek *new* approaches and methods of specific value in evaluating JTPA at the state and local level.

- They are to replicate, to the extent possible and feasible, the issues and measures reflected in Federal monitoring and evaluation decisions.
- They are to make selective use of the results of relevant CETA studies, national studies of JTPA, and issue papers on JTPA evaluation by national public interest organizations in the employment and training area.
- They are to rely on the professional literature in applied social research.

THE ISSUE PAPERS

Volume VII contains two issue papers which serve as companion pieces to the preceding volumes on net impact evaluation. The first paper on cost-benefit issues is designed to help users identify, measure and analyze relationships between monetary and nonmonetary costs and benefits in determining the program's return on investment. The second paper examines the pros and cons of different research strategies associated with the net impact approach. The final volume on MIS issues is to assist users in better understanding how JTPA and other employment and training *management information systems* can efficiently support the evaluation of program implementation and impact.

THE SET OF VOLUMES

The set is *integrated*, but affords *flexible use*. The user can utilize the entire set for comprehensive evaluations over a two-year planning cycle or longer planning period, or the user can apply the information in each volume independently, based on the most pressing evaluation priorities and timeframes and given the extent of resources, during a particular fiscal year or biennium.

It should be understood that although evaluation products have been developed for JTPA, their basic principles and methods can be applied more broadly by states and local areas to evaluate other employment and training programs and other social programs.

GENERAL ACKNOWLEDGMENTS

The JTPA EVALUATION DESIGN PROJECT was developed and carried out based on the partnership philosophy that underlies the JTPA legislation. Several partnerships should be recognized for their substantial contributions to the products previewed here: the project development and coordination partnership; the public-private funding partnership; the interdisciplinary design partnership; and the advisory partnership.

The Development and Coordination Partnership: Washington Employment Security Department

Project Coordinators: Gary Bodeutsch, Ann Blalock

Assistant Coordinators: Deborah Feldman, Chris Webster

Interdivisional Consultants:

Training Division: Martin McCallum, Brian O'Neall, Ross Wiggins

Research and Information Systems: Jeff Jaksich, Mike Gioimo, Irv Lefberg

The Public/Private Funding Partnership

Funders:

IBM Corporation: Corporate Support Program, Armonk, NY

National Commission for Employment Policy, Washington, D.C.

Washington State Employment Security Department, Olympia, WA

Contributors:

Safeco Insurance Company of America, Seattle, WA

Seattle/King County Private Industry Council, Seattle, WA

SPSS, INC., Chicago, IL

The Interdisciplinary Design Partnership

Consultant:	*Area of Expertise:*	*Affiliation:*
Burt Barnow	Labor Economics	ICF, INC., Washington, D.C.
David Grembowski	Planning	University of Washington, Seattle, WA
Terry Johnson	Labor Economics	Battelle Memorial Institute, Seattle, WA
Brian O'Sullivan	Planning	Seattle/King County Private Industry Council, Seattle, WA
Carl Simpson	Sociology	Western Washington University, Bellingham, WA
Bonnie and David Snedeker	Management	Snedeker Scientific, Inc., Seattle, WA
Ernst Stromsdorfer	Labor Economics	Washington State University, Pullman, WA

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1
INTRODUCTION

CHAPTER 1. INTRODUCTION

The Job Training Partnership Act (JTPA) of 1982, which officially replaced the Comprehensive Employment and Training Act (CETA) on October 1, 1983, continues CETA's stated objective of providing job training to improve the economic well-being of trainees. In particular, Section 106 of the Act explicitly states that "job training is an investment in human capital," and that the "return on the investment is to be measured by the increased employment and earnings of participants and the reduction in welfare dependency." Although the major objectives of the two programs are similar, the passage of JTPA significantly changed the administration and operation of employment and training programs, as well as the types of services provided.

The major characteristics of JTPA that distinguish it from CETA and earlier employment and training systems include: (1) the expanded role of the private sector as a partner in the planning and overseeing of job training services; (2) the transfer of responsibility for program administration, operation, and accountability from the federal government to the state and local level; (3) requirements on how funds are expended that result in programs that emphasize skills training rather than work experience or public service employment; and (4) a central role for performance standards to ensure that the employment and training system meets its objectives. As part of the performance standards application process, the authorizing legislation requires governors to use six percent of the funds allocated to each state to either provide incentive grants for service delivery areas (SDAs) that exceed their performance standards or technical assistance to SDAs that do not meet their standards.

As a result of these numerous changes, states and local areas have been given much greater responsibility for program accountability under JTPA than in the past. This has, in turn, increased the need for reliable analysis of program effects and outcomes at the state and local level. For the most part, the only reliable analyses of employment and training programs have been conducted at the national level. Although these national studies have provided much useful information on the effectiveness of employment and training programs overall, they are of limited use to state administrators and local program operators for several reasons. First, the results of these studies have not been available on a timely basis. For example, several recently completed studies of the impact of CETA programs examined the early post-program experiences of 1976 and 1977 enrollees. Because of the many changes in CETA since that time, and the major differences between CETA and JTPA, the usefulness of these results for guiding managerial decisions is limited.

Second, because the studies have been conducted using nationally representative samples of program participants, the sample sizes

available at the state and local level are too small to support reliable program analysis at a disaggregated level. Moreover, because of confidentiality concerns, local identifiers have not been available. This data constraint has prevented analysts from both controlling for differences in program environments as well as from examining the extent to which program impacts vary by labor market or other characteristics of the environments in which the programs operate.

Finally, most of these studies have only examined earnings gains, whereas the authorizing legislation also requires an examination of employment gains and reductions in welfare dependency. As a result of these limitations, previous national studies are of little use in assisting states and local areas to make informed and objective judgments of the effectiveness of JTPA programs to meet their increased oversight responsibilities under the Act. Moreover, while the need for reliable program analysis at the state and local level is increasing, the federal government is greatly reducing its role in providing evaluation research information.

In response to the lack of useful program evaluation information at the state and local level, the National Commission for Employment Policy awarded a grant to the Washington State Employment Security Department to develop a set of integrated models that could be used by states and SDAs in evaluating their JTPA programs. These models are presented in a set of complementary evaluation guides which are outlined in the introductory context section of this volume. Included in this set are: (1) this volume on state-level net impact evaluation, (2) guides for state and local process evaluations, and (3) a guide for state and local gross impact evaluations. Since gross impact and process evaluations can supplement and inform the net impact analysis in important ways, this guide will periodically refer to these additional evaluation approaches.

In this guide we outline the issues involved in designing a reliable net impact evaluation of JTPA and present a specific evaluation approach or "model" for carrying out such an evaluation. More specific details on net impact data collection and data analysis are contained in a companion volume, An Implementation Manual for Net Impact Evaluations (Volume 6). While this guide and the manual accompanying it assume a state-level analysis, larger SDAs or a consortium of SDAs may also readily adapt this net impact design to meet more local-level analysis needs.

Before describing the net impact evaluation model, it is important to understand the significance of the term "net," and how a "net" impact model differs from a "gross" impact model. A gross impact analysis essentially compares the post-program labor market experiences of participants with their pre-program experiences and attributes all gains to the program. Although a carefully designed gross impact analysis can provide useful information on the relative effectiveness of different JTPA program activities, because numerous other factors may have changed from the pre-program to the post-program period that affect participants' labor market experiences (e.g., improvement in labor market conditions), such an analysis is not likely to isolate the true benefits of program participation per se and should not be used to

measure the return on the job training investment. On the other hand, a net impact analysis compares the labor market experiences of participants with the experiences of a comparison group of otherwise similar nonparticipants. The comparison group is used to approximate what the labor market experiences of participants would have been in the post-program period had they not participated in the program. As such, a net impact analysis only attributes to program participation the incremental gain in labor market experiences that occurs over and above what would have happened had these individuals not participated in the program. This is the appropriate concept for providing information on the return on the investment of job training programs funded under JTPA.

In developing a state-level net impact model of JTPA, we were guided by several considerations. First, in order for the model to assist states in meeting their new accountability responsibilities, the model must be usable. In particular, it must provide meaningful information that can be understood by a nontechnical audience in a cost-effective and timely manner. Second, to maximize usability, the model must recognize the severe resource constraints that states and SDAs face, as well as certain other practical considerations. The two most important practical considerations that affected the recommended approach are (1) states and local SDAs will not generally be willing to implement an experimental design in which eligible applicants are randomly assigned to treatment-control status, and (2) states and local SDAs will not generally be willing to conduct follow-up interviews with a large sample of participants and comparison group members. Finally, the model must be designed to produce valid estimates of the net impacts of JTPA programs on the post-program outcomes of participants.[1]

In developing a state-level net impact model we have attempted to meet these objectives while recognizing the inherent tension between them. To the extent that we are successful in designing a model that is usable and provides valid results, then an important by-product, namely consistency in application across states, will occur. This will maximize the information obtained from the evaluations and make an important contribution to what is known about the effectiveness of employment and training programs in states that face different environments.

The remainder of this report is organized as follows. In Chapter 2 we present a review of the previous studies that have analyzed the impacts of various employment and training programs. This review documents what is known about program impacts by indicating the outcomes examined, the comparison groups used, the different types of data

[1] Three types of validity can be distinguished: internal validity is achieved when the models' estimates are unbiased estimates of the net impacts of JTPA; external validity is achieved when the results can be generalized to the state as a whole; statistical validity is achieved when the power of the analysis is sufficient to produce statistically reliable estimates of the impact of JTPA. Our approach to designing a net impact model to achieve these three types of validity is described in Chapters 4 and 5.

sources used, the subgroups for which separate impacts were estimated, the statistical techniques used, and the net impact results obtained. The implications of previous research for developing a state-level net impact model are also discussed.

In Chapter 3 we present a conceptual framework for developing a state-level JTPA net impact model. The conceptual framework builds on the results of the literature review and indicates the key outcome measures that should be examined, the subgroups of trainees for which impacts should be separately measured, the program activities (treatments) that should be examined, the types of economic and demographic characteristics that may affect these outcomes, and the data sources that will be used to measure these elements.

In Chapter 4 we discuss several issues involved in developing a research design for analyzing the net impacts of JTPA and present our recommended approach. The most difficult research design task is the development of a reliable comparison group. The comparison group of nonparticipants must be similar to participants on characteristics that affect program outcomes, and comparable data must be available for both groups. In this chapter we discuss alternative comparison group strategies and present our recommended approach. We also describe a plan for selecting a representative sample of JTPA participants and sufficient numbers of participants and comparison group members to provide valid results.

In Chapter 5 we describe a general approach to analyzing the data to estimate the net impacts of JTPA programs on participants' post-program outcomes. This chapter describes methods for examining the adequacy of the comparison groups selected, as well as procedures for estimating the net impacts of JTPA on the key participant post-program outcomes indicated in the legislation. We also discuss potential threats to the validity of the analysis and techniques for adjusting for such problems.

CHAPTER 2
LITERATURE REVIEW

Evaluations of Pre-CETA and Other Employment and Training Programs

Evaluations of CETA Programs

Recent Methodological Developments

CHAPTER 2. LITERATURE REVIEW

The literature analyzing the impact of various employment and training programs has grown considerably over the past ten years. With the exception of several recent impact evaluations of CETA, most of these studies examined the impact of CETA's predecessors, for example, MDTA institutional and on-the-job training programs. Some of the evaluations dealt with CETA contemporaries, such as the WIN program, the Job Corps, or special employment and training demonstrations. Even though none of these studies dealt explicitly with JTPA programs, and many focused on pre-CETA programs, they are of interest, both because the programs that were examined have many characteristics in common with JTPA programs and because the evaluation issues addressed are in many cases identical to those that must be faced in developing a JTPA state-level net impact model. However, the differences between JTPA and CETA, and the even greater differences between JTPA and pre-CETA programs and other employment and training programs, must be kept in mind when determining the implications of these studies for analyzing JTPA.

In this chapter we review the most relevant previous studies of the impact of employment and training programs on participants' post-program outcomes. Our review summarizes what is know about program impacts by documenting the outcomes examined, the comparison groups used, the different types of data sources used, the subgroups for which separate impacts were estimated, the statistical techniques used, and the net impact results obtained. It should be noted that because virtually all of these studies relied on large-scale data bases on program participants that were created well before the analysis was undertaken, little (if any) information was generally available on the specific content of the services provided, whether the program services were provided as planned, and the extent to which the services provided varied across sites or over time. The analysis of such information (which is the subject of Volume III in this series, A Guide for Process Evaluations) would have been particularly useful in understanding the large variation in net impact estimates described in this chapter.

At the same time that the program evaluation literature has been growing, a complementary theoretical literature focusing on methodological techniques that can be used to adjust for various potential selection biases in measuring program impacts has developed.[2] In this chapter we also briefly summarize the implications of the most relevant methodological studies for conducting

---

2 The exact nature of selection bias, what it is, what its sources are, and what can be done to account for its presence, is discussed in Chapters 4 and 5.

a state-level net impact analysis of JTPA.

We begin with a discussion of evaluations of pre-CETA programs and of other non-CETA programs. We then review the recent studies of the net impact of CETA programs on participants' post-program labor market experiences and briefly summarize a few recent contributions to the methodological literature. We conclude this chapter with a discussion of the implications of these studies for conducting a net impact analysis of JTPA at the state level.

EVALUATIONS OF PRE-CETA AND OTHER EMPLOYMENT AND TRAINING PROGRAMS

Two important contributions to measuring the impact of training programs on participants' post-program earnings are those by Ashenfelter (1975, 1978). Ashenfelter's approach builds on the availability of longitudinal Social Security Administration (SSA) earnings records to measure the net effect of training programs on earnings. He uses data on MDTA classroom training participants who entered training in the first three months of 1964 and a sample of nonparticipants drawn from the Continuous Work History Sample (CWHS). Because the participant and comparison groups are drawn from very different populations, Ashenfelter attempts to control statistically for differences between the two groups by specifying an earnings function that would prevail for both groups in the absence of training. In his model, current earnings are assumed to be the sum of an autoregression in earnings (i.e., several years of pre-program earnings are included as control variables) plus a components-of-variance error term to control for personal and time effects.[3] Assuming that the earnings generating functions are of the same form for participants and nonparticipants, Ashenfelter presents numerous estimates of net program impacts.

Ashenfelter's model produces estimates of the net program impact in each of the five years following training taken in 1964, disaggregated by race and sex groups. His results indicate that training increases the earnings of all subgroups. For men, the effect is between $150 and $500 per year immediately following training, but declining to approximately half this figure five years later. For women, the effect ranges between $300 and $600 immediately following training and does not appear to decline in subsequent years.

Although Ashenfelter's empirical results are interesting, the more important contributions of his work are methodological. A basic point in Ashenfelter's study is that participation in the program is not random, and specifically that individuals whose earnings are unusually low just prior to participation in training are the ones most likely to

---

3 The error term includes a component that is specific to an individual and that does not vary over time (known as a "fixed-effect" and included to capture differences in unmeasured factors such as motivation and ability); a component that is specific to a time period but does not vary over individuals (to capture economy-wide changes in earnings); and a random component that varies over individuals and time.

enter training. This "dip" in earnings introduces statistical bias in any model that contains a fixed-effect in the error component in the estimation equation for current earnings. To deal with this problem, Ashenfelter capitalizes on the availability of detailed pre- and posttraining earnings records for participants and nonparticipants and (in some cases) estimates first-difference models to remove the fixed effect, which eliminates the bias under certain assumptions.[4] The results of the first-difference models generally yield somewhat larger net impact estimates than those obtained from the autoregressive earnings models described above and also indicate no decline in impact in subsequent years for either men or women. Ashenfelter also emphasizes the importance of identifying the structure of the selection process for participants. This observation is extremely important in evaluating program impacts, because program impact estimates may be very different and potentially misleading if the selection process changes from year to year and such changes are not accommodated in the evaluation methodology.

Cooley, McGuire, and Prescott (1979) provide an alternative methodology to analyze the MDTA data. A major focus of their study is whether the CWHS data utilized by Ashenfelter and others to draw a comparison group for MDTA trainees provide an adequate comparison group. Their results indicate that, for 1969-71 cohorts of institutional MDTA trainees, program "no-shows" (i.e., individuals who applied to the program but did not participate) are a more reliable comparison group than individuals drawn from the CWHS. The basis for their choice is the empirical observation that the autocorrelation coefficients of earnings are very similar for the no-show and participant groups in the pre-program years, whereas the coefficients differ considerably between participants and the CWHS comparison group.[5]

Cooley, McGuire, and Prescott utilize a simple analysis-of-variance model to estimate the net impact of MDTA classroom training and on-the-job training programs on participants' post-program SSA earnings. Unlike most other studies, the estimated net effects of training are very similar for men and women, and range between $200 and $500 annually, with some indication that training effects increase over time. The authors also find that the training programs increased earnings of men primarily by increasing their probability of employment, defined as having SSA earnings of at least $50. For women, however, the positive impact on the probability of employment accounts for only one-eighth of the total increase in earnings, suggesting that

---

4 First-difference models are regression equations in which all variables (dependent and independent) are expressed as changes between two time periods. This is often a convenient form for estimating program impacts because it enables one to remove the effects of variables that do not change over time.

5 The autocorrelation coefficients represent the correlation between earnings in a given year and earnings in each preceeding year. Such information is often useful in understanding how earnings change over time.

the primary impact of training for women is through higher wage rates or more hours worked per year.

Although the Cooley, McGuire, and Prescott study is an important contribution to the literature, some potential limitations of their analysis must be recognized. For example, although no-shows appear to be a reliable comparison group, at least to the extent of having similar earnings autocorrelation patterns, no-shows introduce another potential selection bias in that they may be systematically different from trainees in unobserved characteristics that may be responsible for the observed program impacts. It should also be noted that these authors' rejection of the CWHS data for drawing comparison groups for the 1969-71 cohort of trainees does not imply that Ashenfelter's use of these data to draw a comparison group for the 1964 cohort is inappropriate. That is, since the composition of trainees shifted drastically between 1964 and 1969 toward a younger, more poorly educated, and more disadvantaged group of individuals, the CWHS data could have generated a reliable comparison group for the earlier sample of trainees that had more job skills on average.

A paper by Kiefer (1979) analyzes the effect of training on the earnings of trainees in four programs: (1) MDTA Institutional Training Program, (2) Job Opportunities in the Business Sector (JOBS), (3) the Job Corps, and (4) the In-School and Summer Neighborhood Youth Corps (NYC). His analysis is based on a data set developed by the Office of Economic Opportunity (OEO) and the Department of Labor (DOL) that collected survey data on trainees and members of a comparison group from ten standard metropolitan statistical areas (SMSAs) judged to be representative of major SMSAs. The comparison group was developed from a household screening of a large sample of individuals in each SMSA who were eligible for the program but did not apply, and then matched to participants on the basis of age, race, and sex. Kiefer presents evidence indicating that the comparison group and the trainee group are different in important characteristics, which reduces the validity of the program impact findings. By controlling for earnings history prior to program participation, however, he is able to statistically reduce the potential differences between the two groups.

Kiefer's results indicate that the effects of training on earnings varied considerably by program type, sex, and racial group. For the adult training programs, MDTA and JOBS, he concludes that the increase in annual earnings for women is approximately $500 and is somewhat higher for nonblack women. The impact of training on the earnings of adult men is less clear. Kiefer finds that men trained through MDTA or JOBS generally did not regain the earnings position they held prior to training relative to the comparison group, although black men who participated in the JOBS program are estimated to have obtained modest earnings gains of approximately $100 to $150 per year.

In a companion paper that focuses exclusively on the male MDTA classroom training sample of the OEO/DOL data base, Kiefer (1978) estimates the effect of training separately on employment and weekly earnings. His model estimates the effect of training as a nonlinear function of the number of weeks an individual stays in the program (i.e., the number of weeks and the number of weeks squared are included

as independent variables in the regression equation). Kiefer's estimates indicate no effect of training on the probability of employment for either black or white men and negative impacts on weekly earnings for both groups, although the impact is only significant for blacks. Although these negative net impact findings may result partly from the comparison group problem indicated above, they are not inconsistent with results from recent evaluations of CETA described later in this chapter.

A recent paper by Gay and Borus (1980) provides an interesting reanalysis of data from the OEO/DOL sample. They estimate separate earnings equations for race/sex subgroups for each program and then attempt to identify short-term economic indicators that are capable of predicting participants' long-term post-program earnings gains--an important policy question. The long-term impact of program participation is measured by comparing the 1973 SSA earnings of the participant and comparison group members. The beginning of 1973 corresponds to approximately one-and-one-half to three years after participants left the program. The average earnings impacts reported by Gay and Borus for participants in the adult training programs (MDTA and JOBS) are comparable to those found by Kiefer and indicate that women (particularly nonblack women) benefited most from program participation, and that the impact of the youth programs (Job Corps and NYC/OS) on earnings are negative.

Gay and Borus' attempt to identify reliable short-term indicators of long-term program performance, however, was unsuccessful. Specifically, they conclude that short-term indicators such as "placement" are not reliable predictors of long-term program impacts. It should be noted, however, that their measure of various performance indicators depends only on the mix of clients served. Although the mix of clients served is likely to be an important determinant of variation in service-provider effectiveness, it is also important to adjust the performance indicator for differences in local labor market conditions. As such, their conclusion that placement is not a reliable predictor of long-term earnings impacts should be interpreted with caution.

Schiller (1978) reports a longitudinal analysis of the impact of WIN on participants' post-program earnings and welfare dependency. This study attempts to identify the specific dimensions of relative effectiveness by explicitly examining the interaction of program activity and economic conditions. The study compares the earnings and welfare dependency of WIN participants with that of a comparison group of nonactive registrants to evaluate the short- and long-term impacts of the program and to determine the cost-effectiveness of WIN. Because nonactive registrants are likely to differ systematically from WIN participants on characteristics that affect program outcomes, however, the study results should be interpreted with caution.

Schiller distinguishes five levels of treatment provided by WIN: (1) no service; (2) advice and assistance in job placement; (3) classroom education; (4) vocational training; and (5) assignment to on-the-job training or public service employment. He finds modest earnings gains and reductions in welfare dependency as a result of program

participation, although effects vary considerably with participant characteristics and the services provided. His results indicate that men benefited only from OJT or subsidized public employment, whereas women benefited from all of the services. In addition, his results indicate that individuals without recent employment experience generally showed the largest gains.

A major experimental evaluation of a job training program was conducted as part of the Supported Work Demonstration, a transitional work experience program that provided training in basic work skills. The demonstration offered supported work jobs to a random sample of eligible individuals in four major hard-to-employ groups: ex-offenders, ex-addicts, high school dropouts, and women who were long-term recipients of AFDC payments. The unique feature of the Supported Work concept was the provision of a subsidized but productive job in a setting featuring peer support, close supervision, and a gradual increase of work pressure and demands on the job over time. In addition, the experimental design used allows one to overcome the problem of selection bias and obtain valid estimates of the impacts of program treatments.

Evidence from the Supported Work Demonstration indicates that of the four target groups enrolled in the study, long-term recipients of AFDC benefited the most (Masters, 1981). In particular, participation in Supported Work for the AFDC group led to an increase in their employment rate, hours worked, and earnings, both while they were in the program and after they left it. Concurrent with these positive earnings effects was a net reduction in welfare dependency for those in the experimental AFDC group. There was also some indication that the program improved the quality of employment of the experimental population.[6] The program was not generally successful for disadvantaged youths or other predominantly male target groups, although there were positive earnings impacts for former drug users who had the least employment opportunities (Dickinson, 1981).

The literature evaluating the effects of employment and training programs for disadvantaged youths has also grown considerably in recent years. In addition to the Supported Work Demonstration, and the studies described above that evaluated the youth programs contained in the OEO/DOL sample, there have been major evaluations of the Job Corps and the Neighborhood Youth Corps (NYC), among others. Because these programs have several elements in common with JTPA youth programs, we briefly review what is known about them below.

The most in-depth evaluation of the impact of the Job Corps program was conducted by Mallar et al. (1982). The Job Corps provides economically

---

6 Although these results appear quite promising, their generalizability is likely to be limited. This is primarily because the sample of long-term AFDC recipients who volunteered for the demonstration and were subsequently assigned to experimental/control status is not representative of AFDC recipients overall. For example, Masters (1981) estimates that only 17 percent of all AFDC recipients met the Supported Work eligibility criteria.

disadvantaged youths between 16 and 21 years of age with basic education, vocational training, and support services within a residential setting in order to improve their earnings potential and help them become better citizens. As such, Job Corps programs have several elements in common with JTPA programs for disadvantaged youth, although corpsmembers are generally much more disadvantaged and the residential living feature is a major distinguishing characteristic.

The methodology used in the Mallar study involves comparing the post-program experiences of a sample of Job Corps participants with the experiences of a sample of disadvantaged youths selected from comparison sites. The comparison group of youths was developed from lists of school dropouts (70 percent) and from applicants to local Employment Service offices (30 percent). Based on detailed survey data collected over a four-year post-program period, the findings indicate that corpsmembers worked an average of three additional weeks per year and had higher earnings than nonparticipants by approximately $600 per year. The overall effects appear to persist throughout the four-year observation period. The Mallar study results also indicate that men and women without children consistently benefited more in terms of employment and earnings than did women with children. In fact, participating women with children were employed less in the post-program period than otherwise similar comparison group members. The Mallar study also finds that program completers consistently benefited more than noncompleters in terms of employment and earnings and that early dropouts from the Job Corps benefited little or not at all.

Although the general approach used in the Mallar study is quite reasonable, the empirical results should be interpreted with caution for several reasons. First, because ES applicants are likely to be more job ready on average than high school dropouts, such a mixed comparison group strategy is not likely to result in unbiased impact estimates. Second, the initial sample design called for a cross-section of participants at a point in time, which systematically excludes a large fraction of short-term participants. For example, 40 percent of all enrollees in FY 1977 (the year in which the Mallar sample was drawn) dropped out of Job Corps during the first 90 days, compared with nine percent of corpsmembers in the Mallar study. Both of these procedures create an analysis sample that is unrepresentative of the population of Job Corps enrollees, which necessarily limits the generalizability of the net impact estimates. Furthermore, the quality of the comparison group developed for young women was particularly deficient; females in the comparison group were much more likely to be married or to have children. Because of the well-documented evidence that marital status and presence of children affect the work effort of women, the Mallar study's inability to match well on marital status for women creates additional biases in assessing program impacts for this subgroup.

The Neighborhood Youth Corps Program (NYC) was created by the Economic Opportunity Act of 1964 to provide part-time work experience, remedial education, and job training assistance to disadvantaged youths who did not complete high school or were likely to drop out of high school. In addition to a summer employment program, NYC had both an in-school

program (providing part-time employment while attending school to encourage youths to remain in school and graduate) and an out-of-school program (providing work experience and some vocational training). The in-school component of NYC was examined by Somers and Stromsdorfer (1970), and the out-of-school component was examined by Borus, Brennan, and Rosen (1970).

The Somers and Stromsdorfer study is based on a nationwide sample of 576 individuals who enrolled in in-school and summer NYC programs in 60 areas between July 1965 and June 1967. For a comparison group, the authors used students who met the NYC eligibility requirements but did not participate in the program. After controlling for differences in measured characteristics through multiple regression techniques, they find that the NYC in-school program had a significant impact on the earnings of participants after high school. The program had no effect on the wage rates of participants; rather, the earnings gains were a result of increased employment of participants. Benefits to black women accounted for much of the overall gain for the total sample. These results should be interpreted with caution, however, because statistical tests provided by the authors indicate that there were systematic differences between the participant and comparison samples. Although the regression analysis method used by the authors in part corrects for this problem, at least for measured variables, biases may remain because of differences in unmeasured characteristics (e.g., ability, work attitude) that affect post-program outcomes.

Borus, Brennan, and Rosen (1970) evaluated the out-of-school NYC program using data on 604 participants in five urban areas in Indiana. Participants in this study are defined as individuals who had worked at least one day and had left the program by December 31, 1966. Post-program earnings of participants (obtained from UI earnings records) are compared with those of a comparison group of 166 individuals who applied for the program and were eligible but did not enroll because (1) they were placed on a waiting list and were never called, (2) they could not be reached for a job assignment, or (3) they did not report when assigned to a job.

In contrast to the Somers and Stromsdorfer study of the NYC in-school program, as well as to most other evaluations of employment and training programs, the authors find no earnings gains for female participants and relatively small earnings gains as a result of NYC participation for young men. These unexpected results could in part be due to lack of comparability between participants and comparison group members. There is some evidence that individuals with less prior schooling benefited more from the program and that length of participation is positively correlated with post-program earnings gains.

Another program that was designed to help disadvantaged youths was the Concentrated Employment Program (CEP). CEP provided youths with work experience and job search assistance. The only major attempt to evaluate CEP, by Kirschner Associates (1969), compared participants' behavior before and after program participation to infer program impacts. Although such a design can provide useful information on gross program impacts, the large post-program employment and wage increases reported should not be interpreted as measures of net program

impacts. The limitations of a before-after comparison design for measuring program impacts are particularly severe for youths, who have steep age-earnings profiles and would have had considerably higher earnings at later periods even without participating in the program.

EVALUATIONS OF CETA PROGRAMS

The recent availability of data from the Continuous Longitudinal Manpower Survey (CLMS) has made it possible to evaluate the net impact of CETA programs. The CLMS collected extensive information on the socioeconomic and demographic backgrounds, the in-program experiences, and the subsequent labor market experiences of a nationally representative sample of CETA participants beginning in FY 1975. To measure net program impacts, Social Security Administration (SSA) earnings records (both pre- and post-program) have been attached to the CLMS public use files and to various March Current Population Surveys (CPSs), which have been used for drawing comparison groups. Until recently, most net impact evaluations of CETA based on these data have been conducted by Westat, Inc. In this section we describe the Westat studies and other recent net impact evaluations of CETA using the CLMS data.

Through a series of studies, Westat has examined the post-program earnings impacts for FY 1975, FY 1976, and FY 1977 CETA enrollees who participated in a program activity for at least eight days. Because their basic approach and findings are quite similar across these reports, we discuss below their methodology and findings as described in their July 1982 report. In that report, Westat presents impact estimates for the first and second post-program years (1977 and 1978) for FY 1976 enrollees who terminated by December 31, 1976. Their analysis is based on a comparison of the Social Security Administration earnings records of CETA participants and those of a matched comparison group drawn from the March 1976 CPS.[7] Their net impact estimates are derived from weighted regression equations.[8]

Westat's results indicate that CETA terminees earned approximately $300 more in both 1977 and 1978 than individuals in the matched comparison group. There appear to be considerable differences in impact by program activity, with OJT trainees experiencing the largest earnings

---

7 The comparison group was developed using stratified matching techniques separately for low, medium, and high pre-program earners. The highest priority matching variables for low earners were demographic characteristics (e.g., sex, race/ethnicity, age, education), whereas, prior earnings was emphasized in the matches for both medium and high earners.

8 CPS cases in each cell were weighted so that the sum of the weights equalled the number of CLMS cases in that cell, and where cells were defined by combinations of characteristics on which the match was based. The weighting was done to create a comparison group that was distributionally more similar to CETA participants on background variables believed to be associated with earnings potential.

gains, followed by classroom training and PSE trainees. Individuals enrolled in work experience programs had lower post-program earnings than individuals in the comparison group, although the differences are not statistically significant. Westat also finds that the impacts were much larger for individuals who remained in the program longer. In terms of subgroup effects, Westat's results indicate that women experienced larger earnings gains than men, and that the effects for minority men were particularly small. They also observe a general pattern of increasing earnings gains with age, with generally negative estimates of program impacts for youths.

In interpreting the implications of Westat's findings, one should use caution for several reasons. In particular, the procedures used to align and match the CLMS and CPS samples and the specific sample frame and sample exclusion decisions followed have certain disadvantages. In addition, the specification of the weighted regression model to estimate net impact was inappropriate. Because many of the recent evaluations of the impact of CETA adopted Westat's methodological choices on these issues, below we indicate some of the limitations of their approach.

Westat drew comparison groups using a stratified matching technique. The stratified technique, which selects several variables and divides each sample into cells based on factorial combinations of those variables, has several potential disadvantages. First, higher-priority variables must be matched exactly, even if this results in a very large difference in a lower-priority variable. Second, only categorical variables can be used in stratified matching. Because the categorization of a continuous variable is necessarily arbitrary, individuals who are close on a continuous variable may be judged to be far apart when categorized (e.g., if an age cell included individuals 30 to 44 years of age, an individual aged 44 could be matched to another aged 30 but not to one aged 45). Third, very few variables can be used to match the samples. Fourth, a very large number of potential comparison group members is required to generate exact matches. Despite these potential deficiencies, evidence presented in Dickinson, Johnson, and West (1986) indicates that the net impact estimates obtained from stratified matched comparison groups are quite similar to the results obtained when more sophisticated matched comparison groups are used.

A second potential problem concerns the way in which the CLMS and CPS samples are aligned. Westat divides the CLMS sample into cohorts based on the fiscal year in which participants enrolled in CETA and matches these individuals with CPS sample members who were interviewed in March of that fiscal year. For example, the FY 1976 CLMS cohort (individuals who enrolled in CETA between July 1, 1975 and June 30, 1976) is matched with individuals in the March 1976 CPS. The way in which the samples are aligned is important because SSA earnings, which are used to measure the impact of CETA and are also the primary matching variables, are defined on a calendar year basis. Given that SSA earnings are measured on a calendar year basis, it is very difficult to match well on the pre-program decline in earnings experienced by CETA participants when using a fiscal year alignment, without introducing additional analytical complications. For example, Westat matches on calendar year

1975 SSA earnings rather than matching on 1974 SSA earnings and potentially miss matching on the pre-program decline in earnings experienced by CETA participants. The procedure of matching on 1975 SSA earnings results in biased estimates of the impact of CETA, however, because 1975 earnings includes up to six months of in-program earnings, which may be particularly low (e.g., for classroom training participants) or high (e.g., for PSE participants) and even includes post-program earnings for some participants. It should be noted that the change (in 1977) to an October 1 to September 30 fiscal year should reduce the bias introduced by matching on postenrollment earnings in analyzing more recent fiscal year cohorts of CETA enrollees, although the problem must be dealt with in analyzing the net impact of JTPA because program year cohorts are defined on a July 1 to June 30 basis.

A third issue relates to the specification Westat uses to estimate their weighted regression impact model. Specifically, although the participants and the comparison group members are matched on 1975 SSA earnings, this variable is not included as a control variable in the weighted impact regression model. Because the weights are highly correlated with 1975 SSA earnings and because 1975 SSA earnings are highly correlated with post-program SSA earnings, the weights are correlated with the error term in the true model for CPS cases. As shown in Dickinson, Johnson, and West (1986), this procedure leads to overstating the impact of CETA.

A fourth issue concerns decisions regarding the individuals to be included in the sample from which the comparison group is selected (i.e., sample frame decisions). An important CPS sampling frame issue involves restrictions to help ensure that the comparison group more closely matches CETA participants in terms of pre-program labor market experience or attachment. Such restrictions are necessary because many CLMS sample members (particularly women and youths) had little recent labor market experience, yet, by definition, were in the labor force when they applied to CETA. In an attempt to ensure that the comparison groups match participants on pre-program labor market attachment, Westat excludes from the CPS individuals who were not in the labor force during the interview week, unless they had worked part of the previous year. However, it is possible that such a sample criterion includes in the comparison group individuals who have dropped out of the labor force and who are much more likely to continue out of the labor force in the post-program period. This would lead to an upwardly biased estimate of CETA's net impact.

A final issue concerns decisions to exclude certain cases from the analysis (i.e., sample exclusion decisions). The major sample exclusion issue for CLMS cases involves restrictions on length of stay in CETA. Westat restricted their analysis to individuals who participated in CETA for at least eight days. Although CETA clearly cannot have a large impact on those who participate only a few days, it is possible that excluding participants who stayed fewer than eight days (14 percent of the trainee sample) could have introduced further selectivity biases into the analysis if short-term participants differed from other participants on unmeasured characteristics, such as motivation or job readiness. Evidence presented in Dickinson, Johnson, and West (1986) indicates, however, that the decision to exclude

individuals who stayed fewer than eight days does not affect Westat's net impact estimates.

A recent study by Dickinson, Johnson, and West (1986) estimates the net impact of CETA on participants' post-program earnings using an approach that differs from Westat in several respects. Their impact estimates, using 1978 SSA earnings as the outcome measure, are obtained from a CLMS sample of participants that enrolled in CETA during calendar year 1976 and that terminated by December 31, 1977. Separate matched comparison groups are drawn overall and by program activity for adult men, adult women, young men, and young women using a nearest-neighbor matching technique based on a modified Mahalanobis distance metric that has several advantages over stratified matching.[9]

In general, Dickinson, Johnson, and West find that individuals who enrolled in CETA in 1976 do not have significantly better post-program employment experiences than comparable individuals in the matched comparison groups. This conclusion is based on evidence of statistically significant but negative earnings impacts for both adult and young men, and modest but not statistically significant earnings gains for adult and young women. Adult women in PSE programs and young women in OJT programs are estimated to have experienced significant earnings gains; negative program effects are estimated for work experience participants for all age-sex subgroups. The pattern of large negative impacts for men and relatively small positive impacts for women is invariant to the use of alternative procedures and assumptions.[10] It should be noted that the negative impacts for men are not unreasonable if CETA diverted male participants from productive job search, and men might have continued to enroll in CETA despite the

---

9 The nearest-neighbor technique calculates a distance between an individual in the CLMS and each individual in the CPS, based on a number of dimensions, and then matches individuals in the CLMS to the closest individual in the CPS. The Mahalanobis distance metric essentially determines how many standard deviations apart individuals are on each variable, adjusting for the observed covariances of the matching variables, and requires that individuals be closer on variables with smaller standard deviations and allows individuals to be farther apart on variables with greater standard deviations. The Mahalanobis metric has been used extensively in developing matched comparison groups. See Cochran and Rubin (1973), and Rubin (1979).

10 Dickinson, Johnson, and West (1984a, 1984b) provide considerable evidence that the negative impacts of CETA on SSA earnings for men cannot be attributed to potential data deficiencies. Specifically, these negative impacts are due neither to the omission of uncovered earnings from the outcome measure, nor to the cap on SSA earnings or to potential contamination of the comparison group to the extent it contains some unknown proportion of CETA participants. Moreover, the results from alternative econometric models that attempt to correct for potential selection basis (i.e., differences in unmeasured characteristics between CLMS and CPS sample members) provide support for the negative impact findings for men and provide some evidence that the true impacts for adult women could be somewhat larger.

negative post-program effects because of substantial training subsidies and in-program earnings. In fact, consistent with this hypothesis, Dickinson, Johnson, and West (1984a) report positive post-program impacts of CETA for adult men who did not receive stipends.

Although the overall impacts reported by Dickinson, Johnson, and West are interesting and suggest that early CETA programs were not very effective in increasing the earned income of participants, the more important contribution of their study lies in their detailed examination of the sensitivity of CETA net impact estimates to alternative methodological procedures. In a recent summary paper, Dickinson, Johnson, and West (1985) compare the methodologies and resulting net impacts of several of the recent CETA studies discussed in this chapter. They find that the range of overall CETA impacts varies from roughly -\$300 to +\$300 and that these differences largely can be explained by the different methodological procedures used. Although it is comforting that the differences in net impact estimates can be accounted for by the different methodological procedures used, it is troublesome that the range of estimates is so large. Below we describe the various methodological procedures that appear to affect net impact estimates.

Dickinson, Johnson, and West find that CETA net impact estimates are quite sensitive to whether individuals without recent labor market experience are included in the comparison groups. In particular, they demonstrate that including individuals with less attachment to the labor force (e.g., disabled, or dropped out of the labor force, or never worked) in the comparison groups causes estimated program impacts to increase.[11] They also demonstrate that CETA impacts are sensitive to the alignment of the CPS and CLMS samples, and in particular they find that estimated impacts are more positive for early enrollees for whom the pre-program decline in earnings is more accurately measured. This indicates the importance of developing matched comparison groups on a quarterly or semi-annual basis in order to avoid biases due to misalignment.

Dickinson, Johnson and West also provide considerable evidence on the sensitivity of estimated program impacts to the matching technique used. An important finding of their study is that the estimated impacts are quite robust to the matching procedures used. Specifically, they obtain similar net impact estimates using either their overall nearest-neighbor matched groups, their closely matched nearest-neighbor by-program matched comparison groups, or using the large (unmatched) CPS eligible-for-match sample. Moreover, based on a detailed replication of Westat's analysis, they find that when other methodological factors are treated comparably, one obtains similar

---

[11] The main approach used by Dickinson, Johnson, and West (1986) requires adults in the CPS to be in the labor force in the survey week to be eligible for inclusion in the matched comparison groups. It should be noted that this is a much more stringent requirement than Westat's decision to include in the comparison groups individuals who worked in the previous year but were out of the labor force during the interview week.

estimated net impacts using either the nearest-neighbor matches or Westat's stratified cell matches. Thus, the estimated impacts do not appear to depend on the actual matching technique used. This has important implications for future analyses.

Bassi (1983, 1984) builds on the methodological work by Ashenfelter and uses the matches developed by Westat to estimate the net impact of CETA for white women, minority women, and minority men who were at least 23 years old when they enrolled in CETA in FY 1976. Her results indicate that a simple fixed-effects estimator is sufficient to eliminate the bias introduced by nonrandom selection for all subgroups except white men, although there seem to be substantial differences in pre-program earnings between high-income trainees and their matched comparison group.[12] Because of the lack of an adequate comparison group for white men, she does not estimate the impact of CETA on this subgroup. Her overall findings indicate that women experienced substantial earnings gains, with gains for white women somewhat larger. Although Bassi's fixed-effects estimator is a reasonable approach provided the underlying assumptions are satisfied, it appears that her results are biased upward because the sample is not limited to individuals who terminated by December 31, 1976, and thus the impact estimates for 1977 and 1978 include in-program earnings for CLMS cases. This is a particularly severe problem for PSE participants.

A recent report by Bassi et al. (1984) analyzes CLMS data for FY 1977 CETA enrollees to estimate the impact of employment and training programs on two key target groups for JTPA, namely youths and the economically disadvantaged. To estimate the impact of CETA on the post-program earnings of youths, the authors use matched comparison groups developed separately for each program activity. For their analysis of the impact of CETA programs for economically disadvantaged adults, they include in the comparison group all individuals from the March 1977 CPS who were economically disadvantaged. Although evidence presented on the reliability of the comparison groups indicates that considerable creaming may have occurred, the authors argue that a fixed-effects estimator is generally sufficient to obtain unbiased impact estimates.

Using both a fixed-effects and a random-effects model, and with different base years, Bassi et al. find women have larger earnings gains than men, and that PSE and OJT result in the largest earnings

---

12 If the error term in the impact model contains a fixed-effect (i.e., an unobserved component that is specific to an individual and does not vary over time--perhaps reflecting permanent differences in ability or motivation) that is correlated with other exogenous variables, then ordinary least squares estimates of a standard impact model are generally biased. The "fixed-effects estimator," or "first-difference estimator," suggested by Ashenfelter involves differencing the model over time (from the pre- to post-program period) to purge the error term of the fixed component. The fixed-effects estimator will produce unbiased program impact estimates provided certain assumptions are met (see Chapter 5).

gains.[13] The large positive impact for PSE programs, however, appears to be in part a result of including in the analysis sample individuals who were still participating in PSE programs during the outcome year. Results for youths follow the same relative pattern, although the overall impact of CETA programs on earnings appears to be negative for youths.

Another recent evaluation of CETA programs using a subset of the CLMS data base was conducted by Bloom and McLaughlin (1982). Their analysis focuses on adults (over 24 years of age) who entered either classroom training, on-the-job training, or work experience programs between January 1975 and June 1976, and who stayed in the program for more than seven days. They compared the post-program Social Security earnings of participants with the earnings of a comparison group drawn from the March 1976 CPS. Their findings indicate large positive program effects for women of between $800 and $1,300 per year that are similar across type of program, and small but statistically insignificant earnings gains for men.

Although these results are consistent with the general pattern that women gain more from CETA programs than men, the unusually large program impact observed for women, as well as the atypical result of similar impacts across programs, appears to be the result of two problems in their approach. First, no attempt was made to ensure that individuals included in the comparison group were in the labor force at the same time that individuals in the CLMS sample were enrolling in CETA. By including in the comparison group a substantial proportion of individuals who were out of the labor force (particularly for women), the results can severely overstate the impact of CETA. A second problem concerns the inclusion in the estimation model of a time-trend term that extrapolates the pre-program decline in earnings for participants into an expected further decline in the post-program period. A comparison of actual post-program earnings with expected earnings based on a trend term could result in large estimated impacts even if trainees' earnings did not return to their pre-program level. A third problem concerns the decision to exclude from the comparison group any person who earned the Social Security taxable maximum in any of the years from 1970 to 1975. This systematically excludes comparison group members with relatively high earnings, particularly those whose earnings were increasing just before the program, and given the inclusion of the trend term in the model, also results in an upwardly biased estimate of program impacts.

In a recent paper, Geraci (1984) uses Westat's matched comparison groups for FY 1976 CETA enrollees and estimates the impact of CETA on post-program earnings for adults separately for men and women. The specification allows the net impact to vary within program activity by race, age, and length of stay. The post-program outcome measure is the

---

13 A "random-effects" model assumes that the error components in the regression equation are indeed random (i.e., the fixed-effect is not correlated with other exogenous variables in the model). If this is the case, ordinary least squares techniques will produce unbiased net impact estimates.

average of SSA earnings from 1977 to 1979, and the model controls for average earnings from 1972 to 1974 and other demographic characteristics. Geraci finds significant earnings gains for adult women, but generally insignificant or negative net gains for men. Given that the net impact model is estimated by weighted least squares and does not control for SSA 1975 earnings (one of the key matching variables), however, this causes the estimated program impacts to be upwardly biased, as indicated earlier.

In addition to estimating net program impacts, Geraci also examines alternative short-term indicators of long-term earnings impacts. This analysis is motivated by JTPA's focus on long-term outcomes and the recognition that, since program management decisions regarding rewards and sanctions cannot wait for definitive long-term results, it is important to identify short-term indicators that are reliable proxies of long-term program impacts. Geraci tests several measures of short-term indicators, including immediate outcomes such as whether placed at termination and the wage rate at termination if placed, as well as short-term post-program measures (measured alternatively at three, six, or nine months after termination) such as the proportion of time employed, the proportion of time in the labor force, the average wage rate, and total earnings.

Unlike Gay and Borus, Geraci finds that placement status at termination is significantly correlated with long-term post-program earnings gains for both adult women and adult men. Like Gay and Borus, however, Geraci's results contain a potential statistical bias because he was unable to adjust for differences in local labor market conditions. His results also indicate that although the correlation between short-term post-program indicators and other measures of long-term post-program outcomes (e.g., post-program earnings, pre-post changes in earnings, and indirect net impact gains) increases considerably as the length of the follow-up period is increased from three to nine months, there is little change in the correlation between these indicators and net impact measures as the length of the follow-up period increases. Although this suggests that reasonably valid indicators of long-term net impacts could be obtained from follow-up data three to six months after termination, it would be important to obtain similar results after adjusting for differences in local labor market conditions before using his findings to make decisions on the appropriate length of follow-up.

Although all of the recent CETA net impact evaluations have been based on comparison groups developed from the CPS, the usefulness of the CPS as a source for drawing matched comparison groups has been challenged in a recent study by Fraker and Maynard (1984). Fraker and Maynard draw several different comparison groups from the CPS separately for youths and AFDC recipients who participated in the Supported Work demonstration and compare the resulting net impact estimates to those obtained when the true control group is used. Because the net impact estimates obtained from various CPS-based comparison group methodologies are generally quite different from the estimated impacts when the true control group is used, Fraker and Maynard conclude that comparison group methodologies should be used with extreme caution and that limitations of the CPS are in part responsible for the failings of this approach.

In determining the implications of the Fraker and Maynard study, it is important to keep in mind that their results are based on a very tough test of the general usefulness of the CPS for selecting comparison groups for evaluating the impact of employment and training programs. By design, Supported Work participants differ considerably from participants in other employment and training programs in that they are more likely to be high school dropouts, less attached to the labor force, and more involved in criminal activities. For example, nearly 60 percent of the youths enrolled in Supported Work had been arrested and nearly 40 percent had been convicted of criminal activities in the preenrollment period, whereas only five percent of CETA participants had been previously arrested. Because of the unique attributes of Supported Work participants, Fraker and Maynard are unable to develop comparison groups from the CPS that match experimentals well on measured characteristics.

Given that the various CPS-based comparison groups do not generally match Supported Work experimentals even on measured characteristics, it is not surprising that the net impact estimates obtained when these groups are used often differ from the impact when the true control group is used. In fact, given the unique attributes of Supported Work experimentals, it is very reassuring that CPS-based comparison groups tend to yield valid impact estimates for AFDC women, which are a very difficult group to match. Their conclusion that the CPS is not a good source for drawing matched comparison groups for Supported Work youths is in part a result of the inability to match youths on measured characteristics (e.g., criminal behavior) and may also in part be due to the difficulty of developing reliable comparison groups for youths in the absence of detailed pre- and post-program data on schooling behavior.[14]

RECENT METHODOLOGICAL DEVELOPMENTS

As the empirical literature described above has developed, there has been an accompanying methodological literature focusing on procedures that can be used to correct for various selection biases to obtain unbiased estimates of program net impacts. This literature includes the Ashenfelter paper discussed above, recent work motivated by the Ashenfelter paper, and other papers that address selection bias issues in a more general context. In this section we briefly describe some of

---

14 The large negative net impact estimates obtained for youths when each of the CPS-based matched comparison groups is used, which are not obtained when the randomly assigned control group is used, may also be due to an alignment mismatch. Specifically, because approximately 85 percent of the youths in Supported Work were enrolled after March 1976 and because more than half of the cases eligible to be included in the comparison groups came from the March 1976 CPS, this alignment causes matched comparison group members to experience the pre-program dip in earnings much earlier on average than the experimentals. As a result, regression to the mean would cause the comparison group to recover their earnings position earlier than the experimental group and bias the impact estimates downward for youths.

the more relevant methodological developments that may have some applicability to evaluating the net impact of JTPA. Because of the inherently technical nature of this discussion, uninterested readers may choose to skip this section and proceed directly to our summary of the literature and its implications.

The general literature on selection bias has developed from a concern to obtain statistically unbiased estimates of program treatment effects within a nonexperimental design. As we indicated above, in the absence of a randomly assigned control group, it is necessary to statistically control for all characteristics that influence the likelihood of selection into the treatment group when measuring program effects. If all variables that affect selection into the treatment group are included in the regression equation that measures program impacts, or if those omitted are unrelated to the outcome, then one would obtain an unbiased measure of the treatment effect. If one has not properly accounted for the variables (observed or unobserved) that affect selection into the treatment group, then the treatment variable will be correlated with the error term in the impact equation and ordinary least squares estimates will be biased.

To solve the problem of the correlation between the outcome measure and the error term in the regression equation, instrumental variable procedures can be used. Papers by Heckman (1979), Maddala and Lee (1976), and by Barnow, Cain, and Goldberger (1980) provide instrumental variable approaches that can produce consistent estimates of program impacts under certain circumstances or assumptions. Each procedure requires the creation of a predicted value of program participation from a prior probit equation and inclusion of either the predicted program participation variable, the inverse of Mills' ratio, or another constructed variable based on the probit results, in the earnings regression.[15] Although these instrumental variable procedures produce consistent estimates of program impacts, they are only useful if one can find variables that affect participation in the treatment group <u>but do not affect</u> earnings or other outcome variables of interest. Unfortunately, in most cases the instrumental variable created will be correlated with other variables in the equation, making it extremely difficult to obtain precise estimates of treatment effects.

Perhaps a more useful approach to solving selection bias problems is to take advantage of the longitudinal nature of the data to control for potential differences between the treatment and the comparison groups. Such an approach builds on Ashenfelter's generalized difference estimator using detailed pretraining earnings data from participants and nonparticipants, as discussed earlier. The notion of using the longitudinal nature of the data has also been extended by Heckman and Robb (1982, 1985). For example, Heckman and Robb (1982) demonstrate that a difference estimator that is symmetric about the year in which the decision to enroll in training is made can provide unbiased estimates of program impacts even when the random component of earnings

---

[15] For a description of the probit technique see Finney (1964). The construction of the inverse of Mills' ratio is described in Heckman (1979).

is correlated over time. That is, if certain assumptions regarding the program participation decision and the error term in the earnings equation are met, one can obtain unbiased net impact estimates by regressing earnings in year s+t less earnings in year s-t on a participant dummy variable and the difference in exogenous variables between these two periods, where s is the year in which the decision to enroll in training is made, and t represents the number of years after s for which post-program earnings are being measured.

As demonstrated in Dickinson, Johnson, and West (1986), net impact estimates obtained from the symmetric difference estimator are very sensitive to the choice of the decision year. Moreover, it is likely that the appropriate decision year differs among participants, depending on the date of enrollment into training. That is, for individuals who enroll early in a given year, one can reliably use the prior year as the decision year; for individuals who enroll late in the year, it is likely that the enrollment year is the decision year. This problem would be less serious if one could measure earnings on a quarterly or semi-annual basis.

The more recent paper by Heckman and Robb (1985) details the assumptions required to use various cross-section and longitudinal models to estimate the impact of training on earnings free of selection bias. This paper is important because it provides a framework for comparing the assumptions used by previous researchers that were not often stated explicitly. It is likely that some of the differences in net impact estimates reported above are due to differences in assumptions regarding the underlying structure of earnings, the decision rule governing program participation, the time homogeneity of the environment or the distribution of unobservables. It is important to carefully describe the assumptions underlying the proposed state-level net impact model, so that others can evaluate the reasonableness of the assumptions and potentially determine the extent to which different impact estimates are due to different underlying assumptions.

SUMMARY AND IMPLICATIONS

In this chapter we have reviewed several recent evaluations of the impact of employment and training programs on participants' post-program outcomes, as well as briefly summarized a few recent contributions to the methodological literature. The results of these studies generally indicate large earnings gains for women, particularly nonblack women, whereas the effect of employment and training programs on the earnings of adult men is less clear. Although almost all studies have found the earnings gains of men to be considerably less than those obtained by women, several recent evaluations have found that male trainees never regain the earnings position they held prior to training relative to otherwise comparable nonparticipants. Why did men continue to enroll in these employment and training programs? Perhaps because of in-program earnings and the substantial training subsidies offered by CETA.

These studies have several important implications for designing a net

impact evaluation model of JTPA programs. Some of the implications concern the conceptual framework to be used (e.g., outcome measures, subgroups and treatments to be examined), and others relate to research design and data analysis issues (e.g., issues concerning comparison group selection such as the alignment of samples, the sample frame choice and sample exclusion decisions, and the statistical models used). Below we briefly summarize these studies and their implications as they relate to developing a conceptual framework for evaluating JTPA at the state level. This brief summary serves as a natural introduction to the next chapter, which discusses the conceptual framework for a state-level net impact model. The implications of previous studies concerning research design and analysis issues are discussed in Chapters 4 and 5 as appropriate.

In Table 1 we summarize several key features of the employment and training impact evaluations described above. In particular, we list the outcome measures examined, the participant groups selected, the subgroups of trainees for which separate net impact models were estimated, and the ways in which the treatments were included in the models. As this table indicates, annual earnings has been the primary outcome measure, which is very consistent with recent federal employment and training legislation. Moreover, Social Security Administration records have been the main source of earnings data. Although SSA earnings records have several advantages (e.g., they are a cost-effective source of longitudinal data that are measured comparably for participants and comparison group members), they have several potential disadvantages including coverage problems, exclusion of earnings beyond the taxable maximum, and considerable delays in obtaining reliable data (of up to three to four years). Moreover, when SSA earnings is the only outcome measure available, the evaluation is limited to estimating impacts on an annual basis and, except for examining the probability of working at all during a year (indicated by positive SSA earnings), it is not possible to examine the effects of employment and training programs on the components of earnings (e.g., hourly wage rate, hours worked per week, weeks worked per year). As the first column in Table 1 indicates, the only previous studies that were able to examine the impact of employment and training programs on the components of earnings were those for which relatively expensive primary data collection efforts were undertaken. Moreover, information on short-term earnings impacts--those within approximately three to six months after termination--can only be provided through primary data collection efforts or through the use of UI earnings records.

The second column of Table 1 indicates the subgroups of participants for whom separate impact models were estimated. These studies suggest that net impact models should be developed separately for men and women. This is in part because the relationship between earnings and various socioeconomic and demographic characteristics is different between men and women and also because of the considerable evidence indicating that employment and training programs result in sizable earnings gains for women (particularly nonblack women), whereas the impact on earnings for men is consistently less, and in several instances has been estimated to be negative.

Although the evaluations of MDTA programs tended to estimate separate

Table 1

SUMMARY OF SELECTED FEATURES OF PREVIOUS IMPACT EVALUATIONS

(s)	Outcome Measure(s)	Subgroups	Participant Group(s)	Treatment Measures
elter 1978)	SSA earnings	White males Black males White females Black females	Jan. 1964 - March 1964 MDTA classroom training enrollees	Participant dummy
, McGuire, and tt (1979)	SSA earnings and whether employed (SSA greater than $50)	Males Females	1969 - 1971 MDTA classroom training and OJT enrollees	Participant dummy
(1979)	Interview-reported annual earnings	Separate major program types	1968 - 1970 MDTA classroom training, JOBS, Job Corps, and NYC enrollees	Participant dummy for separate program models
(1978)	Interview-reported weekly earnings and employment status	White males Black males	1968 - 1970 MDTA classroom training enrollees	Weeks participated (and weeks-squared)
d Borus	SSA earnings	White males Black males White females Black females (separately for each major program type)	1968 - 1970 MDTA classroom training, JOBS, Job Corps, and NYC enrollees	Participant dummy and weeks participated interacted with all socioeconomic characteristics for separate program models
er (1978)	Interview-reported annual earnings and welfare grants	Males Females	Spring 1974 WIN enrollees	Participant dummy (sometimes interacted with services received)
s (1981)	Interview-reported earnings, employment, hours worked, wage rate, welfare grant	None	Supported Work long-term AFDC recipients	Participant dummy
et al.	Interview-reported earnings, employment, hours worked, hourly wage, schooling, criminal activity	Males Females without children Females with children	April 1977 Job Corps participants	Participant dummy
and dorfer (1970)	Interview-reported wage rates, earnings and employment status	White males Black males White Females Black Females	1965 - 1967 NYC/IS and NYC/OS enrollees	Participant dummy
Brennan, and (1970)	Annual UI earnings	None	1966 NYC/OS terminees in Indiana	Participant dummy (sometimes interacted with hours in program and sex)
ner Assoc.	Interview-reported earnings and wage rates	None	CEP participants	None

Table 1 (cont.)
SUMMARY OF SELECTED FEATURES OF PREVIOUS IMPACT EVALUATIONS

	Outcome Measure(s)	Subgroups	Participant Group(s)	Treatment Measures
)	SSA earnings	None	FY 1976 CETA enrollees who stayed at least 8 days and terminated by Dec. 31, 1976	Participant dummy (sometimes interacted with program activity dummies or participant characteristics)
ohnson, 84a)	SSA earnings	Adult males Adult females Young males Young females (sometimes separately for each major program type)	Calendar year 1976 CETA enrollees who terminated by Dec. 31, 1977	Participant dummy (sometimes interacted with participant characteristics)
1984)	SSA earnings	Adult minority males Adult white females Adult minority females	FY 1976 CETA enrollees who were at least age 23	Participant dummy (sometimes interacted with program activity dummies)
(1984)	SSA earnings	Youths: White males Black males Hispanic males White females Black females Hispanic females Economically Disadvantaged: White males Minority males White females Minority females	FY 1977 CETA youth enrollees (under age 23), and economically disadvantaged and welfare recipient enrollees	Participant dummy (sometimes interacted with program activity dummies)
1982)	SSA earnings	Adult males Adult females	Jan. 1975 - June 1976 CETA enrollees who were at least age 25 and stayed at least 8 days	Participant dummy (sometimes interacted with program activity dummies)
)	SSA earnings	Adult males Adult females	FY 1976 CETA enrollees who were at least age 22, stayed at least 8 days, and terminated by Dec. 31, 1976	Program activity dummies interacted with race dummies, age, and a quadratic in length of stay in days

net impact models for whites and blacks, more recent evaluations have tended to estimate separate models by age groups. In fact, because of the difficulties involved in developing valid net impact estimates for youths, many recent studies have estimated models only for adult men and women. The problems for youths are primarily (1) that earnings is not the appropriate outcome measure for individuals who may return to school--the relative mix of schooling, market work, nonmarket work, and leisure changes rapidly over time for youths--and (2) that it is very difficult to draw a reliable matched comparison group for youths with limited and highly variable earnings histories. Moreover, a considerable amount of pre- and post-program data on schooling behavior are required to obtain valid net impact estimates, much more data than are generally available on existing data sets. As such, it may not be feasible to develop a state-level JTPA net impact model for youths that provides valid results and can be implemented in a cost-effective manner.

Finally, in the last two columns of Table 1 we summarize the participant groups chosen for analysis and the variables included in the model to measure the treatment effects. For the most part, these studies focused on estimating the average impact of program participation on earnings for the selected subgroups. Because in many cases the subgroups of interest were specific program activities, this resulted in numerous net impact estimates by program (treatment) type. The only other dimension of the treatment that was examined in a few of these studies is length of program participation. The results from these studies indicate that net impacts vary by program activity and length of stay. Although fewer programs are generally offered under JTPA (as compared with CETA), and the average length of stay in JTPA is much less than in CETA, it will still be important to develop models that examine these potential differences.

CHAPTER 3
CONCEPTUAL FRAMEWORK

General Research Questions

Participant Outcomes

Earnings and Employment Outcome Measures

Welfare Outcome Measures

Participant Subgroups

Program Activities

Program Environmental Conditions

CHAPTER 3. CONCEPTUAL FRAMEWORK

An important element in the design of a state-level JTPA net impact model is the conceptual framework. The conceptual framework defines the scope and focus of the analysis and guides the research design and analysis methods. In particular, the conceptual framework identifies the key research questions to be addressed, the outcomes to be examined based on those questions, the participant groups and program activities (treatments) to be included to address the questions of interest, and the specific definitions of the outcomes, treatments, and variables that affect the relationship among treatments and outcomes. In this chapter we describe a conceptual framework for conducting a state-level JTPA net impact analysis.

In describing the various components of the conceptual framework, we first discuss the alternatives considered and then indicate our recommendation. Each recommendation is in part based on the results of the literature review described in Chapter 2. This enables us to focus on the major outcome measures, treatments, and participant groups of interest so that key research questions that have been examined in the literature can be replicated. In addition, our recommendations are in part based on the recognition of the severe resource constraints that states and SDAs face, as well as certain other practical considerations such as data availability. The two most important practical considerations that affect the conceptual framework are: (1) states and local SDAs will not generally be able to implement an experimental design; and (2) states and local SDAs will not generally be able or willing to conduct follow-up interviews with a large sample of participants <u>and</u> comparison group members. Although these considerations directly affect the comparison group strategy described in Chapter 4, they also indirectly serve to limit the scope of the conceptual framework as described below.

GENERAL RESEARCH QUESTIONS

In specifying the major research questions that an analysis of the net impacts of JTPA programs should address, it is important to recognize that JTPA may affect different groups in different ways and at different levels. For example, one could specify the general research questions from any one of the following four perspectives: participants, employers, the government, or society as a whole. The questions of interest vary considerably among these four groups. Society as a whole is most concerned with whether JTPA results in a net increase in the resources available (i.e., increases output--GNP). Government is primarily concerned with whether JTPA increases tax revenues and reduces transfer payments. Key employer questions concern whether JTPA reduces hiring and training costs, or increases productivity. At the individual-participant level, the key research

questions relate to improvements in an individual's post-program labor market experiences. Because evaluating the impact of JTPA from the social perspective is beyond the scope of the state-level model, and because employer benefits are being examined separately in Volume 4 in this series, in this section we discuss the general research questions that should be included in an analysis of the net impact of JTPA on participants' post-program labor market experiences.[16] Research questions that are of concern to state and local governments are included only to the extent that JTPA programs affect the likelihood that trainees receive welfare payments.

Our primary goal in developing a state-level JTPA net impact model is to determine the extent to which JTPA employment and training programs improve the labor market experiences of participants relative to what their experiences would have been in the absence of the program, that is, relative to a comparison group of otherwise similar nonparticipants. The average net impact of JTPA programs on participants' post-program labor market experiences will give policy-makers an indication of the overall effectiveness of these programs that will more than meet a state's increased accountability responsibilities under JTPA. Moreover, since most previous analyses primarily focused on estimating average impacts for the nation as a whole, this will provide a benchmark for comparison purposes.

In order to maximize the amount of policy-relevant information, however, it is important that the analysis be designed in such a way that it gives policy-makers a disaggregated view of how the program works and the effects it has on different types of participants. That is, although it is important to know whether the mix of JTPA programs is effective on average, for policy purposes it is perhaps more important to probe beneath the average impacts to identify the relative effects on different subgroups. Moreover, because of the many changes from CETA to JTPA (both in terms of program services provided and characteristics of participants served), it is important to identify the composition of net program impacts in order to determine the comparability of the findings with previous studies. Thus, as we describe below, the net impact model will be designed to address research questions concerning two key dimensions: (1) whether certain participant groups benefit more from JTPA than other subgroups, and (2) whether the net impact of JTPA differs among program activities (treatments) for the program as a whole and for certain demographic groups.

---

16 It should be noted that positive net impacts on participants' labor market experiences is a necessary condition for the program to be considered successful from the point of view of government or society as a whole. That is, only if JTPA increases participants' earnings, and there are no off-setting displacement effects (i.e., the earnings of nonparticipants are not reduced), can government obtain additional tax revenues or can society as a whole have access to additional goods and services. Thus, our focus on the net impact of JTPA on participants' post-program outcomes can be viewed as the first step in a more comprehensive evaluation of JTPA.

The analysis of the net impact of JTPA programs by characteristics of participants will determine whether the program is more effective for some types of individuals than others. This analysis of the variation of program impacts will provide important information about the mechanisms through which the program produces its effects. It will also provide information about the generalizability of the results and the comparability of the results relative to previous studies. For example, if the positive impacts are concentrated only among specific groups--for instance, those with very low pre-program earnings--this may have implications for future targeting practices. Moreover, if the average impacts are very different from those reported in previous studies, the subgroup analysis may indicate reasons why. For example, the net impacts may be largest for a subgroup that was not previously targeted for training services.

The program activity analysis will examine key questions about the relative effectiveness of various JTPA activities. This analysis will shed light on whether the average effects of JTPA are homogeneous across program activities, whether all program activities result in improved labor market outcomes of participants, and which program activities result in the largest net benefits. Although such analyses will provide important information on relative program effectiveness that may have implications for targeting practices, one must recognize that the costs of various program activities differ considerably and, in the absence of a cost analysis, no conclusions can be made about which program activity is most cost-effective.

The state-level net impact model will also examine key questions related to the timing of net impacts and the extent to which net impacts vary by length of program participation. Such analyses will provide additional information on the mechanisms through which JTPA produces its effects. Information on the timing of net impacts is necessary to distinguish whether (a) the program has only a short-term effect, from (b) the impact is expected to persist in the long-term. This is particularly important when making judgments on the long-term cost-effectiveness of various programs. The differential analysis by program length of stay will help distinguish whether (a) program activities generate different impacts and length of stay is unimportant, from (b) length of stay in any program is important, rather than the particular program activity. Such information should be of considerable use to program operators who have limited information to accurately judge the outcomes of their programs.

Finally, the model will be designed to address important research questions concerning the extent to which net impacts vary by local environmental conditions such as the unemployment rate or urban/rural location. Because previous national studies did not have the necessary data to address such issues, it may be possible for the state-level model to make a unique contribution to what is known about the effectiveness of employment and training programs that operate in different environmental conditions.

The general objectives of the state-level net impact model can be summarized by a series of key research questions to be addressed:

- What is the overall net impact of JTPA programs on participants' post-program labor market experiences?

- How do the net impacts change over time?

- For which program activities (treatments) are the net impacts the largest? That is, which program activities result in the largest net benefits to participants?

- For which groups of participants are the net impacts the largest? That is, which subgroups gain the most from participating in JTPA?

- Do individuals who remain in JTPA longer experience greater net gains in labor market outcomes?

- How does the net impact of JTPA vary by local program environmental conditions?

These are the general research questions that the state-level JTPA net impact model will examine. Below we begin to make these questions more specific by indicating the outcomes that will be used to measure participants' labor market experiences, by indicating the participant subgroups to be examined, by identifying the types of program activities that will be examined, and by identifying the local program environmental conditions of interest.

PARTICIPANT OUTCOMES

In selecting the appropriate participant outcomes to be examined in a net impact analysis of JTPA, one should begin with the goals set forth in the legislation. As indicated earlier, Section 106 of JTPA explicitly states that "job training is an investment in human capital," and that the "return on the investment" for adult training programs funded under Title II-A is to be measured by the "increase in employment and earnings and the reduction in welfare dependency resulting from participation in the program." The Act also provides examples of indicators of these basic outcome measures that include placement in unsubsidized employment, job retention, and hourly wage rates. In addition to the three outcomes listed for adult programs, the Act also specifies other outcomes for youths aged 16-21, including: attainment of employment competencies recognized by local private industry councils (PICs); completion of a major level of schooling or the equivalent (e.g., elementary, secondary, or postsecondary); and enrollment in other training or apprenticeship programs or enlistment in the Armed Forces.

In determining the general outcome measures that are to be used to represent the changes in the labor market experiences of participants due to JTPA, it is important to keep several factors in mind. First, one must ensure that the outcome measures used are consistent with the objectives stated in the legislation. Second, to the extent possible, the outcome measures should be consistent with those examined in previous studies so that key research questions can be replicated.

Third, in order to be used in a net impact analysis, the outcomes must be available and comparably measured for participants and comparison group members. Finally, the outcomes selected must recognize the limited resources available to states and local SDAs for conducting evaluation activities. Taken together, the latter two factors suggest the importance of relying on agency administrative data to the extent possible.

Based on these considerations, we recommend that the general participant outcome measures in a state-level JTPA net impact model include:

- Whether employed;
- Earnings;
- Whether receiving welfare grants; and
- Welfare grants received.

These outcome measures are consistent with the major objectives of the legislation and for the most part greatly exceed the outcomes examined in previous national studies of the net impact of employment and training programs. Moreover, as we describe below, comparable data on these measures can be obtained for participants and comparison group members from state administrative records, a cost-effective data source. Although potentially useful information on the mechanisms through which employment and training programs increase earnings could be provided by examining other outcome measures such as employment intensity (e.g., hours worked per week, weeks worked per year), hourly wage rates, job retention, or the additional outcomes listed in the legislation for youths, such outcome measures require the collection of survey data from both participants and comparison group members, which would greatly increase the resources required to implement the net impact model, and thus reduce its practical usefulness.[17] However, states with additional resources, or particular interest in some of these other dimensions of labor market experiences, are encouraged to collect the necessary survey data and follow the research design and analysis plans described in subsequent chapters to estimate the net impacts of JTPA on these additional outcomes.

---

17 An additional outcome measure that could also be created from state administrative records is receipt of Unemployment Insurance (UI) payments. Although this outcome measure may be very appropriate for Title III programs, we excluded it from the core set of outcomes for the Title II-A net impact models because: (1) it was not included in the legislation as a key goal; (2) it was not used as an outcome measure in any of the employment and training net impact studies described in Chapter 2; (3) any potential impact is likely to be small because only approximately ten percent of JTPA applicants are UI claimants; and (4) the development of measures of UI payments received over time generally requires the accessing of the payment-history file, which can be a very expensive process.

As indicated above, in constructing the core outcome measures we recommend that available program administrative records be used whenever possible. In particular, we recommend that the earnings and employment measures be constructed from State Unemployment Insurance Wage Records and that measures of welfare grants received and welfare dependency status be created from State Welfare Administrative Grant Records. Below we describe the advantages of these data sources, their limitations, and the specific outcome measures to be created.

Earnings and Employment Outcome Measures

State UI Wage Records are an excellent source for developing measures of earnings.[18] In fact, they have several advantages as compared to other administrative data sources and to survey data. For example, unlike SSA earnings records, that have been used in most previous studies, all earnings in covered employment are reported in the UI wage records (i.e., earnings are not "top-coded" at the Social Security taxable maximum), and data are available on a relatively timely basis (usually within three months of the end of the quarter of interest). Moreover, unlike SSA data that are only available on an annual basis, UI wage records are available on a quarterly basis. This enables one to construct both short- and long-term outcome measures to better examine the timing of JTPA impacts. Finally, although survey data allow for much more flexibility and a greater range of outcomes, UI wage records are not subject to interviewer biases, to problems that arise from some respondents reporting net (after-tax) earnings and others reporting gross (before-tax) earnings and, for the most part, they are not affected by response-rate problems that can plague survey data collection efforts.

The primary disadvantage of using State UI Wage Records for developing earnings measures is that these data are not available in all states. For example, in 38 states and the District of Columbia, employers are required to report the quarterly earnings of all employees covered by State Unemployment Insurance laws. These states are known as "wage-reporting states." The remaining 12 states are known as "wage-requesting states," and they request wage information from employers only for individuals who file a UI claim. The current list of wage-reporting and wage-requesting states is provided in Table 2. Although one could not currently construct meaningful earnings measures for all participants and comparison group members from UI records in wage-requesting states (since they are only available for individuals who apply for UI--a very nonrepresentative sample), the Deficit Reduction Act of 1984 requires that all states effectively become wage-reporting states by 1988. Thus, by 1988 all states should be able

---

18 Although State UI Wage Records are an excellent source for developing measures of earnings, they are generally not a very good source for developing measures of employment intensity. In almost all states, the only employment measure that can be developed is whether the individual worked in a given quarter. The only exception is Washington State, which also has information on the number of hours worked in the quarter.

Table 2

CURRENT LISTING OF WAGE-REPORTING AND WAGE-REQUESTING STATES

Wage-Reporting States		Wage-Requesting States
Alabama	Mississippi	Hawaii
Alaska	Missouri	Massachusetts
Arizona	Montana	Michigan
Arkansas	Nevada	Minnesota
California	New Hampshire	Nebraska
Colorado	New Mexico	New Jersey
Connecticut	North Carolina	New York
Delaware	North Dakota	Ohio
Florida	Oklahoma	Rhode Island
Georgia	Oregon	Utah
Idaho	Pennsylvania	Vermont
Illinois	South Carolina	Wisconsin
Indiana	South Dakota	
Iowa	Tennessee	
Kansas	Texas	
Kentucky	Virginia	
Louisiana	Washington	
Maine	West Virginia	
Maryland	Wyoming	

to use the net impact model to examine earnings gains using UI Wage Records. Moreover, some wage-requesting states (e.g., New York) could use the model at this time by accessing comparable earnings records maintained by the Department of Revenue, provided the necessary interagency agreements could be worked out.

Other disadvantages of using State UI Wage Records for developing earnings measures include potential nonreporting problems, problems caused by individuals living near state borders, and differences in state practices for collecting and retaining these data. The nonreporting problems (e.g., wages of employees of federal, state, or local governments and self-employed individuals are generally not reported), are not major: approximately 80 percent of all state wages are included in State UI Wage Records. Thus, earnings measures constructed from these data should be comparable among participants and comparison group members who work within a given state. Because the system is state-based, however, it is impossible to distinguish individuals who work across the border in a different state from individuals who do not work in covered employment. Thus, unless inter-state agreements can be worked out to access UI Wage Records, there may be some problems in estimating the net impacts of JTPA on earnings for large SDAs located near state borders.[19]

The final issue, retention of UI Wage Records, is particularly important because detailed pre-program earnings information is critical for obtaining valid estimates of the net impacts of JTPA on post-program earnings. Because the wage records are created for determining UI eligibility and benefit level, they are actively maintained for at least the "base period" used to determine UI eligibility and benefits in the state, which generally varies from four to six quarters. In many states (including the State of Washington), however, these wage data are actively maintained for at least three years. Although procedures can be designed to access adequate pre-program earnings information for analysis even in states that have only four quarters of data available at any one time, these procedures are operationally awkward and potentially burdensome. Instead, states that are seriously interested in conducting a net impact analysis might consider adjusting their archiving and retrieval practices so that at least three years of UI wage records are easily accessible.

Using State UI Wage Records, measures of pre- and post-program employment status and earnings will be developed for both participants and comparison group members. In selecting the specific periods of measurement for the earnings and employment outcome variables, we were guided by several factors. First, the periods of measurement should be consistent with those examined in previous studies so that key research questions can be replicated. Because, as described in Chapter 2, most previous studies examined net impacts on earnings approximately one to

---

19 It should be noted, however, that because net impact estimates are based on differences in earnings between participants and comparison groups members, as long as JTPA does not affect the probability of moving out of state or working across state borders, this problem should not bias the results.

two years after program termination, it was felt that the state-level model should include outcome measures at least a year after termination. Second, to examine the timing of impacts, some short-term and intermediate measures must be developed. Finally, in selecting the specific measurement periods, it must be recognized that the UI wage records are available only on a quarterly basis.

Based on these considerations, we recommend that the state-level JTPA net impact model examine the following earnings and employment outcome measures using UI Wage Records:

- Quarterly, semi-annual, and annual earnings; and

- Quarterly, semi-annual, and annual employment status (i.e., whether employed--based on whether UI Wage Records are positive in the particular period).

These outcome measures capture the range of short-term and relatively long-term impacts that could be observed within a two-year program analysis cycle and will provide valuable information on the benefits to participants from JTPA. States that are interested in additional information on the duration and timing of program benefits should consider including annual earnings in a second post-program year as another outcome measure.

Welfare Outcome Measures

The other major outcome measures to be examined relate to welfare dependency. We recommend that these outcome measures be developed from State Welfare Administrative Payment Records, which appears to be the only cost-effective source of welfare data for both participants and comparison group members. However, unlike the State UI Wage Records that are quite consistent across the wage-reporting states, there is considerable variation in state and local welfare administration and record-keeping practices, as well as differences in the degree of data automation and retrieval capabilities. As such, it will clearly not be feasible for all, or perhaps even for most states to implement this component of the net impact model. Nevertheless, we believe that states with the capability and interest can learn much about the net impact of JTPA programs on welfare dependency by implementing this component of the model. Below we briefly discuss some of the issues involved in accessing and using welfare administrative records to construct reliable measures of welfare dependency and indicate the recommended outcome measures.

In developing operational definitions of outcomes to measure the reduction in welfare dependency due to JTPA, it is important to keep in mind the focus of the legislation. JTPA explicitly refers to measures of reductions of the number of individuals and families receiving cash welfare payments as well as the amounts of these payments. Because there are several public assistance programs that provide cash welfare payments, one is immediately confronted with the issue as to which welfare programs to include in the analysis. The primary cash

assistance programs include Aid to Families with Dependent Children (AFDC), General Assistance (GA), and Supplemental Security Income (SSI). It is important to recognize that these programs are administered at various levels of government and have different eligibility criteria, payment levels, and potential lengths of participation. Because of these differences, the programs often do not share a common data base, which leads to additional complications for analysis purposes. As a result of these complications, and recognizing that reduction in AFDC grants is the most policy-relevant potential outcome of JTPA as it relates to welfare dependency issues, we recommend that the welfare dependency outcome measures only be concerned with the AFDC program.

It must be recognized, however, that even after limiting the scope of the analysis to AFDC grants, it will generally be possible to implement this component of the net impact model only in states that administer the AFDC program (i.e., not in states in which the program is county-administered) and that have sufficiently sophisticated data systems so that necessary pre- and post-program AFDC grant records can be retrieved. Moreover, even in state-administered AFDC programs with automated grant records available at the state level, several additional complications must be dealt with. First, although the program data system should contain the actual AFDC grant to recipients for the current month, grant amounts for earlier months may be very difficult to obtain in some states and will generally require searching a large historical file. The focus of the data system on recording current grants to AFDC recipients is easily understood because of the program's emphasis on providing assistance to the current case load. For analyzing the net impact of JTPA on AFDC participation, however, historical grant records (over at least two to three years) are necessary to control for potential differences between participants and comparison group members on pre-program AFDC participation. As such, the model cannot be reliably implemented for AFDC programs without detailed historical information unless reasonably expensive personal surveys are administered to obtain both pre- and post-program AFDC grants for participants and comparison group members.

A second potential complication involves the ability to identify the AFDC grants received by specific JTPA participants and comparison group members. In most states, a welfare case number is assigned to identify the "assistance unit," and the history file contains grants to that assistance unit over time. Problems arise, however, because not all state AFDC data systems have the Social Security number (SSN) for every member of the assistance unit. That is, because the SSN is the only link between an individual in the analysis sample and his/her welfare data, it is very difficult to determine whether a specific individual is receiving AFDC when the SSNs are not available for all members of the assistance unit. Moreover, even if it is possible to determine whether a specific person is currently in the assistance unit, because the history file usually corresponds to an assistance unit and does not have information on who is in the unit over time, it is possible that the specific individuals of interest (i.e., JTPA participants and comparison group members) may not have been in that unit in earlier months. As a result, the pre-program AFDC history for the unit may not accurately reflect a person's welfare recipiency status during this

period. This is particularly a problem for individuals who experience a marriage or divorce, or who change living arrangements.

In creating AFDC dependency outcome measures it is also important to recognize the implications of the earned income disregard provision of the program. Specifically, during the first four months after an AFDC recipient begins working, the "$30 and 1/3 rule" is applied. This provides that the first $30 of earnings and one-third of the remainder will not reduce the welfare grant. In addition, work expenses (up to a maximum of $75) and child-care costs (up to a maximum of $160 per child) are deducted from earnings before the earnings are counted against the welfare grant. Thus, for the first four months, AFDC recipients can earn a considerable amount with a relatively modest reduction in the grant. After the first four months, however, the grant is reduced dollar-for-dollar after work and child-care expenses are deducted.[20] Because of the four-month cutoff of the earned income disregard provision, and because the AFDC system generally takes a month or more to adjust grants to reflect earned income, one might consider adjusting the outcome periods for the AFDC dependency measures and make them somewhat longer than for the employment and earnings measures. This would be particularly useful in states with relatively brief JTPA program treatments. However, in states with average-to-long JTPA treatments (e.g., four months or longer), the earned income disregard should not introduce any serious analytical complications for post-program net impact analysis.

Based on these considerations, we recommend that the outcome measures related to reductions in welfare dependency be as follows:

- Quarterly, semi-annual, and annual AFDC grants received; and
- Quarterly, semi-annual, and annual AFDC participation status based on whether AFDC grants were received during those periods.

Because the AFDC grants data are available monthly, and the UI Wage Records are only available quarterly, the AFDC measures will be aggregated over the appropriate months to correspond to the same calendar quarters as the post-program earnings measures developed from UI Wage Records. These measures capture the short-term and relatively long-term welfare dependency impacts that could be observed within a two-year program analysis cycle and are consistent with the objectives of the legislation. It should also be noted that the six-month measure is consistent with the recommendation of a recent study for DOL concerning the measures to be used in setting post-program performance standards for Title II-A programs for adult welfare recipients (Berkeley Planning Associates, 1984).

---

20 It should also be noted, that if the person earns 150 percent or more of the state's standard of need, she or he will automatically be removed from the welfare rolls without any transition period of reduced grant levels.

PARTICIPANT SUBGROUPS

The next conceptual framework issue concerns the participant groups for which the outcome measures described above will be examined to determine the net impacts of JTPA programs. Several issues concerning participant groups must be addressed. For example, it must be decided whether all JTPA participants are to be included or whether the average net impact estimates are to be based on a sample of participants that excludes certain types of individuals. In addition, decisions must be made on the participant subgroups for which separate net impact estimates will be derived to identify the types of individuals who gain most from participating in JTPA. A related issue concerns the subgroups for which entirely separate net impact models are to be estimated. Finally, operational definitions of the participant characteristics to be included as control variables in the net impact model must be developed. Because the specific individual variables to be included in the model must be available and comparably measured for both participants and comparison group members, the operational definitions of individual characteristics are constrained by the intersection of common elements in the JTPA MIS and in the agency data base for the comparison group selected (i.e., assuming surveys of comparison group members are not cost-effective). As such, this issue depends heavily on the specific comparison group strategy followed, and we therefore defer our discussion on variable definitions to Chapter 4. In the remainder of this section we discuss issues related to participant groups to be included in the analysis and the subgroups for which separate net impact estimates are to be developed.

The first issue concerns whether the net impact model should be based on all JTPA participants or whether certain participant subgroups should be excluded. Our recommendation is to develop a net impact model only for adults. This recommendation is in part because the outcome measures listed above (earnings, employment, and AFDC dependency) are not appropriate for youths (particularly in-school youths), and no existing data sets include information on more appropriate outcome measures (e.g., schooling attainment, employment competencies) for both participants and comparison group members. In addition, our recommendation is in part due to the extensive evidence presented in Chapter 2 documenting the difficulty researchers have had in developing a reliable matched comparison group for youths who have limited, and highly variable, pre-program earnings histories. Rather than design an approach that is highly likely to yield biased and misleading results for youths, we recommend that the net impact model be developed only for adults.

The omission of youths from the net impact model is unfortunate as JTPA requires that 40 percent of all funds be expended on this important target group. It must be pointed out, however, that the gross impact evaluation guide (Volume 4) contains more relevant outcome measures that can be examined to determine the gross impacts of JTPA for youths. In addition, states that are very interested in developing net impact estimates for JTPA Title II-A youth programs might consider implementing an experimental design, or alternatively, administering relatively expensive interviews of participants and comparison group members to collect the detailed pre-program and post-program employment

and schooling data necessary for reliable analysis.

The second issue concerns whether separate net impact models need to be developed for any specific adult subgroups. As described in Chapter 2, most previous studies estimated separate net impact models for men and women; some of these studies also estimated separate models for youths and adults, and others estimated separate models by race or ethnicity group. Separate net impact models should be estimated for subgroups of participants that have different earnings functions, that is, that have a different relationship between earnings and other demographic characteristics. We recommend that separate net impact models be developed for adult men and women because of extensive evidence indicating that the relationship between earnings and other demographic characteristics is very different for these two groups. On the other hand, as described in Chapter 2, we do not believe it is necessary to estimate entirely separate net impact models for other subgroups such as race or education. This is because the earnings functions for these subgroups of employment and training participants are generally not sufficiently different to warrant the reduction in sample size and statistical power that would occur by estimating entirely separate models, and because the major differences between the groups can usually be accounted for through including appropriate interaction terms in the net impact model.

Although completely separate net impact models are not recommended for other subgroups, as we describe in Chapter 5 the analysis can be designed to examine whether the net impact of JTPA differs among subgroups of interest. Based on the evidence presented in Chapter 2, it will be important to investigate whether the impact of JTPA varies by the following participant characteristics:

- Age (e.g., less than or equal to 35 as compared to over 35);
- Ethnicity (whites as compared to blacks and Hispanics);
- Educational level (at least a high school graduate as compared to others);
- Marital status (married as compared to unmarried); and
- Welfare status (welfare recipients as compared to nonrecipients).

To the extent possible, we recommend that net impact estimates be derived for these subgroups of adult men and women in order to replicate the research questions examined in previous studies and to provide valuable information on targeting issues. It must be recognized, however, that the extent to which these subgroup impacts can be reliably estimated depends on (1) the availability of the data items for both participants and comparison group members; (2) the availability of sufficient sample sizes to generate reasonably precise net impact estimates; and (3) the availability of adequate state resources to support the additional data processing and analysis required.

PROGRAM ACTIVITIES

Another important element of the conceptual framework is the determination of the key program activities (treatments) to be examined and the development of consistent definitions of the treatment variables. Such decisions have important implications for the types of analyses that can be conducted to examine relative program effectiveness that may in turn provide evidence on targeting practices. In this section we discuss the major training activities being offered by JTPA and provide general guidelines for developing operational definitions for the specific program activity variables.

Section 204 of the Job Training Partnership Act authorizes the expenditure of Title II-A funds for over 20 types of employment and training program activities. Although the list of potential program activities is quite extensive and allows for a large combination of potential services, the major program activities provided under JTPA include classroom training (CT), on-the-job training (OJT), and job search assistance (JSA).[21] In fact, nearly 90 percent of adult FY 1984 Title II-A enrollees participated in one of these programs. Although work experience was used extensively under CETA, only 3 percent of the adults in FY 1984 were assigned to work experience programs. A brief description of each of these program services is provided below:

- <u>Classroom training</u> involves basic or remedial educational training, or occupational skills training to ensure that individuals acquire the ability and knowledge necessary to perform a specific job for which there is a demand. Such programs are usually provided in a classroom or institutional setting.

- <u>On-the-job training</u> emphasizes the development of occupational skills in an actual work setting, normally in the private sector. The programs are designed for participants who have been first hired by the employer, and the training occurs while the participant is engaged in productive work that provides knowledge or skills essential to the full and adequate performance of the job.

---

21 It should be noted that JTPA has considerably altered the mix of program activities relative to CETA. In particular, the elimination of public service employment programs, and the 15 percent cap placed on support services (with a large portion of work experience expenditures being regarded as support services), and the introduction of performance standards that emphasize immediate placement and low costs have resulted in major changes in the mix of program activities. For example, in FY 1982, 13 percent of all CETA participants enrolled in OJT programs and 30 percent enrolled in work experience programs. In FY 1984, however, 22 percent of the individuals in JTPA Title II-A programs enrolled in OJT and only 7 percent enrolled in work experience programs. In addition, job search assistance is being used much more extensively under JTPA.

- Job search assistance includes any training activity that focuses on the development or enhancement of employment-seeking skills. This service is provided to participants who need practical experience in identifying and initiating contact and interviewing with prospective employers. It is usually conducted in a structured setting and can include approaches such as job-finding clubs or instructions for self-directed job-search methods.

- Work experience is a short-term or part-time work assignment designed to enhance the employability of participants by developing good work habits and basic work skills. It is primarily intended to assist participants in entering or re-entering the labor force. Work assignments may be with a public employer or with a private non-profit agency; work experience programs are prohibited in the private-for-profit sector.

In creating specific treatment variables to represent these program activities, a couple of potential problems must be recognized. First, although the broad definitions of the major program activities are generally accepted, there is likely to be considerable variation across SDAs in the contents of specific program activities such as length of assignment, occupation, or hours per day. This makes it particularly difficult to create meaningful variables that represent a homogeneous treatment. At the same time, however, it is necessary to do a considerable amount of aggregation of activities that are generally similar but perhaps far from identical, because it is simply not possible to reliably estimate the net effects of the virtually unlimited number of program activities. Second, not only are there differences in the degree of treatment within program activity, but there are likely to be large differences across SDAs in the nature of programs provided such that work experience programs in a particular SDA may be more similar to OJT programs in another SDA. Another complication concerns the way in which the actual training activities provided are recorded in the program MIS. For example, due to the lack of uniform national reporting requirements, some SDAs record participation in a job search workshop as job search assistance, while others record it as classroom training. Such differences in the content and recording of program activities across SDAs emphasize the importance of conducting a process analysis concurrently with the net impact analysis in order to develop meaningful and consistent measures of program activities.

As indicated above, the ways in which the treatment variables are defined will in large part be determined by the structure and content of the MIS. In addition, they will depend on the specific research questions of interest and the sample sizes of individuals who participate in the given program activity. For example, to ensure that the treatment variables are as homogeneous as possible, it may be desirable to separate classroom training activities that focus on remedial education and basic skills from classroom training activities that provide specific occupational skills training. At the same time, however, if the number of individuals participating in each of these programs is too small to produce statistically reliable net impact estimates for the separate activities, it may be necessary to collapse

these two variables into one that represents classroom training programs in general.

Thus, although the specific definitions of the treatment variables will depend on several factors, we recommend that the following variables should be created for participants for potential inclusion in some models to examine impacts by program activity and other characteristics of the treatment:

- Participant dummy variable (1 if JTPA participant; 0 otherwise);
- classroom training dummy variable (1 if CT participant; 0 otherwise)
  - Classroom training remedial education dummy variable (1 if CT program in remedial education, English as a second language, or basic skills; 0 otherwise)
  - Classroom training institutional skills dummy variable (1 if CT program in specific occupational skills; 0 otherwise);
- On-the-job training dummy variable (1 if OJT participant; 0 otherwise);
- Job search assistance dummy variable (1 if participated in JSA job search or placement-related activities; 0 otherwise);
- Work experience dummy variable (1 if WE participant; 0 otherwise);
- Occupation of training dummy variables (1 if in specific 1-digit DOT code; 0 otherwise);
- Length of program participation in weeks;
- Number of hours of training per day; and,
- Completed training program activity dummy variable.

Such treatment variables would enable one to replicate all of the questions examined in previous national studies of employment and training programs, as well as several additional questions of interest.

PROGRAM ENVIRONMENTAL CONDITIONS

The final element of the conceptual framework concerns the program environmental conditions to be included in the net impact model. By program environmental conditions, we are primarily referring to characteristics of the labor markets within which the program operates, although major SDA characteristics could also be considered. As discussed earlier, because of data limitations, previous national studies have been unable to include any program environmental variables in their models. As such, little is known about how the net impact of employment and training programs varies by program environmental

conditions. At the same time, however, because of the nature of local program environmental conditions (i.e., there is no within-SDA variation on these conditions), it is important to recognize that it will only be possible to obtain reasonably precise estimates of a few key conditions, and only in states that have a large number of SDAs and that exhibit considerable variation in the conditions across SDAs. Given these limitations, below we discuss some of the more important comparisons that should be examined in a state-level JTPA net impact model.

In choosing among the many characteristics of labor markets to identify the few that should be included in the net impact model, it is important to focus attention on those factors that are most likely to affect the employment and earnings experiences of adult men and women. Although there are many factors that may affect the employment opportunities or demand for an individual with certain skills, economic theory and limited empirical evidence suggest that the most important factors are likely to be (1) the local unemployment rate and (2) whether located in an urban or rural area. The first factor has been demonstrated to be a key variable in affecting the immediate post-program outcomes of JTPA (West and Dickinson, 1985). Moreover, Johnson, Dickinson and West (1985) also provide evidence suggesting that the net impact of the ES is larger in urban areas and in areas with lower unemployment rates. Thus, we recommend that readily available labor market information be used to construct measures of the local unemployment rate for the appropriate period of analysis so that its effect can be controlled for in the model and so that one can estimate separate net impacts for programs that face different unemployment conditions.

The unemployment rate can be obtained from the Local Area Unemployment Statistics (LAUS), published by the U.S. Bureau of Labor Statistics (BLS). This information is available monthly at the state and county level and for over 1,000 cities with a population of at least 25,000. Aggregate measures of the unemployment rate corresponding to the outcome periods of interest can be calculated as weighted averages of the monthly values. In constructing these aggregate measures it is important to recognize that the unemployment rate will not generally be available for the precise area of interest. Depending on the size of the SDA, the area it serves may be either larger or smaller than the county or the city for which the information is available. In cases where the SDA serves multiple counties, one should calculate the appropriate unemployment rate variable by aggregating over the counties served by the SDA. For example, one would simply sum the number of individuals unemployed in the various counties served by the SDA and divide by the total number of individuals in the labor force in those counties during the appropriate months. In cases where the SDA serves only part of a given county, one is generally constrained to use the county value unless data for the specific cities served are available.

It may also be possible to provide some information on how the net impact of JTPA varies by different SDA service delivery strategies. The service delivery strategies to be examined should be based on their policy importance to the particular state doing the analysis. Moreover, to ensure that the strategies of interest are distinct and

quantifiable, and that there is sufficient variation among SDAs to support the analysis, it is important that a process analysis be conducted. Thus, if states with a large number of SDAs (roughly 30 or more) are interested in obtaining some information on how the net impact of JTPA varies by a key service delivery strategy, they should first ensure that there are significant differences in this strategy across SDAs. Provided it is possible to quantify these differences, one could then use the variable created to determine how the net impact of JTPA varies across SDAs that differ in this strategy using the approach to estimate subgroup effects described in Chapter 5. In states with relatively few SDAs, it is very unlikely that such an analysis would provide reasonably precise estimates of the differential effects of the strategy of interest.

CHAPTER 4
RESEARCH DESIGN

CHAPTER 4. RESEARCH DESIGN

To provide valid estimates of the net impacts of JTPA programs on the earnings and AFDC dependency of adult men and women, a research design must be developed that contains several elements. First, a sample of participants must be chosen so that the results can be generalized to the state level and the necessary data can be collected efficiently, with minimum burden to state and SDA staff. Second, a valid comparison group must be chosen so that the impact of JTPA can be distinguished from the impacts of other factors that also affect earnings and welfare dependency. Third, the sizes of the participant and comparison samples must be determined so that program impacts can be measured with precision. Finally, an overall estimation strategy must be developed, and analysis models must be specified that can provide valid estimates of the net impacts of JTPA programs on the post-program outcomes of participants. In this chapter, we discuss the first three of these research design issues for developing a state-level JTPA net impact model. The overall strategy to estimate net program impacts and the specific models to be estimated are the subject of Chapter 5.

THE PARTICIPANT SAMPLE

An important component of the research design is the development of the participant sample. The two major issues to resolve in selecting the participant sample concern: (1) the individuals to include in the sample frame and (2) the procedure to select participants from the sample frame for inclusion in the analysis. These issues are discussed below.

The Sample Frame

The sample frame from which the sample of JTPA participants will be drawn is an important determinant of the external validity of the net impact analysis, that is, the ability to generalize findings to the program as a whole. The sample frame should be representative of all JTPA participants so that the analysis results can be generalized to the state level, rather than to only particular subpopulations of participants or individual SDAs. We considered several potential sample frames, including:

- All individuals who are participating in JTPA at a given point in time;
- All JTPA terminees during a given time interval; and
- All JTPA enrollees during a given time interval.

As indicated in Chapter 2, these are the three major sample frames used in previous studies of the net impact of employment and training programs.[22] Below we discuss the advantages and disadvantages of these sample frames as they relate to the external validity of the analysis and to the ease of collecting the necessary data.

All Participants at a Given Point in Time. A sample frame that contains all JTPA participants at a given point in time has a number of disadvantages. The most important disadvantage of this approach is that it substantially undersamples individuals who participate in JTPA for only a short period and results in a non-representative sample. Such an approach also oversamples individuals who may only participate in JTPA during a particular time of the year, e.g., construction workers. If program impacts or the characteristics of participants or program activities that are related to JTPA impacts are in any way seasonal, such oversampling would lead to a seasonality bias. Because the average JTPA participant stays in the program for only approximately three to four months, this seasonality bias is potentially important.

This sample frame also creates data collection problems that make it difficult to draw an appropriate comparison group. This is because a sample of individuals who are in JTPA at a given time contains some individuals who have just entered JTPA and others who are nearing completion. Because some individuals are likely to stay in JTPA for up to six months or more, such a sample frame would result in extending the in-program period for the participant group as a whole to a year or more. This would make it very difficult to select a comparison group that is similar to JTPA participants on the timing of the pre-program decline in earnings, which we argued in Chapter 2 is very important to achieve. In addition, if the long-term stayers were retained in the sample this would delay the availability of net impact results. On the other hand, excluding such individuals from the sample would result in undersampling long-term participants, which would reduce the external validity of the analysis. Because of these numerous disadvantages, we do not recommend that the sample frame be comprised of all individuals who are participating in JTPA at a given point in time.

All Terminees during a Given Time Interval. The second potential sample frame includes all participants who terminate from JTPA during a given time interval. Unlike the previous approach, because a sample frame of terminees would include all program completers and dropouts during the time interval selected, such a sample frame is representative of all JTPA participants and does not lead to oversampling of either long-term or short-term participants. It does, however, share a possible seasonality bias with the previous approach, unless a one-year interval is used to define the sampling period. In addition, like the previous approach, because some individuals remain in JTPA for very short periods while others stay for up to six months or more, with a sample frame of terminees it would be very difficult to

---

22 A fourth potential sample frame of participants who enroll and terminate during a given time interval was rejected because it severely overrepresents short-term participants.

draw a comparison group that is similar to participants on the timing of the pre-program decline in earnings. Moreover, depending on the time period selected, and the data retention practices followed, it may be difficult in some states to obtain the necessary pre-program earnings and welfare dependency data for long-term participants.

A sample frame of terminees has some important advantages over the other two approaches that must be recognized. For example, by using a sample of participants that terminated during a particular period (e.g., the last quarter of the calendar year), it is possible to estimate post-program impacts that are measured more closely to the point of termination. For example, if the outcome period of interest is three months following the calendar quarter after program termination, this period would correspond to from three to six months after termination for all individuals if a terminee-based sample frame were used. On the other hand, using an enrollee-based sample and assuming that all individuals terminate from between one to six months after enrollment, the first three-month period that would be entirely post-program for all enrollees would vary from between three to eleven months after termination for different individuals depending on program length of stay. An additional advantage of a terminee-based sample is that in states that maintain three to five years of UI Wage Records and AFDC Grant Records, it would be possible to make just one request for the necessary pre-program and post-program outcome measures for a given group of terminees and thus minimize the burden on state data processing staff. However, in states that maintain only the minimum amount of UI Wage Records necessary to calculate benefits (i.e., four to five quarters), one would have great difficulty in implementing a terminee-based sample frame.[23]

All Enrollees during a Given Time Interval. The final potential sample frame--all JTPA enrollees during a given time interval--also has the advantage of yielding a representative sample of JTPA participants in which neither short-term nor long-term participants are oversampled. Provided the time period selected is not too long (e.g., a quarter), this approach has the major advantage of allowing one to select a comparison group that closely matches participants on the timing of the pre-program decline in earnings. As discussed above, this is particularly important for ensuring valid net impact estimates. In addition, using an enrollee-based sample maximizes the amount of pre-program earnings and AFDC data available for the model, which is particularly important for states that do not retain extensive historical data.

---

23 One possible approach would involve the following steps: first draw quarterly samples of program enrollees; then request the pre-program UI Wage Records available for those samples at enrollment; then ex post construct quarterly samples of terminees based on termination dates; and then obtain post-program outcome data for the terminee samples. It should be noted that although such a roundabout approach would work in the long term (after six months or so), the first few quarterly samples of terminees constructed would greatly overrepresent short-term participants.

On the other hand, however, this approach shares the problem of possible seasonality bias with all procedures unless enrollees during a one-year period are included in the sample frame. In addition, as indicated above, because the time required for analysis results to be available increases with the length of the interval used to define the sample frame, such an approach would delay findings more than a terminee sample frame, and result in outcome measures that are further away from the point of termination.

Sample Frame Recommendation. As the above discussion indicates, there are advantages and disadvantages to using either an enrollee-based or a terminee-based sample frame. The particular approach to use depends primarily on data availability and on whether one considers it more important to match participants and comparison group members on their experiences in the pre-program period, or whether it is more important to measure outcomes at a common point after termination. Although both approaches are viable, based on data availability considerations and the importance of ensuring that the two groups are similar on the timing of the pre-program decline in earnings, we recommend that an enrollee-based sample frame of JTPA participants be used for the state-level net impact model.

In order to precisely define the sample frame to be used in the net impact analysis, two additional decisions must be made. First, if the sample frame consists of all participants who enroll in JTPA during a given time interval, that interval must be specified. As discussed above, a long interval reduces potential seasonality bias problems whereas a short interval can result in more timely evaluation findings and also allows one to draw a comparison group that is matched more closely on the pre-program decline in earnings. Because the shortest interval for which the UI and AFDC outcome measures are both available is a three-month period, we recommend that the sample frame consist of adult men and women who enroll in JTPA during various calendar quarters.

The second decision concerns the specific calendar quarters to be used for the sampling frame. This decision affects the timing of project results, the length of the post-program observation period within a two-year program evaluation cycle, and the likelihood of seasonality bias. We recommend that the sample frame of participants be comprised of adult men and women who enroll in JTPA in each of the four quarters of a given program year. Given the delays involved in obtaining the outcome measures from agency records (approximately three months) and the time required to build the data files, conduct the analysis and prepare written reports (approximately six months), we expect that with an enrollee-based sample of participants one can obtain net impact estimates for the period one year following the calendar quarter after termination for the longest stayers only for the first quarter cohort, and that only a three-month net impact estimate can be obtained for all four quarterly cohorts in approximately a two to two-and-one-half year cycle. Of course, by obtaining additional post-program UI and AFDC records for sample members, one could estimate longer-term impacts, although the analysis period would have to be extended even further.

Sample Exclusions

Once the general sample frame is chosen, one must then determine whether certain types of individuals should be excluded. Although such exclusions reduce the representativeness of the participant sample, there may be valid reasons for excluding some participants. For example, it may be desirable to exclude cases because they lack data on critical items or because they are clearly inappropriate to include in the analysis due to policy or statistical reasons. For the most part, previous studies of the impact of employment and training programs discussed in Chapter 2 consistently imposed an age restriction on participants and also excluded individuals for whom data were missing on the outcome measure(s) or any of the key treatment variables such as program activity or length of stay. An issue of importance for which there is less consensus concerns whether individuals with limited exposure to the program should be included in the analysis. Below we briefly discuss these potential sample exclusion issues and offer a recommended approach.

Age. Most previous studies imposed some restrictions on participant age. Although there is no universal agreement on the specific age range to use, very young participants (e.g., under age 16) have been excluded because earnings is not an appropriate outcome measure for those likely to return to school, and very old participants (e.g., those age 65 and older) have been excluded because participation in employment and training programs among individuals eligible for retirement is very rare, and it is unlikely that a proper comparison group could be identified. Because the state-level net impact model is designed to focus on adults only, the participant sample will be restricted to individuals of at least age 22. Because of the difficulty of obtaining a valid comparison group for older participants, we also recommend that any individuals age 65 and older be excluded from the participant sample frame.

Missing Data. Individuals should also be excluded from the sample frame if they have missing data on key variables. For the most part, we do not expect that there will be severe missing data problems with the agency records to be used, and most problems can be dealt with in straightforward ways in the analysis.[24] The limited amount of missing data is in part a result of the procedures used by many agencies to assign "default" values when data are missing. Such procedures, however, lead to measurement error, which can also introduce analytical complications as discussed in Chapter 5.

---

24 For example, one could substitute the mean value of an independent variable for cases with missing data. Although such a procedure preserves the overall mean of the variable, it does not preserve the variance and results in biasing downward the standard errors of the estimated coefficients of the variable in a regression model. A better procedure, which preserves more of the variance of the variable, involves estimating an auxiliary regression equation to predict the variable in question using cases with complete data, and then using the predicted value for those with missing data.

A more difficult problem arises when information is missing on the treatment provided by JTPA. For example, one cannot estimate the impact by program activity or by length of stay for individuals with no designated program activity or for those who have missing data on the start and end dates of JTPA participation. We expect there will generally be very few problems concerning the omission of program start and end dates, in part because length of stay is necessary for adjusting certain performance standards for Title II-A programs. However, because there are no uniform reporting requirements for program activities, it is likely that some cases will contain missing (or unusable) program activity data. Although individuals with missing data on the JTPA treatment could be included in an impact model that only examined the average effect of JTPA participation, we recommend that they be excluded from the sample frame to preserve the statistical power of the program activity analysis, provided the reason the data are missing is unrelated to the program outcome. That is, if the reason the data items are missing is systematically related to the impact of the program, excluding such cases would reduce the internal validity of the analysis. As a result, it will be important to examine the missing data problems before making a final decision.

<u>Multiple Program Activities</u>. In some cases, individuals will be assigned to multiple program activities. Provided data are available on each of the activities that an individual participates in, multiple activities do not introduce new problems for the analysis. That is, one could include all cases with multiple activities and directly estimate the marginal effects of each program activity using the approach described in Chapter 5.[25] However, because of data limitations, persons who participated in multiple program activities caused analytical complications for some previous studies. For example, the series of recent CETA net impact studies only had available the initial program activity and whether the individual participated in multiple activities, but had no information on the number or types of other activities.

In response to this limitation, some of these studies excluded cases that participated in multiple activities. However, such an approach is not recommended because it reduces the validity of the overall analysis. Other studies included participants with multiple activities and considered them as a separate treatment group or used their initial program activity to represent the treatment. Should a state JTPA MIS contain similar limitations, information would be needed on

---

25 It must be recognized, however, that such analysis would be subject to potential selection bias as discussed in Chapter 5. That is, individuals who receive multiple program activities may be systematically different from those who only receive a single program treatment, and these differences in personal characteristics may be responsible for the observed treatment effect. For example, individuals who are less motivated could receive multiple activities because the initial activity was not effective. As a result, a small (or negative) estimated net impact for those who participated in multiple activities would in part reflect the lack of motivation and would not measure the true effects of the program activities.

the extent to which individuals receive multiple activities and the type of initial activity to make an informed judgment on which procedure to follow. For example, if a very small proportion of the sample were assigned to multiple activities, then it would not be possible to precisely estimate the impact for this group. In this case, if most of the initial program assignments involved relatively lengthy treatments such as OJT or classroom training, then the use of initial program activity for those with multiple activities would seem to be a reasonable compromise.

Limited Program Participation. The final issue concerns whether to exclude individuals who participate in JTPA for a very limited period. As described in Chapter 2, some studies chose to include all employment and training participants in the sample, while others imposed somewhat arbitrary restrictions that resulted in excluding individuals who participated in the program for only a short period (e.g., a week). Although JTPA clearly cannot have a large impact on those individuals who participate only a few days, exclusion of such individuals may introduce a selection bias into the analysis because short-term participants are likely to differ from other participants on unmeasured characteristics, such as motivation and attitude toward work. For example, if individuals leave JTPA early because they found a job, a negative bias in assessing JTPA impacts may result because those participants who would do relatively well on their own are excluded from the participant sample. Alternatively, short-term participants might consist of individuals who would do less well on their own than other JTPA participants, and hence their exclusion would result in a positive bias in the estimated impact. Both of these selection biases threaten the internal validity of the analysis. We believe it is preferable to keep the sample of JTPA participants as representative as possible and not exclude cases based on length of stay in the program. It is then possible to examine whether, and in what ways, short-term participants differ from long-term participants. This will help to determine how much confidence to attach to net impact estimates by length of program participation.

Selecting the Participant Sample

Once the participant sample frame has been determined, the next step involves the procedures to use in selecting JTPA enrollees for inclusion in the analysis. In many states, this step will be trivial, as all enrollees in a given program year will be necessary to provide reasonably precise estimates of the average effect of JTPA programs. This will also be true for medium-size states that are interested in obtaining reliable net impact estimates for various subgroups and that have the resources necessary to support the analysis. The issue of sampling primarily arises in states that serve large numbers of JTPA participants. Below we discuss alternative procedures for selecting a sample of JTPA participants from the sample frame described above to assist states that would find sampling desirable.

The principal methods that could be used for selecting a participant sample include (1) random selection, (2) stratified random selection, and (3) clustering. Because the primary advantage of a clustered

sample is to reduce the costs of data collection when relatively expensive in-person surveys are used, a clustered sample is not appropriate for the net impact model that relies on agency data. Therefore, below we discuss the relative advantages of selecting a random sample versus a stratified random sample of JTPA terminees.

Simple Random Sample. Choosing a random sample from the sampling frame has the advantage of providing a representative sample of all JTPA participants. As a result, estimates of the net impact of JTPA and of the differential impacts by program activity can be obtained in a straightforward manner; no weighting of the estimated impacts is necessary. In addition, an analysis of the types of individuals who are assigned to various JTPA program activities (i.e., targeting) is straightforward because the probability of selection into the analysis sample is not related to the assignment to specific program activities. The only potential disadvantage to simple random sampling is that it is possible that the resulting sample may be inadequate to estimate the impact for certain subgroups. For example, the impacts of program activities that occur rarely may have low power even if the total sample size is reasonably large.

Sample Stratification. Sample stratification can potentially increase the statistical power of the net impact analysis of JTPA programs. Depending on the specific research questions of interest, one could consider stratifying on the basis of participant characteristics, program activity, or even by SDA. In general, stratification is desirable only when the research questions of interest relate to subgroups that occur rarely or that occur so frequently that their nonoccurrence is rare. This is because by altering the design to increase the efficiency of the estimated effects of those subgroups that occur rarely, stratification necessarily reduces the power of estimates of the overall impact of JTPA and also complicates various analyses because the probability that an individual is included in the analysis sample is affected by the stratification. Thus, whether a state ultimately decides to stratify, and the particular strata chosen, will depend on the research questions of interest and the distribution of the character..tics of participants served and program activities provided.

Sample Selection Recommendation. We recommend that the quarterly samples of JTPA participants be selected on a random basis from the groups of adult men and women enrollees included in the sample frame. Although adult men and women are served in approximately equal numbers by JTPA overall (i.e., 52.8 percent of all Title II-A adult participants were women in PY 1984), because this varies considerably across SDAs, we believe it would be prudent to first stratify the participant sample by sex before the analysis samples are selected, or else there may be insufficient numbers of either men or women for analysis purposes.[26] As described above, because the net impact models will be estimated separately by sex, choosing random samples

[26] For example, in some SDAs women comprised as little as 25 percent of adult JTPA terminees in PY 1984, while in other SDAs women were over 80 percent of all adult terminees in Title II-A programs in PY 1984.

from the sampling frames, separately by sex, has the major advantage of providing representative samples of JTPA adult men and women participants, so that the results can be generalized to the state level and so that differential impacts by program activity can be obtained in a straightforward manner.

In addition to stratifying the sample by sex, states that want to focus their efforts on specific subgroups of adult men or women (e.g., female welfare recipients, male or female high school dropouts) may also want to consider stratifying the participant sample and oversampling the subgroups of interest. For example, because of the wide variation across states and SDAs in the use of work experience programs, states that are interested in examining the net impact of work experience programs would likely need to stratify and oversample participants of such programs. Moreover, because job search assistance generally constitutes a less intensive treatment that is likely to have a smaller average net impact, a much larger sample of participants in JSA would be needed to precisely measure the lower expected effect. Thus, for states that are very interested in precisely measuring the marginal benefits from JSA participation, such participants would have to be oversampled. States that may be interested in stratifying the participant sample and oversampling certain groups should consult a statistician or sampling expert to better understand the advantages and disadvantages of such an approach and the specific steps to be followed in drawing the sample and conducting the analysis.

COMPARISON GROUP STRATEGY

To estimate the net impact of JTPA on participants' post-program outcomes, a method is needed to gauge what would have happened to participants had they not participated in JTPA. The standard approach for determining the net impact of a program is to compare the experiences of persons influenced by the program (JTPA participants) with the experiences of persons who are not influenced by the program (the comparison group). The comparison group is used to estimate what the experiences of the participants would have been in the absence of the program. An evaluation that estimates the effects of JTPA without reference to a comparison group essentially attributes <u>all</u> gains (or losses) to the program, when other factors, such as improvement in labor market conditions, might also have contributed to this gain. This procedure generally results in greatly overstating the benefits of program participation per se. To ensure that the differences between the experiences of the participant and comparison groups can be attributed to the program, the comparison group must have characteristics similar to JTPA participants, and the data available must be comparably measured for the two groups. In this section, we describe the advantages and disadvantages of alternative comparison group strategies and present a recommended approach.

As described in Chapter 2, a few recent evaluations have randomly assigned applicants to either a treatment group that could participate in the program or to a control group that could not. Such an experimental design can potentially allow a valid determination of the

program's net impact because the only systematic difference between the treatment and control groups is the opportunity to participate in the program. However, because JTPA is an ongoing program, serious ethical and legal concerns preclude the use of a randomly assigned control group, particularly for an evaluation that would be conducted on a biennial basis. Moreover, to the extent that crossover occurs (i.e., control group members reapply and join the treatment group) or that substantial attrition occurs, this would threaten the internal validity of the experimental design. Thus, alternative methods that can approximate a true control group must be explored.

One method that has been widely applied in other contexts is to use each individual as his or her own control by comparing the individual's pre-program experience with his or her post-program experience. Although such comparisons can provide useful information on gross program impacts, this methodology is not appropriate for examining the net impact of JTPA for two reasons. First, the key outcomes--earnings and AFDC dependency--may change systematically over time, for example, as the result of changes in the national economy or changes in state welfare eligibility rules. Pre-program and post-program comparisons will confuse the impact of such changes with the impact of JTPA. Second, it is very likely that individuals' earnings in the period before enrollment in JTPA are atypically low and would increase over time even in the absence of participation in the program. Such pre-program earnings declines of participants have been extensively documented and reflect legislation instructing program operators to serve persons who have recently faced difficulties in the labor market and are most likely to benefit from employment and training assistance. As a result, with a pre-post design it is difficult to determine whether the observed post-program earnings gains should be attributed to training or are merely an artifact of the way in which individuals are selected into the program.

Because a randomly assigned control group is not feasible and a before-and-after comparison is not appropriate, a comparison group must be drawn from a population of program-eligible individuals who do not participate in JTPA. The major disadvantage of using nonparticipants as a comparison group is the potential for selection bias. Specifically, individuals who choose to participate in JTPA may be systematically different from those who are also eligible, but do not choose to participate. To the extent that these differences can be measured, they can be controlled for in the analysis. However, there may also be unmeasured differences, such as motivation or attitude toward work, that affect both whether individuals participate in JTPA as well as their subsequent employment experiences. A comparison of the outcomes of participants with the outcomes of nonparticipants risks attributing the effects of these pre-program differences to JTPA. This indicates the importance of selecting a comparison group that is as similar as possible on measured and unmeasured characteristics to the sample of JTPA participants. In the remainder of this section we elaborate on the criteria that should be used in choosing a comparison group of nonparticipants and discuss the advantages and disadvantages of various alternative comparison groups.

Criteria for Choosing a Comparison Group of Program-Eligible Nonparticipants

The two major considerations involved in choosing a comparison group are the size of the group and its similarity to the participant sample. A large comparison group is desirable because it enables more efficient or precise estimates of net impacts. Consequently comparison groups that can be obtained less expensively are more desirable than comparison groups that require expensive data collection because a larger sample may be available within a fixed budget.

The comparison group should also be similar to the participant group on both observed and unobserved characteristics. Similarity on observed characteristics increases the efficiency of the estimation of net impacts because the correlation between individual characteristics and JTPA participation is small. In addition, the estimates of net impacts are likely to be less affected by unmodeled nonlinearities and interactions. That is, when the distribution of observed variables is similar in the two groups, specification errors affect both the comparison and the participant samples in the same manner, so that the estimated difference in their behavior should not be affected.

It is especially important for the comparison group to be similar to the participant sample in terms of characteristics that are related to eligibility for JTPA. According to the JTPA legislation, to be eligible for Title II-A programs, adults must be 22 years of age or older and be economically disadvantaged. The legislation defines economically disadvantaged to mean an individual who: (1) is a member of a family that in the six months prior to application received a total income of less than the OMB poverty level or less than 70% of the lower living standard income level, whichever is greater, given the person's family size; (2) is a member of a family that receives federal, state, or local cash welfare payments; (3) is receiving food stamps; (4) is a foster child for whom state or local support payments are made; or (5) is a handicapped individual who is economically disadvantaged but whose family is not, as permitted by the Secretary of Labor. The Act requires that at least 90% of Title II-A participants be economically disadvantaged and allows up to 10% of the participants to be individuals who are not economically disadvantaged provided they have encountered barriers to employment (e.g., limited English language proficiency, school dropouts, ex-offenders).[27] In addition, the Act requires that recipients of AFDC grants and school dropouts be equitably served in relation to their incidence in the eligible population. To the extent possible, the comparison group should only include individuals who meet the explicit eligibility criteria and who are similar to participants on characteristics emphasized in the legislation.

---

27 It appears, however, that relatively little use is being made of the 10% "window" for serving persons who are not economically disadvantaged. For example, Cook et al. (1985) report that 94% of all persons served by Title II-A programs during the transition year were economically disadvantaged.

It is also desirable for the participant and comparison groups to be similar in unmeasured characteristics so that selection bias can be minimized. However, such similarity is extremely difficult to achieve. One approach is to measure some aspects of normally unobserved characteristics, for example, through questionnaire items that measure motivation or work ethic. Because the net impact study will rely exclusively on agency data, however, this approach is not feasible. Other approaches involve either choosing comparison group members who have gone through selection processes similar to those used to screen JTPA applicants, or utilizing complex statistical models in an attempt to correct for potential differences in unmeasured characteristics. Alternative approaches to deal with selection bias are discussed in Chapter 5.

Alternative Comparison Groups of Nonparticipants

In this section we discuss alternative comparison groups of nonparticipants that can be drawn from existing data bases. As described above, the primary advantage of drawing a comparison group from an existing data base is that it is relatively inexpensive and allows a greater sample size for a given budget, thus resulting in greater statistical power. The major disadvantage of using an existing data base is that the analysis must necessarily be limited to variables contained in both the JTPA MIS and in the data base for the comparison group. To be useful in the JTPA net impact model, the information in the data base for the comparison group should pertain to the same general time period as the data collected for participants, and the common baseline variables should be measured comparably. In addition, it is important that comparison group members have not received employment and training program assistance in the past or that variables measuring the extent of treatment be available.

We considered several alternative data bases for use in developing a comparison group including:

- Various national surveys (e.g., the Current Population Survey, the Survey of Income and Program Participation, the DOL National Longitudinal Survey);
- State files of UI claimants;
- JTPA applicants who are "no-shows"; and
- State files of Employment Service (ES) registrants.

Below we discuss the advantages and disadvantages of each of these alternative comparison group strategies.

<u>Various National Surveys</u>. Although most previous studies of the net impact of employment and training programs used comparison groups drawn from various national data bases, such an approach is simply not feasible for a state-level net impact model. This is primarily because these data sources do not have local area identifiers, which are necessary to ensure that the comparison group is drawn from the same

state and local area as the participant sample. Moreover, even if this problem could be overcome, these data sets are deficient in that they do not have sufficient sample sizes at the state and local level. Finally, the data sets are generally not available in a timely fashion.

UI Claimants. An advantage of developing a comparison group from the file of UI claimants is that it is possible to ensure that the comparison group members are from the same state and local areas as participants. In addition, there are a few demographic characteristics such as age, race, sex, and occupation on the data file that could be used as control variables in the model. A sample of UI claimants, however, has two important drawbacks. First, claimants are very different from JTPA participants in terms of their work histories. That only approximately 10% of adult JTPA participants are UI claimants supports this view. Second, because claimants are generally eligible to receive UI benefits for some time longer than the average length of treatment in JTPA, such a comparison group would artificially inflate the net impact estimate on earnings.[28]

JTPA Program No-Shows. JTPA applicants who are eligible and assigned to a program activity but who do not participate--"no-shows"--have certain potential advantages as a source for drawing a comparison group. The primary advantage of using these individuals as a comparison group is that they underwent many of the selection processes that participants did. Thus, they were eligible, decided to apply for JTPA, were selected to participate, and were also assigned to a program activity. In addition, a major potential advantage is that because these individuals applied to JTPA, comparable baseline data were collected for them, which maximizes the amount of pre-program information that could be used to construct control variables for the net impact model. The major disadvantages are that (1) these groups introduce an additional selection bias in that they may be systematically different because they chose not to participate (perhaps because they found a job); (2) the samples sizes of such groups may be too small to support precise net impact estimates; and (3) data on no-shows are not usually included in the JTPA MIS. Thus, although this approach has certain advantages, it raises additional questions and cannot be implemented unless states and SDAs consistently include no-shows in the data base.

ES Registrants. Drawing a comparison group of ES registrants in offices in the areas served by the SDAs has several advantages. First, data are available on several individual characteristics of interest, which should generally be comparably measured in the JTPA MIS and, as such, can be included as control variables. Except for some potential exceptions noted below, this usually includes the variables used to define eligibility for JTPA. Second, like JTPA participants, ES registrants also presumably experienced a recent decline in earnings.

---

28 We also rejected the Continuous Wage and Benefit History (CWBH)--a special-use subfile of UI claimants available in certain states that has detailed work history information--because it is available only for claimants and only in a very limited number of states.

Finally, also like JTPA participants, ES registrants are in the labor force at the time they applied for assistance. As discussed in Chapter 2, it is important to ensure that participants and comparison group members are similar in their attachment to the labor force, or else net impact estimates can be severely biased.

Although ES registrants should be more similar to JTPA participants than individuals in other data sources, the use of ES registrants as a source for drawing a comparison group has certain disadvantages. For example, to the extent that ES registrants receive other employment and training services, the comparison group would look less like an "untreated" group. If ES registrants receive significant employment assistance, a comparison of the average post-program outcomes of JTPA participants and ES registrants would measure the incremental effect of JTPA relative to the ES, and not the effect of JTPA per se. Although this is not an uninteresting question to address, particularly at a time of limited resources for employment and training program assistance, there are other procedures that can be used to deal with this problem and estimate the net impact of JTPA. For example, one could obtain data on ES services received and control for the effect of those services in the net impact model.[29] Alternatively, one could restrict the sample of ES registrants to persons who did not receive ES services. Because a substantial proportion of ES registrants generally receive no services other than registration, there are likely to be a sufficient number of ES registrants that receive no services and that are similar to JTPA participants on measured characteristics that could be used to develop a comparison group.[30]

A second potential problem in using ES registrants to develop a comparison group concerns the possibility that the ES registrant file may be dominated by UI claimants in some areas. For example, in states and local areas in which ES offices are co-located with UI offices, this could result in a large proportion of UI claimants in the ES registrant file. As we argued above, it would be inappropriate to compare the outcomes of JTPA participants with a sample that is dominated by UI claimants. Thus, in instances where ES offices are co-located with UI offices, it may be necessary to undersample UI claimants in the ES registrant file to obtain a comparison sample that has a proportion of claimants similar to the JTPA population.

---

29 It must also be recognized that the ES is a potential source of job placement assistance for some JTPA terminees. As such, it could be important to control for ES services received by participants as well as by comparison group members.

30 Because the net impact model relies on a regression approach with dummy variables representing the treatments, the estimated impact ultimately involves a comparison of the earnings of participants and of ES registrants who receive no ES services. As such, the estimated impact should be similar whether one controls for ES services in the model or instead excludes individuals who receive ES services. It is important to note, however, that the power of the analysis is considerably greater for a given sample size if those who received ES services are first excluded from the analysis.

A third potential disadvantage to using ES registrants concerns recent reductions in federal reporting requirements related to the ES that will result in less consistent information across states. In particular, states are no longer required to report the number of economically disadvantaged applicants who are registered and served by the ES. Because economically disadvantaged status is the major criterion for JTPA eligibility, and given the importance of ensuring that the comparison group be similar to participants on all characteristics that affect eligibility, it is important that the economically disadvantaged status variable be available for the net impact model. It appears, however, that many states have decided to continue collecting information on the economically disadvantaged status of ES registrants in the event that reporting requirements change again. For states that no longer collect this information, in order to implement the net impact model they will either have to change their data collection practices or implement more complex procedures to draw a comparison group that is matched to participants on the level of earnings and recent changes in pre-program earnings, similar to the approach used by Westat (1982) and others to evaluate the net impact of CETA programs.

A final potential disadvantage to using ES registrants as a source for drawing comparison groups concerns differences across states in their procedures for retaining historical data. In the past, most states have kept copies of data tapes with individual ES records (including registrant characteristics and ES services received) for a period of three to five years. In some states, however, individual-level data are purged after one year. In such states it would not be possible to conduct the analysis retrospectively, since data on comparison group members for individuals who enrolled in JTPA early in the year may already have been purged. Thus, such states would have to draw the comparison samples on an ongoing basis or alter their practices concerning the retention of historical data.

Despite these potential disadvantages, we believe that ES registrants are the best source for drawing a comparison group to examine the net impact of JTPA. As a result, in the remainder of this section we discuss the variables that are comparably measured in the ES and JTPA MIS data and that can be included as control variables in the model. In subsequent sections, we discuss additional details related to drawing a sample of ES registrants, including the time frame for selection, cases that should be excluded, and procedures for drawing the sample.

In Appendix A, we provide copies of the application forms that are currently being used in the State of Washington for ES registration and for JTPA application. Based on a review of these forms, the following data items appear to be comparably measured and as such can reliably be included in the net impact model:

- Age (in years);
- Race (White, Black, Hispanic, American Indian/Alaskan Native, Asian/Pacific Islander);
- Sex;

- Education (whether received high school degree or equivalent);
- Handicapped status (whether has physical or mental impairment that is a substantial handicap to employment);
- Occupation (primary DOT code of previous job);
- Veteran status (whether a veteran, whether a Vietnam-era veteran, and whether a disabled veteran);
- Food Stamps recipient;
- WIN registrant; and
- Economically disadvantaged status.

In addition to these items that should be available and comparably measured for participants and comparison group members in the JTPA and ES MIS systems, other state agency data bases will provide key control variables for the net impact model. Specifically, from various UI data bases and from the PA data base one will obtain comparable measures of:

- Pre-program quarterly earnings;
- Pre-program UI benefit payments received; and
- Pre-program AFDC grants received.

Although these two lists provide considerable information on individual characteristics in the pre-program period that can be used in the net impact model, the information is not as complete as one would ideally like. In particular, it would be desirable to have measures of marital status, family size, dependent children, ex-offender status, limited English-speaking ability, and detailed data on pre-program employment experiences. It must be recognized, however, that most of these characteristics were also unavailable to previous national studies of the impact of government subsidized employment and training programs. As such, this is not a limitation that is specific to the state-level model.

Comparison Group Sample Exclusions

Prior to selecting the comparison group of ES registrants, certain cases should be excluded from the sample frame to maintain comparability with the participant sample. In addition, it is necessary to exclude from the comparison group individuals who are clearly not eligible for JTPA. It may also be desirable to exclude those who are likely to have earnings functions that are considerably different from the earnings functions of JTPA participants. Below we discuss sample exclusion considerations concerning the comparison group of ES registrants.

<u>Age and Missing Data</u>. To maintain comparability with the JTPA

participant sample, the group of ES registrants should be restricted to individuals who are at least age 22 and who are under 65 years of age. The ES registrant sample should also be restricted to cases that do not have missing data on key independent variables.

<u>JTPA Participation</u>. It is important to exclude ES registrants from the sample who are JTPA participants during either the pre-program, inprogram, or post-program period. This problem, known as "comparison group contamination," results in comparing the outcomes of JTPA participants with the outcomes of comparison group members who also participated in JTPA and yields net impact estimates that are biased toward zero. To minimize this problem, one should compare the social security numbers of current and recent JTPA participants with the social security numbers of ES registrants and exclude all matches from the comparison sample.

<u>Labor Force Attachment</u>. As discussed in Chapter 2, most recent studies have generally imposed a labor force attachment criterion on comparison sample members. Because ES registrants are in the labor force at the time they apply to the Job Service, however, it is possible to align the comparison group so that the sample of ES registrants is selected at approximately the same time as JTPA participants are enrolling in JTPA. Specifically, the comparison group should be drawn from new ES registrants in the same calendar quarter that participants enroll in JTPA. Following this procedure, other registrants who applied for ES assistance more than three months earlier and, as such, experienced their decline in earnings even earlier and who may have already returned to work, will be excluded from the sample frame for the comparison group.

<u>ES Services</u>. Another issue concerns potential exclusions to ensure that the comparison group is as untreated as possible. In particular, in addition to excluding from the sample frame potential comparison group members who have participated in or are participating in JTPA, it is also important to exclude individuals who receive significant ES services. The major ES services include referral to jobs listed with the ES, job development, job counseling, and testing. Because individuals who receive such services may be receiving considerable employment assistance, we recommend that such individuals be excluded from the comparison group sample frame. Since individuals may receive some ES services several months after they apply for assistance, we recommend that a period six months after application be used for determining whether a person received significant ES assistance. Because a large proportion of ES registrants generally receive no services other than registration during a six-month, post application period (Johnson, Dickinson, and West, 1985), there are likely to be a sufficient number of ES registrants who receive no services and who are similar to JTPA participants on measured characteristics to develop a comparison group.

<u>JTPA Eligibility</u>. A final important sample exclusion issue concerns procedures to ensure the similarity of the participants and comparison group members on characteristics related to eligibility for JTPA. As indicated above, the primary criterion for JTPA eligibility is that the person be economically disadvantaged. Although the Act requires that

at least 90% of JTPA participants be economically disadvantaged, in PY 1984, 95% of adults in Title II-A programs were economically disadvantaged. Moreover, of those who are not economically disadvantaged, or who cannot be certified to be economically disadvantaged, the Act requires that they face demonstrated employment barriers. Because virtually all adult Title II-A enrollees are economically disadvantaged, and even those who are not certified to be economically disadvantaged may in fact be so, or may have earnings potential that is most similar to economically disadvantaged individuals, we recommend that the comparison group also exclude all new ES registrants who are not economically disadvantaged at application. This will ensure a comparison group that is similar to JTPA participants on the key characteristic related to JTPA eligibility.[31]

Selecting the Comparison Group Sample

Once decisions have been made on which individuals should be excluded from the comparison sample frame, the next step involves procedures to use for drawing the comparison group. We considered the following three alternative approaches for selecting the comparison group:

- Including all ES registrants in the sample frame;
- Developing a matched comparison group; and
- Using a subsample of ES registrants that has the same distribution as JTPA participants on a few key characteristics related to JTPA eligibility.

As we describe below, each of these approaches has advantages and disadvantages in terms of the resources required to conduct the net impact analysis and the validity of the results. These tradeoffs must be kept in mind in identifying the most appropriate procedure for a state-level net impact evaluation.

<u>Including All ES Registrants</u>. The first alternative has the advantage of maximizing the size of the comparison group and thus the statistical power of the analysis. It has a major disadvantage, however, of generating a comparison group with demographic characteristics and employment histories that are very different from those of the

---

31 As described in Chapter 2, several previous studies excluded from the comparison group individuals with very high pre-program earnings who were clearly ineligible to participate in employment and training programs. By matching participants and comparison group members on economically disadvantaged status, however, such additional exclusions are no longer necessary. It should be noted that if a comparable measure of economically disadvantaged status is not available in some states for ES registrants, then procedures to exclude cases with high pre-program earnings (e.g., one standard deviation above the maximum earnings of participants) would have to be implemented.

participant sample. As a result, the correlation between JTPA participation and individual characteristics is likely to be relatively high, and the anticipated gain in statistical power may not be realized. In addition, as discussed earlier, using the entire comparison sample could result in biased net impact results because of unspecified nonlinearities and interactions that arise from estimating the regression model over a wide range of incomes. As such, we do not recommend this approach.

<u>Matched Comparison Group</u>. An alternative approach involves choosing a comparison group from the sample frame that is specifically matched to the participant sample on the basis of characteristics that are related to program participation and the post-program outcomes of interest. This procedure has the important advantage of reducing the pre-program differences between the participant and comparison samples so that regression estimates are less sensitive to unspecified nonlinearities and interactions. However, the sample size would be greatly reduced, which would result in lower statistical power. Moreover, the principal matching approaches that have been used in developing comparison groups for evaluating the impact of employment and training programs--stratified matching and nearest-neighbor matching--are complex procedures to implement and require considerable staff and computer resources. In addition, unless a one-to-one match is developed, the impact analysis would have to be weighted, which would be an additional complication. Because of the excessive resources involved in conducting the analysis with a comparison group that is specifically matched to the participant sample, we believe it is unrealistic to develop a matched comparison group for the state-level net impact model.

<u>Matching Distributions on Key Variables</u>. Our recommended approach is a compromise between the two alternatives discussed above. Specifically, we recommend that comparison group members be selected from the sample frame to ensure that the distribution of the resulting comparison group is similar to the distribution of participants on key characteristics. This approach maintains maximum statistical power to the extent possible, while ensuring that the participant and comparison samples are similar on key characteristics. As indicated above, because of program eligibility considerations and certain practical issues concerning the relationship between the ES and other programs such as UI and welfare, some of the more important characteristics on which to ensure comparability between participants and comparison group members are economically disadvantaged status, receipt of UI, and receipt of AFDC. Because we will ensure comparability between the two groups on economically disadvantaged status by excluding from the sample frame for the comparison group all new ES registrants who are not disadvantaged, no additional matching is required on this characteristic. Although we do not have any information on the proportion of new ES registrants that are AFDC recipients or that are UI recipients nationwide, there are likely to be relatively too few ES registrants who are AFDC recipients (e.g., 9 percent of adult men and 35 percent of adult women JTPA terminees in PY 1984 were receiving AFDC at application) and relatively too many ES registrants who are receiving UI benefits (e.g., 15 percent of adult men and 8 percent of adult women JTPA terminees in PY 1984 were UI claimants at

application). In order to ensure similarity on these important characteristic we recommend that comparison group members be randomly selected from the sample frames of adult men and women to match the distribution of participants on these characteristics. Thus, for the separate samples of adult men and women, procedures would be used to ensure that the participant and comparison groups are similar on the proportions in the four cells representing combinations of AFDC and UI recipiency status. Operationally, for a given total sample size of participants and comparison group members, sampling rates for each cell would be determined to match the two distributions and then comparison group members would be selected randomly from the cells at the given sampling rates. Alternatively, the maximum combined sample size could be obtained by including all comparison group members in the cell with the smallest ratio of ES to JTPA cases and then adjusting the sampling rate for the other three cells to obtain the same proportion of comparison group members as participants in each cell.

SAMPLE SIZES FOR THE PARTICIPANT AND COMPARISON SAMPLES

An important element of the research design is the determination of the appropriate sample sizes for the participant and comparison groups. As we indicated earlier, in many states there will be little choice involving the size of the participant sample. That is, it may be necessary to use all JTPA participants to obtain reasonably precise estimates of average program net impacts. In such cases, the only issue involves the size of the program impacts that can be detected at a given statistical power for different sized comparison group samples. Provided sufficient staff and computer resources, one should generally use the largest sized comparison group feasible. Because the marginal cost of increasing sample size is very low since data are available in existing agency files, even in medium to large states one should generally use the largest numbers of participants and comparison group members feasible, given available staff and computer resources. In states with large JTPA programs, however, samples of participants and comparison group members will be drawn, which raises the issue of total sample size as well as the allocation of the total sample among the two groups. In the remainder of this section we describe an approach to guide states in selecting appropriate sample sizes for the net impact model.

In order to make an informed decision concerning both the total sample size and the allocation of the total sample between the participant and comparison groups, information on several factors is required. For example, the appropriate sample size for the net impact model

---

32 For example, based on State of Washington data for the first nine months of PY 1985 (July 1, 1984 - March 31, 1985), of the 360,000 ES registrants active during this period, only 7% were welfare recipients and nearly 50% were UI claimants. On the other hand, nearly 30% of Title II-A terminees in Washington during this period were welfare recipients and less than 15% were UI claimants. For the most part, ES registrants and JTPA participants were reasonably similar on other demographic characteristics such as age, race, and education.

ultimately depends on the size of the impact that is important to detect for policy purposes, and the level of statistical accuracy required. With larger sample sizes, one has greater assurance of detecting small differences in overall outcomes between the participant and comparison groups, as well as detecting differences for major participant subgroups or across program activities. The likelihood of detecting a given difference in outcomes also depends on the allocation of the total sample between the participant and comparison groups and the unexplained variance of the outcome measure (i.e., earnings, AFDC grants). In general, the more homogeneous the samples (i.e., the smaller the variance of the outcome measure), the smaller will be the number of cases necessary to detect a given difference in outcomes at a specific level of significance. Below we illustrate how some of these factors can affect the sample size required for testing various hypotheses concerning the net impact of JTPA on earnings.[33]

In Tables 3 through 6 we provide information on the likely precision of estimates of annual earnings impacts for different sized samples of JTPA participants and comparison group members, separately for adult men and women. These estimates of the precision of annual earnings impacts are based on simple difference-in-means tests and assume that the unexplained variance of earnings is the same for participants and comparison group members.[34] The final two columns of these tables indicate the likely observed increase in earnings due to JTPA that would be statistically significantly different from zero at least 90% of the time. That is, one would have 90% power at a .10 significance level of detecting an overall increase in JTPA participants' annual

---

33 Although the sample size required for examining the net impact of JTPA on welfare dependency is likely to be different, information on the unexplained variance of AFDC grants was not available to estimate sample sizes with any degree of confidence. We expect, however, that sample sizes that are adequate for examining earnings impacts will also be adequate for examining impacts on AFDC grants.

34 Although the analysis models described in Chapter 5 rely on multiple regression techniques in which the hypothesis that JTPA increases earnings is examined by performing a "t"-test on the JTPA participation dummy variable, because the comparison group of ES registrants will be similar to JTPA participants on key characteristics that affect eligibility, the t-test is essentially a difference in means test. As such, standard procedures can be used to calculate the standard deviation of a difference in means test. The standard deviation of a simple difference in means test is given by
$S = (S_p^2/N_p + S_c^2/N_c)^{1/2}$, where S_p^2 and S_c^2 are the variances of earnings for the participant and comparison groups respectively, and N_p and N_c denote the sample sizes for the two groups, which are shown in Tables 3 through 6 in the columns headed "Relevant Sample Sizes." Based on an analysis of earnings of CETA participants, we estimate the unexplained variance of earnings for participants and comparison group members would be approximately $(\$6,000)^2$ for adult men and $(\$4,700)^2$ for adult women.

Table 3

PRECISION OF ESTIMATES OF ANNUAL EARNINGS IMPACTS FOR AVERAGE STATE WITH EQUAL SIZED COMPARISON GROUPS

COMPARISONS	Relevant Sample Sizes		Standard Deviation of Mean Difference: Men	Standard Deviation of Mean Difference: Women	Observed Impact Required for 90% Power at .10 Significance Level: Men	Observed Impact Required for 90% Power at .10 Significance Level: Women
I. JTPA vs. Comparison Group[a]						
. Overall Impact	3,000	3,000	$155	$121	$397	$310
. CT Impact	1,200	3,000	205	161	526	413
. OJT Impact	750	3,000	245	192	628	492
. JSA Impact	600	3,000	268	210	687	538
II. Within-program impacts[b]						
. CT compared to other activities	1,200	1,800	224	175	574	513
. High school graduates compared to nongraduates	2,250	750	253	198	741	580
. Blacks compared to other race/ethnic groups	900	2,100	239	187	700	547
. Female welfare recipients compared to nonrecipients	1,200	1,800	---	175	---	513

Notes: a - A one-tail significance test is assumed.
b - A two-tail significance test is assumed.

Table 4

PRECISION OF ESTIMATES OF ANNUAL EARNINGS IMPACTS FOR AVERAGE STATE WITH EXPANDED COMPARISON GROUPS

COMPARISONS	Relevant Sample Sizes		Standard Deviation of Mean Difference		Observed Impact Required for 90% Power at .10 Significance Level	
			Men	Women	Men	Women
TPA vs. Comparison Group[a]						
Overall Impact	3,000	9,000	$126	$ 99	$323	$254
CT Impact	1,200	9,000	184	144	472	369
OJT Impact	750	9,000	228	179	585	459
JSA Impact	600	9,000	253	198	649	508
ithin-program impacts[b]						
CT compared to other activities	1,200	1,800	224	175	574	513
High school graduates compared to nongraduates	2,250	750	253	198	741	580
Blacks compared to other race/ethnic groups	900	2,100	239	187	700	547
Female welfare recipients compared to nonrecipients	1,200	1,800	---	175	---	513

a - A one-tail significance test is assumed.
b - A two-tail significance test is assumed.

Table 5

PRECISION OF ESTIMATES OF ANNUAL EARNINGS
IMPACTS FOR LARGE STATE

COMPARISONS	Relevant Sample Sizes		Standard Deviation of Mean Difference		Observed Impact Required for 90% Power at .10 Significance Level	
			Men	Women	Men	Women
JTPA vs. Comparison Group[a]						
. Overall Impact	9,000	9,000	$ 89	$ 70	$228	$179
. CT Impact	3,600	9,000	118	93	302	238
. OJT Impact	2,250	9,000	141	111	362	285
. JSA Impact	1,800	9,000	155	121	397	310
Within-program impacts[b]						
. CT compared to other activities	3,600	5,400	129	101	378	296
. High school graduates compared to nongraduates	6,750	2,250	146	114	427	334
. Blacks compared to other race/ethnic groups	2,700	6,300	138	108	404	316
. Female welfare recipients compared to nonrecipients	3,600	5,400	---	101	---	296

es: a - A one-tail significance test is assumed.
b - A two-tail significance test is assumed.

Table 6

PRECISION OF ESTIMATES OF ANNUAL EARNINGS
IMPACTS FOR SMALL STATE

COMPARISONS	Relevant Sample Sizes		Standard Deviation of Mean Difference		Observed Impact Required for 90% Power at .10 Significance Level	
			Men	Women	Men	Women
JTPA vs. Comparison Group[a]						
. Overall Impact	1,000	1,000	$267	$210	$ 685	$ 538
. CT Impact	400	1,000	354	279	908	715
. OJT Impact	250	1,000	423	333	1,085	[illegible]
. JSA Impact	200	1,000	465	363	1,192	[illegible]
Within-program impacts[b]						
. CT compared to other activities	400	600	387	303	1,134	888
. High school graduates compared to nongraduates	750	250	438	342	1,281	1,002
. Blacks compared to other race/ethnic groups	300	700	414	324	1,212	948
. Female welfare recipients compared to nonrecipients	400	600	---	303	-----	988

a - A one-tail significance test is assumed.
b - A two-tail significance test is assumed.

earnings as small as the values given in these two columns.[35] Before discussing the information provided in these tables, it must be emphasized that these estimates depend on the unexplained variance of earnings for participants and comparison group members and, if the samples ultimately selected are more homogeneous, one would be able to detect smaller increases in earnings for a given total sample size at the same power.

In Table 3 we present estimates of the precision of the annual earnings gains that would be obtained from a state with an average size JTPA program and with equal numbers of participants and comparison group members. In particular, the illustration in Table 3 assumes 6,000 adult JTPA terminees equally divided between men and women. As this table indicates, with 3,000 adult male JTPA participants and 3,000 adult male ES registrants serving as the comparison group, the standard deviation of the mean difference in earnings between the two groups is estimated to be $155 (see formula described in footnote 34). This indicates that an average gain of approximately $400 in annual earnings due to JTPA would be statistically significant at least 90% of the time. For adult women, we estimate that somewhat smaller average JTPA impacts (approximately $300) can be detected with sample sizes of 3,000 participants and 3,000 comparison group members. This is because of the smaller expected unexplained variance in earnings for adult women. Based on estimates of the average earnings for adult men and women, these results indicate that a five (six) percentage point effect on earnings can be detected for men (women) with 90% power at a .10 significance level.[36]

The remainder of Table 3 illustrates how the precision of net impact estimates differs for various subgroups.[37] As expected, the minimum sizes of within-program impacts that can be detected with the same degree of statistical accuracy are considerably larger. For example, it is only possible to detect an average increase in earnings due to JSA participation of approximately $700 ($540) for men (women) at least 90% of the time. With this sample design it is also possible to detect similar impacts for blacks and high school graduates.

---

35 The statistical power is the probability of detecting an effect at the chosen significance level when an effect of the specified size in fact exists. Formally, it is equal to 1 minus the probability of making a Type II error. The use of 90% power requires that the hypothesis test be correct 90% of the time.

36 The estimates of average annual earnings for adult men of $8,600 and for adult women of $5,200 are derived from data for 1976 CETA enrollees and rescaled to reflect inflation since that time.

37 The relevant sample size estimates reported in these tables assume a random sample of JTPA terminees. Specifically, based on data for JTPA terminees in the nine-month transition period, we assume that 40% of adult terminees are assigned to classroom training, 25% to OJT and 20% to job search assistance. In addition, high school graduates are assumed to comprise 75% of terminees, blacks 30%, and 40% of the female terminees are assumed to be welfare recipients.

In Table 4 we demonstrate the sensitivity of the precision of net impact estimates to expanding the size of the comparison group. In particular, we maintain a participant sample size of 3,000 each for adult men and women and increase the size of the comparison group to 9,000 each. As Table 4 indicates, expanding the comparison group considerably does not appreciably reduce the size of earnings impacts that can be detected with 90% power at a .10 significance level. For example, the overall annual impact that can be detected for men declines from $397 to $323, and for women from $310 to $254. The value of the additional comparison group members is even less for estimating separate subgroup impacts.

In Tables 5 and 6 we describe tne precision of estimated earnings impacts for large and small state JTPA programs respectively. As Table 5 indicates, by increasing the participant sample to 9,000 and maintaining a comparison group sample of 9,000, it is possible to detect small overall effects and moderate sized effects for several subgroups. For example, with such a sample design, we estimate that one could detect an annual earnings impact of as small as $228 for men and $179 for women, which correspond to approximately 2.5 to 3 percentage point impacts. Moreover, such a sample provides the same precision for estimating the separate impact of job search assistance programs as a sample size of 3,000 comparison group members and 3,000 participants provides in examining the overall impact of JTPA.

The results in Table 6 reveal the difficulty small states will have in obtaining precise estimates of the impact of JTPA programs. Because several states are likely to have fewer than 1,000 adult male enrollees and 1,000 adult female enrollees in Title II-A programs in a given program year, should such states undertake a net impact analysis the results must be carefully interpreted. That is, as Table 6 indicates, because of the small sample sizes available, only very large net impacts can be reliably detected--on the order of $1,000 or more--impacts that are considerably larger than those generally reported in previous studies. As such, these states are very likely to find net impacts that are not statistically significantly different from zero at conventional levels of significance. In such circumstances, states should refrain from concluding that JTPA has no impact and focus on the signs and magnitudes of the estimated program impacts.

The above discussion has illustrated how the minimum program impact that can be detected with 90% power at a .10 significance level varies by sample size and the sample allocation between participants and comparison group members. However, the question remains as to the appropriate sample size and allocation for conducting a state-level JTPA net impact analysis. Provided one retains the criteria of 90% power and a .10 significance level, for a given unexplained variance in the outcome measure the choice of sample size ultimately rests on the minimum detectable effect that is considered to be acceptable. For the most part, however, this is a matter of judgment, and little guidance can be provided.

One possible approach, recently offered by Stafford (1985), involves selecting a sample size such that the minimum detectable post-program

effect covers the average cost of JTPA programs. Stafford suggested a post-program time horizon of four years as reasonable in order to capture the longer-term earnings gains of participants compared to nonparticipants.[38] We implemented this general approach using a discount rate of 5% and assuming JTPA costs are similar to the experiences during the nine-month transition period, in which the average cost per terminee was approximately $2,300. Given equal sizes of the participant and comparison groups for men and women, in order to yield a combined average discounted effect of $2,300, this suggests that the impact for men that can be estimated at 90% power is approximately $700 per year, and the impact for women is approximately $550. As the results in Table 6 indicate, a sample size of 1,000 comparison group members and 1,000 participants each for adult men and women would meet this criterion for the overall impact of JTPA, but none of the subgroup impacts would be estimated with any degree of precision. On the other hand, based on the results in Table 3, with 3,000 persons in each cell one could ensure that the minimum detectable effect covered average program costs for all major subgroups. It should also be noted that samples of approximately 6,000 adult men and 6,000 adult women, divided equally between participants and nonparticipants, were used in several recent studies to examine the net impact of CETA on earnings (see, for example, Westat, 1982; Dickinson, Johnson, and West, 1986).

Thus, we believe that a total sample of 12,000--divided equally between adult men and women, participants and comparison group members (i.e., 3,000 each)--should be adequate to meet most states' analysis needs.[39] States that are interested in obtaining precise net impact estimates for subgroups of adult men or women, or that have additional resources should consider larger sample sizes as necessary. Finally, states with relatively small JTPA programs (i.e., less than 2,000 adult enrollees per year) should be very careful in interpreting the results as only very large impacts will be regarded as significantly different from zero. As a result, such states might consider pooling samples with other states or increasing the sample size over time to enhance the reliability of the net impact findings.

---

38 It should be recalled that Ashenfelter (1978) reports program impacts for women that persist over a five-year post-program period and for men that decline somewhat over the five-year period.

39 It should also be noted that with such a sample design it will be possible to detect approximately a 2.4 (2.8) percentage point impact on the probability of employment in the post-program year for adult men (adult women) with 90% power at a .10 significance level. These estimates assume that 83% of the adult men and 75% of the adult women have earnings in the post-program year, which are based on data for CETA participants.

CHAPTER 5

DATA ANALYSIS PLAN

Examining the Adequacy of the Comparison Groups—Obtaining Evidence on Selection Bias

Estimating the Net Impacts of JTPA Programs

Obtaining Net Impact Estimates for Various Subgroups

Measurement Error and Other Statistical Issues

Adjustments for Potential Data and Design Deficiencies

CHAPTER 5. DATA ANALYSIS PLAN

The final step in developing a model to obtain valid estimates of the net impacts of JTPA programs on the earnings and welfare dependency of adult men and women involves the specification of a data analysis plan. In particular, a data analysis strategy must be developed for examining the adequacy of the comparison groups selected and for using the comparison groups to estimate the net impacts of JTPA on participants' post-program outcomes. In this chapter, we describe an overall estimation strategy for obtaining valid estimates of the net impacts of JTPA.

As described in Chapter 4, our proposed approach involves a comparison of the post-program labor market experiences of a sample of JTPA participants with the experiences of a sample of economically disadvantaged ES registrants who do not receive ES services. If the samples of JTPA participants and ES registrants are similar on both measured (e.g., age, race, education) and unmeasured (e.g., attitude toward work, motivation) characteristics, then valid inferences about the impacts of JTPA programs can be drawn from such comparisons. However, whether an individual decides to participate in JTPA or decides to register with the ES is likely to depend on both individual and agency decisions, and because of these selection processes the two groups may differ on a number of dimensions that could result in biased net impact estimates. This is the issue of selection bias.

It is important to recognize that virtually all nonexperimental approaches will contain a certain amount of bias. That is, the formal conditions required to ensure unbiased estimates of program effects are not likely to be met, even if one had extensive data on the characteristics of program participants and comparison group members. This is particularly true for the proposed research design, which does not involve detailed survey data and, as a result, the amount of information available to control for selection bias is limited. As such, the fundamental issue is not whether any bias exists, but how large it is, whether some bounds can be placed on the likely direction and magnitude of the bias, and whether the assumptions under which it is possible to correct for the bias are at all realistic.

To provide information on such issues requires a well-designed analysis strategy and a careful interpretation of the results that recognizes the inherent limitations of the basic approach. Below we describe an analysis strategy for examining the adequacy of the comparison group selected to get a better understanding of the likely direction and magnitude of selection bias. We then describe alternative statistical models for estimating the average net impacts of JTPA and impacts for important subgroups, and indicate the conditions under which the various models will produce unbiased estimates of program impacts. The chapter concludes with a discussion of measurement error and other

statistical issues, and of potential adjustments for certain data and design deficiencies.

Before describing the analysis strategy, it is important to note that some of the material presented in this chapter is quite technical. Although we have attempted to make the discussion as accessible as possible to practitioners, a certain amount of technical presentation and statistical jargon is unavoidable. As such, program administrators will need to discuss issues of concern with individuals on their staffs who are qualified to carry out a net impact analysis. Those who have difficulty with the technical material should also consider consulting the implementation manual for the net impact model (Volume VI) which discusses the recommended approach in a somewhat less formal manner.

EXAMINING THE ADEQUACY OF THE COMPARISON GROUPS--OBTAINING EVIDENCE ON SELECTION BIAS

As indicated above, the most important threat to obtaining valid estimates of the net impacts of JTPA programs is the potential of selection bias, that is, systematic differences between the participant and comparison groups. Although a group of local ES registrants who do not receive ES services is likely to be the best comparison group available (given the limited resources states have for evaluation activities), it must be recognized that several selection processes affect who is included in the JTPA and ES groups. For example, JTPA participants must decide to apply to the program; they must meet certain legislated eligibility criteria; they must be selected by the agency for program participation and assigned to a program; and they must decide to accept that assignment and enroll in the program. Although ES registrants do not have to meet any formal eligibility criteria, certain individuals, such as those receiving benefits from UI, are required to register with the ES, and some offices follow selective registration policies. Furthermore, whether an ES registrant receives ES services depends on several factors, including the availability of suitable job openings and the qualifications and persistence of the applicant. Because of these various selection processes, it is unlikely that the resulting samples of JTPA participants and ES registrants who do not receive services are equivalent on all measured and unmeasured characteristics that affect labor market outcomes. As a practical matter, therefore, one should not focus on the fact that the two groups are not identical, but identify the major dimensions on which the groups differ; determine the extent to which the net impact estimates are likely to be sensitive to those differences; and adjust statistically for those differences to the extent that data and analysis resources allow.

In this section we describe three different criteria that can be used to judge the adequacy of the comparison groups selected and outline some recommended analyses to determine whether these criteria are met.[40] It should be noted that the criteria and analyses recommended

---

[40] It should be noted that the analyses described below could also be used as input to choosing among alternative potential comparison

here are very similar to those used by Westat (1982), Dickinson, Johnson, and West (1985, 1986), and Fraker and Maynard (1984) to evaluate the quality of matched comparison groups for estimating the net impacts of various employment and training programs. Before describing these analyses, however, it is important to recognize their inherent limitations. The fundamental limitation is that these analyses provide evidence only on the comparability of the two groups in the <u>pre-program</u> period, whereas the crucial assumptions for deriving unbiased net impact estimates concern the comparability of the two groups in the post-program period, which, of course, are inherently untestable. That is, these criteria are necessary, but not sufficient, conditions for the comparison groups to overcome the problem of selection bias. Thus, even if the two groups are comparable in the pre-program period, this should not be interpreted as definitive evidence that there is no selection bias, since changes in the underlying behavioral relationships could have occurred from the pre- to post-program period. Nevertheless, the analyses described below should provide useful information on the extent and types of differences between participants and comparison group members that must be kept in mind when interpreting the net impact results.

Criterion One: Similarity in Measured Characteristics

The first criterion that we recommend be used to judge the comparability of the participant and comparison groups is the similarity of the two groups on measured characteristics during the pre-program period. For example, one could compare the means and the distributions of measured characteristics for participants and comparison group members, separately for adult men and women.[41] It

---

groups. That is, because the choice of an appropriate comparison group is not entirely clear a priori, such analyses could be used to determine, for example, whether all ES registrants or JTPA no-shows would potentially be better comparison groups than ES registrants who receive no services.

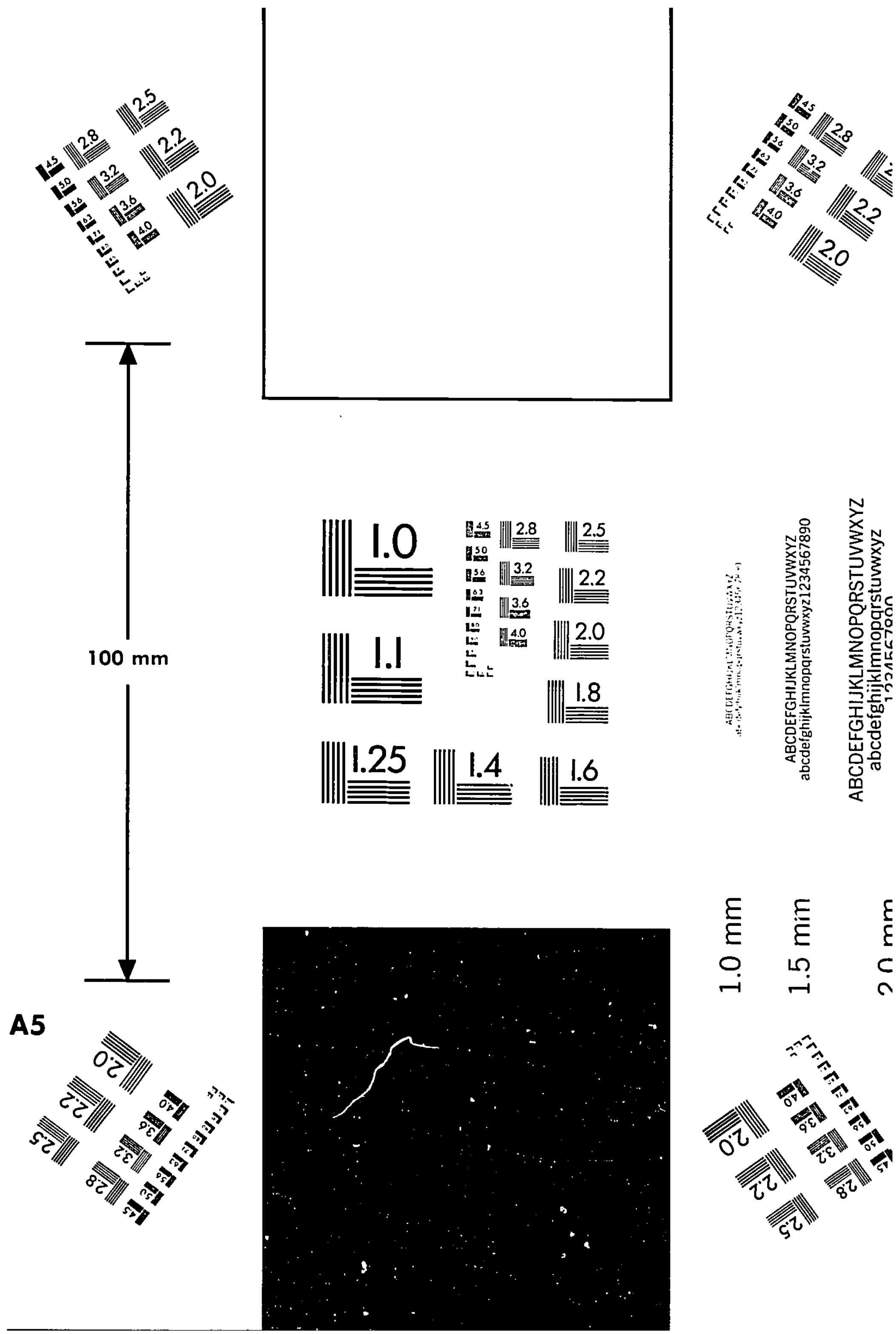
41 It would also be possible to estimate an OLS linear probability model of the likelihood of participating in JTPA to determine the major differences between the two groups. That is, one would estimate a regression equation with the dependent variable equal to 1 for JTPA participants and 0 for comparison group members, and the independent variables would be all measured characteristics included in the net impact model described later in this chapter. This approach has the advantage of estimating the independent influence of each measured characteristic, while controlling for the influence of all other characteristics, which eliminates the confounding effects of other variables that may be present when comparing mean characteristics. That is, a comparison of mean characteristics could indicate, for example, that JTPA participants are more likely to be minorities and less educated, whereas the regression approach would account for the differences in education by race and could reveal that, after adjusting for differences in race, there are no differences between participants and comparison group members in terms of education levels.

is particularly important to compare the participant and comparison groups on available measured characteristics that are known to affect earnings and AFDC grants. For example, because of the well-known differences in earnings and AFDC participation by race, one would compare the participant and comparison groups in terms of the proportion white, black, and Hispanic. Although differences in measured characteristics such as race can be controlled for directly in the net impact model and, as such, should not necessarily be of primary concern, major differences in measured characteristics between the two groups could be indicative of a selection process that also holds for unmeasured characteristics, which would cause a more serious analytical problem. For example, if one found that the JTPA samples of adult men and women had much larger proportions of minorities and high school dropouts than the samples of ES registrants, it could be that JTPA participants were also more disadvantaged on unmeasured characteristics (i.e., less motivated). This would result in underestimating the net impacts of JTPA.

It would also be useful to conduct similar analyses across JTPA program activities. That is, one would compare the characteristics of participants separately by program activity (e.g., CT, OJT, JSA). This would provide some indication of the additional selection bias that could arise in estimating separate net impacts by program activity. For example, if it were determined that more advantaged individuals were being sent to OJT, the net impacts of this program activity would be somewhat inflated because of this assignment process. That is, the observed positive relationship between earnings and the OJT treatment would not be entirely causal, but in part would reflect the fact that OJT participants were more advantaged (i.e., more motivated). On the other hand, if there were relatively few differences in measured characteristics by program activity, this evidence, in combination with other tests described below, would provide some confidence that no additional selection biases would be introduced in deriving estimates of the net impacts by program activity.

Criterion Two: Similarity in Preprogram Earnings and AFDC Grants

The second criterion we recommend be used to judge the adequacy of the comparison groups is the similarity of the pre-program earnings and AFDC grants of participants and comparison group members. This involves a test of whether there is a significant difference in the pre-program earnings and AFDC grants of the two groups, controlling for measured characteristics. Such a test provides valuable evidence on whether the two groups are comparable on the basis of the lagged dependent variables (after controlling for measured characteristics) or, alternatively, whether there are differences in the outcome variables between the groups in the pre-program period that are due to unmeasured characteristics. If there are any differences in adjusted pre-program earnings or AFDC grants between the two groups, then this analysis will also provide some hints as to the direction and magnitude of the selection bias. For example, the extent to which JTPA participants have larger (smaller) adjusted pre-program earnings than ES registrants provides some indication as to whether they are more (less) advantaged on the basis of unmeasured characteristics, and the


A5
100 mm
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz1234567890
1.0 mm
1.5 mm

size of the difference is a reasonable estimate of the amount by which the net program impacts could be overstated (understated) if the difference persisted in the post-program period.

To formally test for differences in the pre-program earnings (or AFDC grants) of participants and comparison group members, one would estimate a regression equation (separately for adult men and women) with pre-program earnings (or AFDC grants) as the dependent variable. For example, one could estimate the following equation using ordinary least squares (OLS) regression techniques:

$$(1) \quad Y_{i,t-1} = a + bX_{i,t-1} + vZ_i + \sum_{j=1}^{k} d_j Y_{i,t-j} + mM_{i,t-1} + cP_{i,t} + e_{i,t-1},$$

where t is the period of enrollment into JTPA; $Y_{i,t-1}$ represents the earnings of individual i in the immediate pre-program period t-1; $X_{i,t-1}$ represents the vector of exogenous characteristics (e.g., age, race, education) included in the net impact model and measured just prior to enrollment in training;[42] Z_i is a vector of interaction terms (e.g., between race and other independent variables in the model); $Y_{i,t-j}$ represents the earnings of individual i in earlier pre-program periods t-j; $M_{i,t-1}$ represents labor market characteristics facing individual i in the pre-program period; $P_{i,t}$ is 1 for JTPA participants and 0 for comparison group members; $e_{i,t-1}$ is a random error term; and a, b, v, d_j, m and c are coefficients to be estimated. The test for differences in pre-program earnings between the participant and comparison groups is based on a t-test of the estimated coefficient c, where the t-ratio is given by the estimated coefficient divided by its standard error.[43] Given that the sample sizes involved are likely to be large, the hypothesis of no difference between the two groups in pre-program earnings or AFDC grants would be rejected at the .05 (.01) significance level if the calculated t-ratio exceeded 1.96 (2.58) in absolute value.

On a more intuitive level, because participation in JTPA during period t can not have an effect on earnings in period t-1, the estimate of c

---

42 It should be noted that some of the $X_{i,t-1}$ characteristics included in the net impact model below should not be included in this pre-program earnings equation because the variables are jointly determined with pre-program earnings. For example, the Food Stamps dummy variable and the female WIN registrant dummy should be excluded from the pre-program earnings equation.

43 One could also estimate a similar pre-program earnings or AFDC grants equation and replace the JTPA participation dummy variable with a set of variables representing different program activities to provide information on the differences in pre-program earnings across program activities that are due to unmeasured characteristics. By testing whether the separate program activity coefficients are significantly different from zero, this would indicate whether there are likely to be any additional selection biases due to the nonrandom assignment of programs to individuals that will affect the net impact estimates by program activity that are described later in this chapter.

in the model described above should be 0. The extent to which c deviates from 0 provides evidence on the direction and magnitude of the likely selection bias. That is, large negative (positive) values of c indicate that participants were less (more) advantaged than comparison group members in the pre-program period on unmeasured characteristics, and if this persisted through the post-program period, it would likely result in understating (overstating) the net impact of JTPA. Thus, for example, if this analysis indicated that, after adjusting for differences in measured characteristics, the pre-program earnings of JTPA participants were $200 less (more) than the earnings of the comparison group, then one might consider adding (subtracting) $200 to (from) the net impact estimate to adjust for differences in unmeasured characteristics. It should be noted, however, that because pre-program earnings and AFDC grants will be included as independent variables in some net impact models, the extent of this bias should be less in the post-program period. As such, adjusting the net impact estimate for the total difference in pre-program earnings is likely to overcompensate for the bias due to differences in unmeasured characteristics.

It should be noted that with several periods of pre-program earnings data, it is possible to examine in more detail the comparability of the earnings dynamics in the pre-program period.[44] In particular, one could estimate a regression equation like the one described above for each pre-program period and derive a vector of estimated coefficients of c. The only difference in the equations would be that earnings in subsequent pre-program periods would have to be omitted from each equation (i.e., the equation for $Y_{i,t-3}$ would not include $Y_{i,t-2}$ or $Y_{i,t-1}$ as explanatory variables). One would expect that the extent of any bias in program impacts would tend to zero as a longer history of pre-program earnings is corrected for in the regression. That is, the largest (in absolute value) estimated coefficients of c should occur in the very early pre-program periods, and the coefficients should tend toward zero as additional lagged values of earnings are included in the equation.[45]

Criterion Three: Similarity of Preprogram Earnings Equations

The third criterion that we recommend be used to judge the adequacy of the comparison groups is the similarity of the earnings equations for

---

44 An alternative method of evaluating the similarity of the earnings dynamics of the participant and comparison groups involves comparing the correlation of earnings over time for the two groups in the pre-program period. See Cooley, McGuire, and Prescott (1979) for a discussion of this approach.

45 Such a pattern was observed in an evaluation of the impact of ES job referrals on earnings by Johnson, Dickinson, and West (1985) for men, but was not true for women, for whom the estimated values of c varied little across the different pre-program periods. Thus, they concluded that including a long series of lagged earnings accounted for much of the potential selection bias for men, but not for women.

JTPA participants and for the comparison groups in the pre-program period.[46] This criterion, which is considerably more strict than the previous two, is quite important because, if the same model is generating earnings in the two groups, it suggests that program impacts will not be sensitive to nonlinearities in the earnings equation. This would provide additional confidence in our ability to obtain unbiased estimates of program impacts.

To test for differences in the pre-program earnings equations of participants and comparison group members, one would estimate separate regression models for the two groups, as well as a regression model for the two groups combined. These models would be estimated separately for adult men and women. For example, to test whether the pre-program earnings models were the same for the two groups in period t-1, one would estimate the following equations using OLS regression techniques:

Participant Model:

$$(2a)\quad Y_{p,t-1} = a_p + b_p X_{p,t-1} + v_p Z_{p,t-1} + \sum_{j=2}^{k} d_{pj} Y_{p,t-j} + m_p M_{p,t-1} + e_{p,t-1},$$

Comparison Group Model:

$$(2b)\quad Y_{c,t-1} = a_c + b_c X_{c,t-1} + v_c Z_{c,t-1} + \sum_{j=2}^{k} d_{cj} Y_{c,t-j} + m_c M_{c,t-1} + e_{c,t-1},$$

and

Combined Model:

$$(2c)\quad Y_{i,t-1} = a + bX_{i,t-1} + vZ_{i,t-1} + \sum_{j=2}^{k} d_j Y_{i,t-j} + mM_{i,t-1} + e_{i,t-1},$$

where subscript p (c) refers to participants (comparison group members), and the third equation is estimated over all individuals (i), participants and comparison group members combined. Similar equations could be estimated for earlier pre-program periods to determine whether the earnings models are the same in these periods as well.

The test that the earnings equations are the same for the two groups is equivalent to testing that the effects of the measured characteristics on earnings are similar for the two groups. To perform this test, one must compute the following test statistic:

---

46 As we describe in a later section, the net impact models assume that this criterion is met in the post-program period.

$$\frac{[RSS - (RSS_c+RSS_p)]/r}{(RSS_c+RSS_p)/(N-K)} ,$$

where RSS is the residual sum of squares from the combined regression equation, RSS_p (RSS_c) is the residual sum of squares from the regression over participants (comparison group members) only, r is the number of restrictions involved (i.e., the number of independent variables in each equation), and N-K is the number of degrees of freedom when no restrictions are imposed (i.e., the sum of the number of degrees of freedom from the separate participant and comparison group models). Under the assumption that the error terms are normally distributed, the test statistic given above follows Snedecor's (1956) F-distribution with r degrees of freedom in the numerator and N-K degrees of freedom in the denominator. If the test statistic exceeded the critical value for the specified level of confidence, then the null hypothesis would be rejected (i.e., we would conclude that the earnings models for the two groups are not similar).[47] For example, suppose the earnings model had 10 independent variables and a total sample size of 6,000. Then, there would be 10 degrees of freedom in the numerator and 5,980 degrees of freedom in the denominator, so the test statistic would follow a F(10, 5,980) distribution. Given the critical values for the F-distribution with these degrees of freedom, one would conclude that the earnings models of the two groups were different at the .05 (.01) significance level if the test statistic exceeded 1.83 (2.32).

How to Interpret Evidence on the Adequacy of the Comparison Groups

The three criteria and related analyses described above should provide considerable information regarding the adequacy of the comparison groups in the pre-program period and the types of likely biases that must be dealt with. It should be emphasized again, that these criteria are relatively strict tests of the comparability of the two groups and one should not generally expect nonexperimentally-derived comparison groups to meet all of them. If the conditions are generally satisfied, then the chances of obtaining unbiased program net impact estimates using the standard statistical models described below are considerably improved. If the criteria are strongly rejected (e.g., F-statistics of 10 or 20 when approximately 1.5 is sufficient for rejection), then one should be very careful in proceeding to estimate net impacts with these comparison groups. Instead, one should first double check to ensure

---

47 An alternative method of conducting this test is available and can be performed with most standard software packages. Specifically, one would estimate an equation that included all of the explanatory variables in Equation (2c), plus each of the variables multiplied by the JTPA participation dummy variable. The formal test of whether earnings in the pre-program period are different for the two groups is based on an F-test of the hypothesis that the coefficients of the interaction terms are jointly insignificant. This is sometimes referred to as a "Chow" test.

that the data processing and analysis guidelines described earlier were followed and, if the criteria are still strongly rejected, one should then consider obtaining assistance from a researcher familiar with these issues. If, as is most likely, the results are somewhere in between (i.e., pre-program differences between the two groups that are sometimes statistically significant, but not exceptionally large), then one will need to understand the implications of these differences for interpreting and adjusting the net impact's results.

ESTIMATING THE NET IMPACTS OF JTPA PROGRAMS

We now turn to a description of alternative statistical models for estimating the average net impacts of JTPA programs on participants' post-program outcomes. In describing the alternative models, we attempt to carefully indicate the conditions under which each model will result in unbiased estimates of program effects. Based on the advantages and disadvantages of the models, we present a recommended model. Using the recommended statistical model, a subsequent section describes how to obtain separate estimates of net impacts for major demographic groups and by program activity. It should be noted that although the specific models described below may not be familiar to state-level analysts, each model can be estimated using standard estimation techniques contained in statistical software packages (e.g., SAS, SPSS) that should be readily available and familiar to them.

The statistical models presented below describe how to derive estimates of the net impacts of JTPA on post-program earnings. It should be emphasized, however, that this is for illustrative purposes only. To obtain estimates of the net impacts of JTPA on the receipt of AFDC grants by individuals, the only change required would be to replace the earnings variable in a particular period with AFDC grants in that period; the other variables to be tested for inclusion in preliminary models would be the same. The same general models would also be estimated to determine the net impacts of JTPA on the probability of employment or on the probability of AFDC participation. Depending on the specific outcome measure, however, some additional statistical issues arise. These issues are discussed in a later section.

It should be noted that some of the net impact models described below impose restrictions on the nature of the earnings generating process. Because these restrictions are more likely to be valid for "real" earnings (i.e., earnings expressed in constant dollar terms) than for nominal or money earnings, the earnings variables should be converted into constant dollar values by deflating them with the Consumer Price Index. This would be particularly important if the recent pattern of very low inflation rates (i.e., 4% per year) were to end and the double-digit inflation rates of the late 1970s returned. Because there are no valid price indices available at the state level, as indicated above we recommend that the BLS Consumer Price Index for All Urban Consumers be used to deflate nominal earnings and create measures of real earnings.[48]

---

48 For example, for a given nominal earnings series, Y_t,

Before describing the alternative net impact models, it is important to briefly discuss the choice of post-program periods to be examined. In estimating the net impacts of JTPA on post-program outcomes, one can either define the outcomes in terms of specific calendar periods, or in terms of specific periods after program termination. Virtually all previous studies of the net impacts of employment and training programs used earnings in a given time period, e.g., a specific calendar year, as the outcome measure for both participants and comparison group members. These studies also generally used enrollee-based participant samples (i.e., samples of participants who enrolled in the program during a given time period), and created a "pre-program period" that was common for all individuals and that was defined to end at the beginning of the enrollment period. As a result of these methodological choices, the estimated net impacts capture earnings changes from a pre-program period to a post-program period that corresponds to the identical calendar months/years for all participants and comparison group members. Because of differences in program length of stay, however, it must be recognized that such net impact estimates do not reflect impacts for a common period since termination. Consistent with this tradition, the models described below can produce net impact estimates for several different points during the post-program period, but these periods will not generally correspond to a similar period after termination for all participants unless there is little variation in program length of stay.

There is likely, however, to be considerable program and policy interest in estimates of net impacts that correspond to fixed periods after termination (e.g., first three months after termination, first year after termination). This is in part because of the recent development of JTPA follow up systems in many states, that track the employment status of participants at fixed points after termination, usually at three months and sometimes again later. Although the UI earnings data do not allow one to create outcome measures that precisely correspond to a given period after termination, it is possible to create earnings measures for various three-month periods after the calendar quarter of program termination. Thus, for individuals who terminate in a given calendar quarter, earnings received during the next calendar quarter, that are from three to six months after termination (4.5 months on average), could be used to approximate the impacts of JTPA for a common period after termination.

An important conceptual problem arises, however, in estimating net impacts for a fixed period after termination. This is because the concept of "date of termination" has no meaning for comparison group members and, as a result, additional complexities arise. In

---

Y_{t-1}, ..., Y_{t-k}, and a corresponding consumer price index series CPI_t, CPI_{t-1}, ..., CPI_{t-k}, with $CPI_t=100$, then a real earnings series in constant period t dollar values is given by Y_t/CPI_t, Y_{t-1}/CPI_{t-1}, ..., Y_{t-k}/CPI_{t-k}. If, instead, $CPI_{t-j}=100$, then it is necessary to first create a revised index with $CPI_t=100$--by dividing the entire CPI series by CPI_t--and then using the revised series as deflators in order to obtain a real earnings series in constant period t dollar values.

particular, to minimize biases using enrollee-based samples, one would have to select the "post-program periods" for the comparison group to match the distribution of participants' termination dates across calendar quarters. For example, suppose that of the individuals who enrolled in JTPA in the first quarter of 1986, 35 percent terminated in the first quarter, 50 percent terminated in the second quarter, and 15 percent terminated in the third quarter of 1986. Then, in estimating the net impact of JTPA on earnings during the 3 months after the calendar quarter of program termination, the real earnings outcome measure for participants would be based on earnings during the second quarter of 1986 for 35 percent of the sample, earnings during the third quarter for 50 percent of the sample, and earnings during the fourth quarter for 15 percent of the sample. To avoid biases, the comparison group would have to be aligned similarly. That is, the 35, 50, and 15% distributions across quarters would have to occur for the comparison group as well. Although this post-program alignment problem could be avoided if a terminee-based sample of participants were used (i.e., since the period after termination would be common for all participants and comparison group members), this would create a comparable alignment problem in the pre-program period. That is, because of differences in length of participation, participants would have enrolled in JTPA over several previous quarters, and comparison group members would have to match this distribution in the pre-program period.

Thus, although it is possible to estimate net impacts for outcome measures that approximate the impacts for common post-termination periods, we do not believe the additional data processing and analysis complications are worth the effort. Moreover, some of the econometric models we describe require that the outcomes be measured for the same calendar periods for all participants and comparison group members. As a result of these considerations, the net impact models described below are based on outcome measures and independent variables that are defined for the same calendar periods (both pre-program and post-program) for both participants and comparison group members.

Autoregressive Earnings Model

Several recent nonexperimental evaluations of the net impacts of employment and training programs (e.g., Westat, 1982; Dickinson, Johnson, and West, 1986; and Geraci, 1984) have primarily relied on analysis-of-covariance approaches that are variants of Ashenfelter's (1978) autoregressive earnings model. Using this approach, these studies included as independent variables in the regression equation the pre-program earnings history and all socioeconomic characteristics that were comparably measured for participants and comparison group members. In addition, interaction terms were also included in order to control as much as possible for potential differences between the two groups. Thus, an autoregressive earnings model that would be applicable for evaluating the net impacts of JTPA at the state level would have the following general form:

$$(3) \quad Y_{i,s} = a + bX_{i,t-1} + vZ_i + \sum_{j=1}^{k} d_j Y_{i,t-j} + mM_{i,s} + cP_{i,t} + e_{i,s},$$

where t is the period of training; $Y_{i,s}$ is the earnings of individual i in a post-program period s; $X_{i,t-1}$ is a vector of pre-program socioeconomic characteristics of individual i, where the t-1 subscript is included to emphasize that the characteristics are measured at application to JTPA (and at a similar time for ES registrants); Z_i is a vector of interaction terms (e.g., between age and race, education and race, pre-program earnings and race) to account for the fact that the effect of various characteristics on earnings may differ by race or personal characteristics; $Y_{i,t-j}$ are the earnings of individual i in various pre-program periods t-j, j=1,...,k; $M_{i,s}$ is a vector of labor market characteristics applicable to individual i in post-program period s (all individuals located in the area served by a given SDA would have the same value during this period); $P_{i,t}$ is a dummy variable that is 1 for JTPA participants and 0 for comparison group members; $e_{i,s}$ is an error term; and a, b, v, d_j, m and c are parameters to be estimated. Such a model could be estimated using ordinary least squares regression techniques, and separately for different post-program periods (s) to obtain information on the timing of the net impacts for a given cohort of trainees or for several cohorts combined.

With such an autoregressive model estimated separately for adult men and women, the estimated coefficient of the JTPA participant dummy variable (c) represents the average net impact of JTPA on the outcome measure for the particular post-program period (s) for both groups. For dependent variables that are expressed in dollar terms such as earnings and AFDC grants, the coefficient of the JTPA participant dummy variable can be interpreted as the average dollar impact on the particular outcome measure. By dividing the estimated dollar impact by the mean earnings or AFDC grants of comparison group members, this provides an estimate of the percentage change in earnings or AFDC grants due to JTPA.

For dummy dependent variables such as whether employed in a particular period, or whether receiving AFDC grants during a particular period, the autoregressive net impact model is equivalent to a linear probability model. That is, the model essentially estimates the effects of various factors on the probability of a certain event occurring (e.g., having positive earnings in a given post-program period). As such, the estimated coefficient of the JTPA participant dummy variable can be interpreted as the average percentage point change in the probability of working or receiving AFDC grants due to JTPA. Once again, by dividing the estimated percentage point change by the mean proportion of comparison group members, one can obtain an estimate of the percentage change in the probability of working (or receiving AFDC) due to JTPA.

With this model, the estimated coefficient of c will provide an unbiased estimate of the average net impact of JTPA on post-program earnings in period s provided two conditions are met: (1) all characteristics that affect earnings and program participation are included in the model, so that the correlation between the error term and the participation dummy variable is zero, and (2) the earnings model that would prevail in the absence of training is the same for JTPA participants and comparison group members (i.e., the third criterion for the adequacy of the comparison group described earlier

must be met in the post-program period).[49] The first condition--that all factors that affect program participation and the outcome measure be included in the model--is equivalent to saying that any unmeasured characteristics that affect earnings are uncorrelated with program participation.[50] Thus, if the measured characteristics included in the net impact model capture the factors that affect the program participation selection process, it is possible to obtain unbiased net impact estimates of program effects using an autoregressive earnings model. For example, Ca'n (1975) has shown that if selection into the program is determined solely on the basis of pre-program earnings (or other measured characteristics in the model), then consistent net impact estimates can be obtained from an autoregressive model. This is true even if the error term in the earnings equation ($e_{i,s}$) is autoregressive (as is likely to be the case), as long as the error term in the program participation equation is uncorrelated with $e_{i,s}$. Thus, although it is well known that the coefficients of the lagged earnings variables will be biased if the error term is autoregressive, it is possible to derive consistent estimates of net program effects from such an autoregressive model.

An advantage of an autoregressive earnings model like the one described above that places no restrictions on the relation between current and prior earnings (except for linearity) is that it controls for differences in measured characteristics between JTPA participants and comparison group members that remain after the matched comparison group is selected. At the same time, however, because of the limitations of the proposed research design discussed earlier, the independent variables that are comparably measured for JTPA participants and ES registrants and that are likely to be available in most states for inclusion as control variables (i.e., $X_{i,t-1}$) in the net impact model are severely limited. For example, based on data currently included in the various program MIS systems for the State of Washington, the following demographic/socioeconomic variables could be created and included in the net impact model:

- Age (in years);
- Age-squared[51];

---

49 In addition, we are assuming that all variables in the model are perfectly measured. The consequences of measurement error are discussed later in this chapter.

50 Note that this condition does not require that all factors that affect earnings must be included in the model, but only those that affect both program participation <u>and</u> earnings. If some characteristics that only affect earnings (but not program participation) are not measured, and therefore omitted from the model, then the coefficients of other control variables could be biased, but not c.

51 Age-squared is included to capture the nonlinear relationship between earnings and age that has been documented extensively in the literature. An alternative procedure would be to create mutually exclusive and exhaustive age categories (e.g., 22-24, 25-29, 30-34, 35-39, 40-44, 45 and older) and replacing age and age-squared with dummy variables for the age categories selected.

- Black dummy variable (1 if Black; 0 otherwise);
- Hispanic dummy variable (1 if Hispanic; 0 otherwise);
- Other nonwhite race dummy variable (1 if not Black, Hispanic, or White; 0 otherwise);
- Grade 12 dummy variable (1 if completed high school only; 0 otherwise);
- Grade 13+ dummy variable (1 if attended some college; 0 otherwise);
- Handicapped dummy variable (1 if person has a physical or mental impairment that is a substantial handicap to employment; 0 otherwise);
- Male veteran dummy variable (1 if male and served in the Armed Forces; 0 otherwise);
- Food Stamps dummy variable (1 if receiving Food Stamps at application; 0 otherwise);
- Female WIN registrant dummy variable (1 if female and a WIN registrant at time of application; 0 otherwise);
- UI payments received in the pre-program calendar quarter;
- 1-digit occupation dummy variables (1 if occupation prior to application is a specific 1-digit DOT code; 0 otherwise). Treating professional and managerial occupations as the left-out category, the occupational dummies to be included would be:

    -clerical and sales
    -service
    -agriculture, fishing, and forestry
    -processing
    -machine
    -benchwork
    -structural, and
    -miscellaneous.

Although this is a fairly extensive list of demographic and socioeconomic characteristics, it is far from ideal. Specifically, these variables would not allow one to directly control for differences in marital status, family size, presence of dependent children, ex-offender status, limited English speaking ability, or pre-program employment intensity such as total hours worked or total weeks worked. To the extent that these unmeasured factors affect program participation, as well as earnings, then the autoregressive earnings model described above could produce biased estimates of net program impacts. At the same time, however, it is important to keep in mind that many of these characteristics were also not available to previous national studies of the impact of employment and training programs, so that this potential limitation is not unique to a state-level model. Moreover, the proposed model has an advantage over previous studies that partially overcomes this data deficiency in that, in addition to the demographic/socioeconomic characteristics listed above, the model will also include quarterly values of pre-program earnings (or AFDC grants) covering as long a period as data permit (e.g., up to 12 quarters if possible). These quarterly histories will enable one to better control for pre-program differences in earnings or welfare dependency patterns that could in part reflect differences in unmeasured characteristics between the participant and comparison groups and, as such, will help to reduce the problem of selection bias.

An additional advantage of the state-level model is that it will be possible to control for differences in labor market conditions. Previous national studies were unable to control for labor market variables in examining the net impacts of employment and training programs since the geographic location of participants and comparison group members was unknown, which prevented the linking of labor market data to specific individuals. The number of different measures of labor market conditions that can be included in the autoregressive earnings model will primarily depend on the data available, the number of SDAs in the state (since all persons who reside in the service area of a given SDA will receive the same value for the local labor market characteristic in a particular period), and the extent to which the local labor markets differ across SDAs. At a minimum, it is desirable to include the local unemployment rate as a control variable in the autoregressive earnings model.

Fixed-Effects Models

As described above, the autoregressive earnings model controls for differences in measured characteristics between participants and comparison group members to the extent possible and can yield consistent estimates of program net impacts under certain conditions. However, such models cannot control for differences in unmeasured characteristics (e.g., ability, motivation) between the two groups that may be introduced by the various selection processes involved. That is, if individuals either self-select into JTPA or are selected by program operators on the basis of unmeasured characteristics, and these characteristics affect earnings, this would introduce a correlation between the error term and the participation dummy variable in the net impact model, which would lead to biased estimates of program effects. In this section we describe an alternative model--known as a fixed-effects model--that under certain conditions can overcome problems caused by differences in unmeasured characteristics between participants and comparison group members.

The fixed-effects model, used by Bassi (1983, 1984), Kiefer (1979), and Ashenfelter (1978), explicitly recognizes that an individual's earnings and the decision to participate in JTPA are likely to depend on unmeasured characteristics. The underlying motivation for the fixed-effects model is the following general earnings equation:

(4) $Y_{i,t} = a + bX_{i,t} + cP_{i,t} + e_i + e_t + e_{i,t}$,

where as before, $Y_{i,t}$ is the earnings of individual i in period t; $X_{i,t}$ represent measured characteristics that affect earnings; $P_{i,t}$ is a dummy for program participation; and a, b, and c are parameters to be estimated. Unlike the model described earlier, however, the error term is assumed to have three components to explicitly recognize that earnings depend on unmeasured characteristics. Specifically, the disturbance term includes a permanent component, e_i, that is specific to an individual and constant over time to capture factors such as ability and motivation; a time-specific component, e_t, that is the same for all individuals and that captures the effects of economy-wide changes in earnings during period t; and $e_{i,t}$, a purely random,

serially uncorrelated error term. It is also assumed that these error components are independent. If the process of selecting individuals into JTPA is unrelated to these unmeasured components, then an unbiased estimate of c could be obtained by estimating the above earnings equation using OLS. If, however, selection is related to the unmeasured components, then OLS estimates of this equation will yield biased program effects.

The fixed-effects model is designed to handle situations in which selection into the program is based on e_i, the individual-specific unmeasured characteristics that are constant over time.[52] Specifically, if individuals either choose to participate in JTPA or are selected by program operators on the basis of e_i, the bias caused by the correlation between the error term and participation can be eliminated by differencing the impact equation over two time periods. That is, for a post-program period s, and a pre-program period k, differencing yields the following estimation equation:

$$(5)\quad Y_{i,s} - Y_{i,k} = (e_s - e_k) + b(X_{i,s} - X_{i,k}) + cP_{i,t} + (e_{i,s} - e_{i,k}).$$

As can be seen from this equation, by calculating the change in earnings from a pre-program to a post-program period, the permanent component of earnings due to unmeasured characteristics is eliminated. In addition, since most exogenous measured characteristics such as race and education are constant over time, the permanent components of earnings due to measured characteristics (bX_i) will also be eliminated, except perhaps for age. Thus, if selection into JTPA depends only on unmeasured characteristics that do not change over time, if the $Cov(e_{i,s} - e_{i,k}, P_i) = 0$, and if the earnings structures of participants and comparison group members are similar, then regressing the difference in earnings on a constant, a dummy variable for JTPA participation, and age will yield unbiased estimates of JTPA.[53]

It is important to note, however, that there is good reason to believe that the $Cov(e_{i,s} - e_{i,k}, P_i)$ will not be zero, particularly for pre-program periods near enrollment. That is, since program operators are evaluated on their success in meeting certain performance standards related to participants' post-program employment outcomes, they may "cream" by choosing individuals who become eligible primarily because of bad luck. That is, they may choose individuals with good long-term earnings prospects but who have large negative transitory components in earnings in the immediate pre-program period, i.e., large negative values of $e_{i,t-1}$. Such a selection process would introduce a correlation between $e_{i,t-1}$ and P_i, and lead to biased estimates of program impacts if the immediate pre-program period were used in the

---

52 In the autoregressive earnings model, the fixed effects are in part accounted for through the inclusion of prior earnings levels as regressors.

53 Differencing reduces the degree of any polynomial by a factor of one, so that if the initial model includes age and age-squared, then the difference model will only include age.

difference equation. In fact, use of the immediate pre-program period as the base period for the fixed-effects estimator would result in an upwardly biased estimate of the net program impact on earnings, because the transitory component of earnings would subsequently return to normal and earnings would rise, even in the absence of JTPA.

It is possible, however, to reduce the effects of creaming on $e_{i,k}$ by using a fixed-effects estimator with a base period considerably prior to program participation (e.g., two or three years) as long as the random component of earnings is not correlated over time. In fact, if the model is correct and the random component of earnings is uncorrelated over time, then net impact estimates using different base periods (e.g., two, three, or four years prior to enrollment) should be equal. Unfortunately, evidence presented by Ashenfelter (1986) and Dickinson, Johnson, and West (1986) indicates that estimates of program impacts on earnings from fixed-effects models vary widely using different base years. This extreme sensitivity of the fixed-effects estimator to the choice of the base year is likely due to the fact that the random component of earnings is correlated over time, as demonstrated by Lillard and Willis (1978) and others. The correlation of the random component of earnings over time brings us to the third approach to deriving unbiased net impact estimates.

Symmetric-Difference Estimators

In a recent paper, Heckman and Robb (1982) demonstrated that a fixed-effects or difference-estimator symmetric about the decision period will yield consistent program effects even if the random component of earnings is correlated over time, as long as the earnings process is covariance stationary.[54] That is, under the assumption of stationarity, the difference in earnings from a pre-program to a post-program period that are equally far from the participation decision period is not affected by the transitory decline in earnings that leads trainees to enroll in JTPA. Thus, if the decision period is t-1, the symmetric-difference estimator is equivalent to the fixed-effects model described above, with $k = s-(t-1)$.

Results from Dickinson, Johnson, and West (1986) indicate, however, that the symmetric-difference estimator is sensitive to the choice of the decision period. This is in part because their results were based on a sample of individuals who enrolled in CETA throughout a particular calendar year, and for whom the appropriate decision year likely differed depending on whether they enrolled early or late in the year. That is, individuals who enrolled very early in the year likely based their decision to enroll on the previous year's earnings--indicating that the difference estimator should be symmetric about period t-1--whereas individuals who enrolled late in the year probably based their decision primarily on the current year's earnings, which would result in a difference estimator symmetric about period t. As a result, very different net impact estimates were obtained when t and t-1 were used

54 Covariance stationary means that the correlation between a variable at time period t and t-k is the same for a given k and all t.

as the decision year for all enrollees.

A major advantage of the state-level model is that because quarterly earnings data will be available, it will be possible to more easily identify the decision period(s). In general, it is likely that the decision to participate will depend primarily on the earnings in the three to six months before the quarter of enrollment in JTPA. If, for example, it is earnings in the six months before the quarter of enrollment that determine program participation, then for a group of trainees who enrolled in the first quarter of 1986, one could obtain unbiased net impact estimates by regressing the change in earnings from a pre- to post-program period symmetric around the midpoint of the decision period (i.e., symmetric around October 1, 1985) on a constant and a JTPA participation dummy variable.[55] For example, to determine the net impact on annual earnings one could use earnings measured two years before and after the decision period. That is, one could use earnings from April 1983 to March 1984 (with a midpoint of October 1, 1983) as the base period and earnings from April 1987 to March 1988 (with a midpoint of October 1, 1987) as the outcome measure. This illustrates a major disadvantage of the symmetric-difference approach, namely its extensive longitudinal earnings requirements. That is, to estimate the net impact on annual earnings using the base periods and outcome periods chosen above, one would need five years of data, and the results would not be available in a timely fashion.

Although it is possible to use other base and outcome periods that are symmetric about the decision period, but closer to the decision period, so that results could be available in a timely fasion, other problems arise. For example, if one used earnings from April 1984 to March 1985 as the base period (i.e., one year prior), then the appropriate outcome period would be from April 1986 to March 1987. This period is unsatisfactory since a large portion of individuals who enrolled during the first quarter of 1986 would still be in the program. Treating program participants as having completed the program when in fact they are still enrolled would lead to a biased estimate of program impacts.

In addition to the extensive longitudinal data requirements of the symmetric-difference estimator, it should be noted that this estimator may be sensitive to potential seasonality biases if less than a year is used as the outcome period. For example, provided all trainees who enrolled in JTPA during the first quarter of 1986 were terminated by October 1, 1986, then to estimate the net impact on quarterly earnings, the first post-program quarter for all participants would be from October 1, 1986 through December 31, 1986. This outcome period has a

---

55 It is also possible to include a set of demographic variables to control for differences in earnings trends among individuals with different measured characteristics. For example, one could include age, ethnicity, and education as control variables in the symmetric-difference model. However, it is important that no pre-program earnings variables or any of the other variables that depend on earnings in the pre-program period (e.g., Food Stamp recipient, WIN registrant) be included in the model, as that would violate the constraints implicit in the model.

midpoint of November 15, 1986, or 13.5 months after the midpoint of the assumed decision period for these enrollees. Therefore, the appropriate base period for the symmetric-difference model would be from July 1, 1984 through September 30, 1984 (with a midpoint of August 15, 1984 or 13.5 months before the midpoint of the decision period). That is, the symmetric-difference model would use earnings in the fourth quarter of 1986 minus earnings in the third quarter of 1984 as the dependent variable.[56] If there are seasonal differences in the characteristics of JTPA enrollees or if the net impacts of the program vary by season, changes in earnings over different calendar quarters could give misleading estimates of the overall net impacts of JTPA.

Summary and Recommended Approach

In this section we have presented three statistical models for estimating the average net impacts of JTPA on earnings. Each of the models rests on certain assumptions (e.g., that the model is correctly specified, that all variables are perfectly measured, and other assumptions concerning the nature of the earnings generating process, the structure of the unmeasured characteristics, and the process that determines selection into JTPA) that may be difficult to meet in practice. In addition, the models differ somewhat in their data requirements and ease of use. Below, we briefly summarize the strengths and weaknesses of the various approaches and present a recommended analysis strategy.

The analysis of covariance or autoregressive earnings model controls for differences in measured characteristics between participants and comparison group members that remain after the match is drawn, but does not control for any unmeasured differences that affect both earnings and program participation that may be introduced by the various selection processes. It can provide consistent estimates of program effects if program participation depends only on pre-program earnings or on other measured characteristics. This estimation model is quite simple to implement, has been used extensively in the literature, and has relatively modest data requirements.

The other two types of estimation models make assumptions about how unmeasured characteristics affect earnings and then propose changes in earnings models from a pre-program to a post-program period to reduce the effects of unmeasured characteristics. The fixed-effects model assumes that the effects of unmeasured characteristics on earnings are constant over time, whereas the symmetric-difference estimator is somewhat more general and allows the effects of unmeasured characteristics to be correlated over time, provided the earnings

---

56 Other potential dependent variables for measuring the net impacts on quarterly earnings include earnings in the first quarter of 1987 minus earnings in the second quarter of 1984, earnings in the second quarter of 1987 minus earnings in the first quarter of 1984, and so on. The point is that the pre- and post-program quarters involved in the symmetric-difference estimator always correspond to different calendar quarters.

generating process is covariance stationary. Each of these approaches is somewhat more complex and is likely to be less intuitive to state-level analysts, which would make the results more difficult for them to explain and defend. In addition, these models have considerably greater data requirements, which will limit their usefulness in states that do not keep detailed longitudinal UI earnings records, and the results will generally not be available on as timely a basis.

Based on considerations related to data requirements, ease of implementation, and understanding, the autoregressive earnings model is preferable. Moreover, based on limited evidence, it appears that net impact estimates from autoregressive earnings models are similar to those obtained from symmetric difference estimators when the appropriate decision year is used.[57] <u>As such, we recommend that the autoregressive earnings model be the primary method for estimating the net impact of JTPA, but that appropriate cautions be made about the potential biases</u>. (For more detailed discussion on how to implement this model for estimating JTPA net impacts, refer to Volume VI in this series.) In addition, we recommend that states with detailed longitudinal UI earnings records (five years or more) consider using the symmetric-difference estimator to determine if correcting for unmeasured characteristics affects the estimated net impacts.

OBTAINING NET IMPACT ESTIMATES FOR VARIOUS SUBGROUPS

The models described above have focused on providing overall estimates of the net impacts of JTPA for adult men and women. It is also of considerable policy importance to determine whether the effectiveness of JTPA varies by the type of service provided or by individual characteristics. Because these factors may change considerably over time, knowledge of how program net impacts vary among them would help interpret time trends in JTPA's impacts. Furthermore, information on which programs work best for which types of participants and under what conditions could provide valuable information for targeting future employment and training programs to the disadvantaged. Although the approach to estimating net impacts for different subgroups is formally identical, whether the subgroup refers to program service type or individual characteristics, additional selection biases are likely to arise in some situations. Below we describe how to modify the autoregressive earnings model described above to estimate the net impacts of JTPA for various subgroups and indicate the additional biases that can arise.

---

57 Results from Dickinson, Johnson, and West (1986) indicate that the autoregressive earnings model and the symmetric-difference estimator yield very similar results for adult men when the appropriate decision year is used. However, the net impact estimates for adult women are somewhat more positive when the symmetric-difference estimator is used. This suggests that correcting for differences in unmeasured characteristics may be more important for women, and to the extent that the selection processes in JTPA are similar to those in CETA, the estimates from an autoregressive earnings model may be lower bound estimates of the net impacts for adult women.

In general, subgroup effects are estimated by including in the regression equation an "interaction" term, that is, a term representing the product of the dummy variable for JTPA participation with the variable for the subgroup of interest. Suppose one is interested in testing whether the net impact varies by the characteristic represented by the three dummy variables Z_1, Z_2, and Z_3. For example, one might think of the three variables as representing race/ethnicity categories (white, black, other), or program services (CT, OJT, JSA).[58] Then, using the autoregressive earnings model described above, one would estimate the following model to determine whether JTPA net impacts differed across these subgroups:

$$(6)\quad Y_{i,s} = a + bX_{i,t-1} + vZ_i + \sum_{j=1}^{k} d_jY_{i,t-j} + mM_{i,s} + c_1P_{i,t}Z_1 + c_2P_{i,t}Z_2 + c_3P_{i,t}Z_3 + e_{i,s}.$$

That is, the only change to the basic model described in Equation (3) involves replacing the program participation dummy variable with three variables that each involve the JTPA dummy variable multiplied by one of the three variables representing the particular subgroup. Then c_1 represents the net impact for the first subgroup (e.g., whites), c_2 represents the net impact for the second subgroup (e.g., blacks), and c_3 is the estimate of the net impact for the third subgroup (e.g., other race/ethnicity groups).

To formally test whether the program net impacts differ significantly across the subgroups of interest, an F-test is used. Similar to the F-test described earlier in this chapter, this test involves a comparison of the residual sum of squares from two regression equations, the main net impact equation (in which the treatment variable is entered directly and is not interacted with any subgroup characteristic), and one in which the main treatment variable is omitted and the treatment is interacted with variables representing each subgroup. Specifically, the following test statistic would be computed:

$$(7)\qquad \frac{[RSS_m - RSS_i]/r}{RSS_i/(N-K)},$$

where RSS_m and RSS_i are the residual sum of squares from the main model and the model with subgroup interactions respectively, r is the number of restrictions imposed by the main model (i.e., the number of

---

58 It should be noted that, in principle, similar analyses could be performed to determine whether net impacts vary across local labor market conditions. However, because the labor market variables would take on the same value for all persons in the same local area in a given time period, there is not likely to be sufficient variation in these variables to obtain precise estimates of how program impacts vary across local labor market conditions, except in large states, with many SDAs, and where there are considerable differences in labor market conditions across SDAs.

subgroups minus one), and N-K is the number of degrees of freedom in the main impact model. This test statistic follows an $F(r,N-K)$ distribution, and the null hypothesis that the net impacts do not vary across the subgroups of interest (e.g., across racial groups) would be rejected for r=2 and sufficiently large sample sizes at a .05 (.01) significance level if the test statistic exceeded 2.99 (4.60).

Although the F-test described above is the appropriate test for the equality of the net impacts across several subgroups, one can also provide information on differences in net impacts by subgroups using an alternative estimation strategy. For example, consider the following regression model:

$$(8)\quad Y_{i,s} = a + bX_{i,t-1} + vZ_i + \sum_{j=1}^{k} d_j Y_{i,t-j} + mM_{i,s} + c_1P_{i,t} + n_2P_{i,t}Z_2 + n_3P_{i,t}Z_3 + e_{i,s}.$$

In this formulation, the program dummy variable and all except one of the interaction terms are included in the model; that is, Z_1 becomes the "left-out category." Then, c_1 is an estimate of the net impact for subgroup Z_1 as before; n_2 is an estimate of the extent to which the program impact differs for subgroup Z_2 relative to subgroup Z_1; and n_3 measures the extent to which the program impact for subgroup Z_3 differs from that for subgroup Z_1. In fact, the two equations are essentially equivalent in that $c_2 = c_1 + n_2$, and $c_3 = c_1 + n_3$. The advantage of estimating the equation in the latter form is that one can directly test which program impacts are significantly different from the left-out category, because the standard errors of n_2 and n_3 are available. That is, separate t-tests of the estimated coefficients n_2 and n_3 provide information on whether the estimated impacts for the two subgroups (Z_2 and Z_3) are significantly different from the impact for Z_1. On the other hand, if one wants to determine whether each of the subgroup impacts is significantly different from zero, then estimates of the standard errors of c_1, c_2, and c_3 are required, and the earlier specification should be used.

It should be noted that in attempting to disaggregate JTPA net impacts across subgroups, it is important that the subgroup characteristics also be included in the model as control variables to account for differences in the general level of earnings across these subgroups. That is, in our illustration, the three Z_i variables must also be in the model separately so that the estimated coefficients (i.e., the c_i's or n_i's) only capture earnings differences due to JTPA across these subgroups and do not include the average earnings differences due to the Z_i's themselves. In addition, in order to ensure that the c_i's are capturing meaningful differences in program net impact by subgroups, it is important that the subgroups be mutually exclusive and exhaustive. That is, if the equation above that has the program participation dummy interacted with <u>all</u> the subgroup dummy variables of interest had omitted one of the interaction terms, then those observations would be in fact treated as part of the comparison group, and biased estimates of the net impacts of JTPA for different subgroups would be obtained. Fortunately, this is a simple problem to avoid.

The interaction analyses described above will identify the types of individuals who benefit most from JTPA and whether there is a general pattern to the variation in program effectiveness. Because individuals' pre-program characteristics cannot be affected by JTPA, no additional selectivity bias is introduced in disaggregating JTPA net impacts by demographic subgroups.[59] However, as we describe below, this may not be the case when examining whether JTPA effectiveness varies by program activity or by other aspects of the treatment provided.

In principle, to probe beneath the average net impacts of JTPA and provide information on the program activities and other aspects of the treatment (e.g., length of stay) that contributed to the average effects, one would perform an identical interaction analysis to the one described above, except that the subgroup characteristics (i.e., the Z_i's) would relate to JTPA services received. For example, to examine net impacts by program activity, then if Z_1, Z_2, and Z_3 corresponded to classroom training, on-the-job- training, and job search assistance, respectively, then c_1 would be the estimate of the average net impact for CT, c_2 would represent the estimated net impact for OJT, and c_3 would represent the estimated net impact for JSA. However, as we describe below, there is a major problem that threatens the internal validity of the by-program activity net impact analysis.

The problem is the familiar one of selection bias. In this context, it relates to the nonrandom assignment of program services to JTPA participants. As described above, the assignment of program activity is likely to be based on the agency's perception of an individual's needs and abilities. To the extent that this assignment process is based solely on the measured characteristics of participants (i.e., age, race, sex, education, pre-program earnings), this will not bias the net impacts by program activity as these characteristics will be included in the net impact model. A much more serious problem arises if the assignment of program activities is based on unmeasured characteristics, such as motivation and ability, and those unmeasured characteristics also affect earnings. If, for example, the more-motivated participants are assigned to OJT programs and are also

---

59 If the characteristics defining the subgroups of interest are not measured equally well for the participants and comparison group members, however, the subgroup impacts will inappropriately reflect these differences. That is, because the presence of measurement error in an independent variable biases its estimated coefficient downward (see discussion in later section), if the amount of measurement error on a subgroup characteristic were greater in the JTPA sample, for example, than in the comparison group, the effect of that characteristic on earnings would be smaller in the JTPA sample than in the ES registrant sample. The interaction term would inappropriately pick up such a difference and misleadingly indicate that JTPA impacts were smaller for individuals with that characteristic. It may be useful, therefore, when examining the quality of the comparison group to make sure that the variables used in the interaction analysis do not have preexisting differential impacts in the two samples.

more likely to have higher earnings, then the estimated coefficient of OJT (c_2) may be large and positive. However, this significant coefficient of OJT on earnings would not indicate a causal relationship, but merely reflect the fact that more-motivated individuals were assigned to the OJT activity. Thus, one must be very careful in interpreting net impacts by program activity because of this additional selection bias that can occur as a result of the assignment process.

As described in the literature review (see Chapter 2), there are statistical procedures that can potentially be used to correct for such selection biases. These procedures rely on instrumental variable techniques to produce consistent estimates of program impacts in the presence of selection bias. However, as noted above, such procedures are only useful if one can identify variables to play the role of instruments. That is, one must find variables that affect the assignment of program activities to particular individuals but that do not affect participants' earnings. Although this is a difficult task in any circumstance, it will be particularly difficult given the limited number of individual characteristics available for the proposed analysis.

It may, however, be possible to identify potential instrumental variables from a carefully structured process analysis. That is, an important aspect of the process analysis will be a detailed description of the process involved in assigning program activities to specific participants. Such an analysis might indicate, for example, that the assignment of program activities is influenced primarily by the availability of program-activity slots in a given SDA, and since the availability of program slots should not affect post-program earnings, such a variable could potentially serve as an instrument.

Even if the process analysis is not successful in identifying specific variables to play the role of instruments, it could shed light on the relative validity of the program-activity net impact estimates and provide some indication of the likely direction of the selection biases. For example, the process analysis should be able to determine to which program activities the more- and less-job ready participants are assigned.[60] Then, if the impact analysis finds that a given program activity has a large effect on post-program earnings and the process analysis indicates that the more disadvantaged individuals are assigned to that activity, it would increase confidence that the observed relationship is causal and not due to selection bias. If, however, the process analysis indicates that the individuals assigned to that program activity are considerably more advantaged, then one should not be very confident that the net impact estimates for that program activity solely reflect the effects of the program.

In addition to providing evidence on how the net impact of JTPA varies by individual characteristics and program activities, it may be useful

60 A comparison of the differences in measured characteristics of participants across program assignments should also provide some information on the probable direction of the selection bias.

to examine whether the impact of JTPA varies by participants' program experiences. For example, several studies have attempted to determine how the impacts of employment and training programs vary with length of stay in the program. Although the results of such analyses could provide important information about the mechanisms through which employment and training programs produce their effects, as we describe below there are important selection bias problems that limit the validity of such analyses.

To investigate whether the net impacts of JTPA vary by program length of stay, one would estimate an autoregressive earnings model with the overall program participation dummy variable interacted with variables representing length of participation. For example, using the notation developed above, one could categorize the length of stay variable into four dummy variables corresponding to stays of less than one month, one to three months, three to six months, and greater than six months. Then the estimates of the four c_i coefficients would represent the average impacts of JTPA for individuals with these different lengths of stay. Alternatively, if the effects of length of stay on earnings are approximately linear, one could estimate a model with a JTPA participation dummy and the participation dummy interacted with actual weeks in the program less average number of weeks in the program, and obtain a more efficient estimate of program effects. In this formulation, the coefficient of the JTPA dummy represents the estimated impact of JTPA at the average length of stay, and the coefficient of the interaction term is an estimate of the dollar impact of an additional week of program participation.

Although it is straightforward to construct such program experience variables and estimate the coefficients of the interaction terms in the net impact models, one must be very careful in interpreting the results. Once again, this is due to the familiar problem of selection bias. Although the autoregressive earnings model controls for differences on measured characteristics between short- and long-term participants, it is likely that some differences on unmeasured characteristics remain. Individuals who leave the program early may be less motivated or, alternatively, may have found employment on their own. On the other hand, individuals who stay in the program a long time may do so because they have fewer other employment opportunities. Length of stay is also likely to depend on the type of program activity and SDA characteristics.

Because of these additional selection bias problems, one should use extreme caution in interpreting the estimated impacts by length of stay as representing causal relationships. In order to overcome these biases and obtain estimates of the net impacts of JTPA by program length of stay that one has confidence in, it is necessary to formally account for the endogeneity of the length of stay variables, which would require instrumental variable procedures. Operationally, one would first estimate a regression equation to obtain a predicted value of length of stay. Such equations would be estimated using the participant sample only, and the predicted values would be entered in the net impact equation for participants, and zeros would be included for comparison group members.

The success of such an instrumental variable procedure relies heavily on the ability to identify variables to play the role of instruments. That is, one must find variables that are highly correlated with length of stay but that do not directly influence earnings. Depending on the types of information obtained from participants and included in the JTPA MIS, one may be able to obtain variables that can play the role of instruments for length of stay. For example, it may be possible to develop instruments from answers to very simple questions about participants' general satisfaction with JTPA services and whether participants were assigned to the type of program they wanted. That is, it seems plausible to assume that individuals who are satisfied with the program, or who are assigned to the program type they were most interested in, would remain in JTPA longer. At the same time, however, there is no obvious reason why satisfaction with the program or assignment to the desired training program should affect earnings, independent of the effect of the training and the actual length of stay. Thus, these variables could be entered into a regression equation along with other demographic characteristics to explain program length of stay, and then a predicted value for length of stay could be constructed and entered in the net impact equation in place of the actual length of stay.

To summarize, estimating how the net impacts of JTPA vary by program experiences introduces an additional selection bias that is in practice difficult to overcome. For the most part, these additional biases have been recognized in the literature, but not explicitly dealt with due to the lack of variables to play the role of instruments. In the absence of adequate instruments, one must not place too much confidence in the estimated impacts. For states that are very interested in overcoming these selection bias problems, they should carefully review what is in the JTPA MIS, as well as the information obtained during the process analysis, to see if variables that affect length of stay, but not earnings, can be identified. If successful, they should implement the instrumental variable approach and determine whether the variables selected in fact strongly affect length of stay. If not, then it is not necessary to go to the second stage of including the predicted value in the net impact model, as the predicted variable will be too highly correlated with other characteristics in the model, and the results will not be reliable. If, however, the instruments do strongly affect length of stay, then the net impact model should be estimated with the predicted length of participation replacing the actual value. If the instruments have been successful, then they will purge the correlation between the error term and length of stay and result in consistent net impact estimates.

MEASUREMENT ERROR AND OTHER STATISTICAL ISSUES

In estimating the net impacts of JTPA programs on earnings and AFDC grants, we have recommended that ordinary least squares (OLS) regression techniques be used. Although such techniques are easy to implement and are generally very robust, it is important to recognize that, depending on the specific outcome measure and the population group of interest, OLS regression may not be the most appropriate net impact estimation technique. At the same time, however, the

alternative models that may be more appropriate are in general not contained in standard statistical software packages and thus will not be readily available to state-level researchers. Below we briefly discuss some of the limitations of OLS regression models for the planned analyses and briefly indicate other statistical models that could be used.

Measurement Error

In the regression models described above, we have treated the variables as though they are perfectly measured, that is, as though they contain no measurement errors. Although this is a very convenient assumption, it is also somewhat unrealistic. In fact, it is likely that most variables contain some measurement error, although the nature and extent of the error may be such as to not affect the estimated program net impacts. Below we discuss the implications of various types of measurement error for the state-level net impact model.[61]

We first consider the implications of measurement errors in the dependent variable. As described in Chapter 3, in developing post-program earnings variables from State UI Wage Records, these variables will be subject to measurement errors for individuals who work in nonreported employment or who work in another state. In both cases, zero earnings will be recorded in the data set when the true value is some positive amount. Such measurement error in the dependent variable can bias the estimated net impacts of JTPA on earnings _if_ treatment status is correlated with the measurement error in earnings. That is, if the likelihood of working in nonreported (or out of state) employment is different for JTPA participants than for comparison group members, OLS estimates of program net impacts will be biased. Specifically, if JTPA participants are more (less) likely to work in nonreported employment or out of state, the estimated impacts will be biased downward (upward). If, however, the measurement error in the dependent variable is uncorrelated with treatment status, the estimated program net impact will not be biased.

It should be noted, however, that even though the net impact estimates may not be affected by measurement error in the dependent variable, the precision of the estimated program effects will generally be affected. Assuming that the measurement error in earnings and the error term in the net impact model are uncorrelated (or positively correlated), the presence of measurement error increases the variance of the estimated net impacts. Thus, if one suspects that the dependent variable is likely to suffer from considerable measurement error, it is possible to compensate for the expected loss of precision by drawing larger analysis samples.

We next consider the complications that are introduced when there is measurement error in an independent variable. Independent variables may be measured with error because of recording or keypunch errors,

---

61 For additional information see Duncan (1975), Goldberger and Duncan (1973), or Rao and Miller (1971).

because of data editing procedures that assign default values when the data item is missing for a person, or as a result of inappropriate aggregation. For example, measurement error can be introduced by collapsing heterogeneous JTPA treatments into one common treatment dummy variable. That is, by representing the effect of JTPA by a simple participant dummy variable that combines diverse treatments such as classroom training, OJT, or Job Search Assistance that have different lengths of participation and intensity, one necessarily introduces measurement error into the treatment variable. As is well known (see, for example, Duncan, 1975), measurement error biases the coefficient of the independent variable toward zero. This is true for all independent variables with measurement error, that is, variables representing the treatments provided by JTPA as well as variables included to control for differences in measured characteristics between participants and comparison group members. The magnitude of the bias depends on the size of the variance in the measurement errors relative to the variance of the error term in the net impact model. Specifically, the greater the variance in the error of measurement--that is, the lower the accuracy of measurement--the greater is the bias towards zero. This emphasizes the importance of developing independent variables that accurately measure the extent of the treatment provided, which in turn reinforces the importance of conducting a process analysis to identify the major differences and similarities in program treatments.

As the above discussion indicates, measurement error is an important conceptual issue that analysts should be aware of when specifying their models and interpreting the results. In addition, it should be noted that approaches have been developed to formally incorporate measurement error into standard statistical models (see, for example, Joreskog, 1973). However, these multiple indicator approaches are very complex and generally beyond the scope of the state-level net impact model. As a result, states that anticipate severe measurement problems should consult a measurement specialist, perhaps from within their university system.

Dichotomous Dependent Variables

The models described earlier in this chapter have concerned continuous outcome measures such as earnings and AFDC grants. However, as described in Chapter 3, some of the outcome variables to be examined are dichotomous. Examples include whether an individual is employed or whether he or she is receiving AFDC payments during a particular period. In the case of dichotomous outcome variables, the prediction from a regression model represents the probability of the individual being employed or on welfare. However, a standard OLS regression model suffers from two deficiencies in this circumstance. First, the predicted probability is not constrained to fall between 0 and 1, and thus nonsensical results can potentially arise. Second, the error term in the model is necessarily heteroscedastic, and OLS estimation is inefficient.

A number of statistical procedures have been designed to deal with these circumstances, the most popular being logit and probit

analyses.[62] Although such models are theoretically more appropriate for dichotomous variables, in practice the predicted probabilities are likely to be quite similar to OLS estimates, provided the dichotomous outcome measure does not represent an extremely rare or extremely likely event. That is, as long as the mean outcome measure is approximately between .2 and .8, OLS estimates of a linear probability model usually yield very similar predictions to logit or probit models. Because this will generally be the case for the outcome measures to be examined in the net impact model, the deficiencies of OLS techniques should not be a serious problem for dichotomous outcome variables in the state-level model.

Bounded Dependent Variables

A related statistical issue arises for outcome measures that although continuous, have an upper or lower bound at which a large number of observations occur. For example, AFDC grants are zero for nonparticipants and earnings are zero for nonworkers. In the relatively short post-program period, many individuals can be expected to have zero earnings, and even more individuals are likely to have zero AFDC grants. If there are numerous observations at the limit, ordinary least squares will be inconsistent and will also be inefficient, because the error term is heteroscedastic. One way of accounting for the statistical problems is to estimate a tobit model (Tobin, 1958), which is designed to handle outcome variables that are truncated normal, i.e., normally distributed but truncated at a particular point. Unfortunately, there is extensive evidence that the distributions of earnings for adult men and women are not truncated normal in which case the tobit specification is inappropriate and could yield misleading results.

Thus, although it is important to recognize the limitations of OLS techniques, the potential deficiencies of alternative procedures, in combination with their much greater expense and limited availability, suggest that the state-level model should rely on OLS techniques.

ADJUSTMENTS FOR POTENTIAL DATA AND DESIGN DEFICIENCIES

The models described above involve a comparison of the UI earnings records of JTPA participants with those of a comparison group derived from ES registrants. In addition to the problem of selection bias, there are some deficiencies in the UI Wage Records and in the ES data that may affect the results. For example, as indicated above, UI Wage Records contain measurement error in that they do not reflect earnings from jobs that are not in reported employment or earnings from jobs that are located across the border in other states. The ES data are deficient because there is inadequate information on whether ES registrants participated in JTPA, which may result in a "contaminated" comparison group. In this section, we briefly discuss the likely extent to which the basic impact estimates will be affected by these

---

62 For a discussion of these alternative models see Theil (1971).

data and design deficiencies and indicate the types of adjustments that may be necessary.

To the extent that it is impossible to identify individuals in the comparison group who participated in JTPA, the comparison group will be "contaminated." Such contamination would lead to an underestimate of the net impacts of JTPA, since it would effectively dilute the treatment as comparison group members would have also received JTPA services. We do not believe, however, that the net impact estimates will be seriously affected by the potential contamination problem. This is so for two reasons.

First, it may be possible to reliably identify ES registrants who have received JTPA services. This could be done either by comparing the social security numbers of comparison group members with current and past lists of JTPA participants and excluding all matches. Alternatively, the ES service file includes information that allows one to identify individuals who recently enrolled in JTPA programs. Although this information is likely to be inadequate both because of underreporting problems and the fact that the information only relates to the current year, it can be used to minimize the contamination problem.

Second, we believe that the magnitude of the contamination problem is likely to be small. Although the ES is one source of applicants for JTPA programs and, as such, one might expect that contamination could be high, existing data suggest otherwise. For example, based on data for the State of Washington for program year 1985, only 0.1 percent of all ES registrants active during the year enrolled in JTPA programs. Moreover, only 0.3 percent of those economically disadvantaged enrolled in JTPA. Although the figures are somewhat higher for enrollment in any training program (e.g., JTPA, Job Corps, WIN, other)--1.0 percent for all applicants and 3.1 percent for those economically disadvantaged --even these participation rates are small enough to be safely ignored.

In states that for some reason have considerably higher probabilities of ES registrants enrolling in JTPA, it may be necessary to make some adjustment to the net impact estimates. Although individual adjustments are not possible, one can make an aggregate adjustment. Specifically, if p is an estimate of the proportion of the comparison group participating in JTPA, the estimated program net impact should be multiplied by $1/(1-p)$ to obtain the true impact.[63]

The major limitation to the UI Wage Records is that some jobs are not in reported employment and earnings from those jobs are therefore not included in the post-program earnings measure. As discussed above, the omission of nonreported earnings (i.e., measurement error) can bias the estimated impacts of JTPA if program participation affects the

---

63 Because the autoregressive model controls for the impacts of previous participation, an adjustment should be made for each period from the decision to enroll through the post-program period of interest. Thus, p is properly thought of as a cumulative participation rate that in this application spans roughly a one to two year period.

probability of working in nonreported employment. Although this was an important concern of earlier studies, particularly given the high likelihood of CETA participants turning their training slots into subsequent jobs in the public sector (which are less likely to be in reported employment), given the focus of JTPA on employment in the private sector, this should be less of a problem for the state-level model.

There is very little information available on the extent to which nonreported earnings is likely to bias net impact estimates. The only study that has provided any evidence on the impact of employment and training programs on the likelihood of working in nonreported employment is Dickinson, Johnson, and West (1986). Using SSA earnings records and interview-reported earnings, they created a measure of whether a person was working in nonreported employment (i.e., SSA earnings of zero and positive interview-reported earnings) and examined whether CETA participants were more likely to be working in uncovered employment in the post-program period. They found that adult male participants were generally slightly less likely to be working in jobs that were not covered by Social Security, but that there were no differences for adult women. Thus, if these results also applied to UI earnings reporting problems and to the sample of JTPA participants and comparison group members selected for the state-level model, this would suggest that the underreporting problem could generate net impact estimates for men that are somewhat overstated, but that no adjustments for women would be necessary. However, since these results do not directly apply to the issues facing the state-level model, one should be reluctant to use such results to adjust net impact estimates for the reporting problems of UI Wage Records.

In the absence of additional information, it will be very difficult to determine the extent to which JTPA impacts are affected by the incomplete coverage of UI Wage Records. To determine whether participants and comparison group members differ in terms of their likelihood of working in nonreported employment, one would need information on interview-reported earnings and UI Wage Records for both groups in the post-program period. Alternatively, if one had information on the industry of employment of all individuals, one might get some sense of whether the earnings patterns are consistent with the nature of nonreported employment across industries. However, because such data must come from surveys of both participants and comparison group members, it is unlikely that such information will be available for the net impact model. In that event, the best one can do is acknowledge the potential problem and indicate that the net impact estimates assume that JTPA does not affect the probability of working in nonreported employment or implement more complex methodologies that directly incorporate measurement error.

Finally, the most important potential adjustment to be considered is for selection bias, that is, systematic differences between participants and comparison group members that cannot be directly controlled for in the autoregressive net impact model. One can attempt to adjust for selection bias by using evidence on pre-program differences between the two groups. For example, one could use estimates of differences in adjusted pre-program earnings or AFDC

grants between participants and comparison group members that are due to unmeasured characteristics that are developed as part of the analysis to examine the adequacy of the comparison groups described earlier in this chapter. The size of the estimated difference in pre-program earnings due to unmeasured characteristics is a reasonable estimate of the amount by which earnings impacts could be overstated or understated (depending on whether it is positive or negative) if the difference persisted in the post-program period. Thus, for example, if adult men (women) JTPA participants are estimated to have earned $100 more ($200 less) in the immediate pre-program year than individuals in the comparison groups, using this first approach one would adjust the main impact estimate (i.e., the coefficient of the JTPA dummy variable) downward (upward) by $100 ($200) for men (women). It should be noted, however, that because pre-program earnings are controlled for in the autoregressive net impact model, this adjustment may overcompensate for the selection bias due to differences in unmeasured characteristics between the two groups.

REFERENCES

REFERENCES

Ashenfelter, Orley, "The Effect of Manpower Training on Earnings: Preliminary Results," in Proceedings of the Twenty-Seventh Annual Winter Meeting of the Industrial Relations Research Association, 1975, pp. 252-260.

__________, "Estimating the Effects of Training Programs on Earnings," Review of Economics and Statistics, February 1978, pp. 47-57.

Barnow, Burt S., Glen G. Cain, and Arthur S. Goldberger, "Issues in the Analysis of Selection Bias," in Ernst W. Stromsdorfer and George Farkas (eds.), Evaluation Studies Review Annual Volume 5, Beverly Hills, CA: Sage Publications, Inc., 1980, pp. 42-59.

Bassi, Laurie J., "The Effect of CETA on the Post-program Earnings of Participants," Journal of Human Resources, Fall 1983, pp. 539-556.

__________, "Estimating the Effect of Training Programs with Non-random Selection," Review of Economics and Statistics, February 1984, pp. 36-42.

Bassi, Laurie J., Margaret C. Simms, Lynn C. Burbridge, and Charles L. Betsey, "Measuring the Effect of CETA on Youth and the Economically Disadvantaged," The Urban Institute, April 1984.

Berkeley Planning Associates, "Performance Standards for Reductions in Welfare Dependency," report prepared for US DOL under Contract No. 99-3-3229-75-103-01, September 15, 1984.

Bloom, Howard S., and Maureen A. McLaughlin, "CETA Training Programs--Do They Work for Adults," Joint CBO-NCEP Report, July 1982.

Borus, Michael E., John P. Brennan, and Sidney Rosen, "A Benefit-Cost Analysis of the Neighborhood Youth Corps," Journal of Human Resources, Spring 1970, pp. 139-150.

Cochran, W. G., and Rubin, D. B., "Controlling Bias in Observational Studies, a Review," Sankhya Series A, 1973, pp. 417-446.

Cook, Robert F., Richard L. Beatty, H. Allan Hunt, Sarah F. Liebschutz, Gretchen E. Maclachlan, Bruce Maclachlan, Susan A. MacManus, Mark J. Morlock, Kenneth T. Palmer, Alex N. Pattakos, V. Lane Rawlins, Kalman Rupp, and Wayne M. Turnage, "Transition Year Implementation of the Job Training Partnership Act," Westat, Inc., Rockville, Maryland, January 1985.

Cooley, T. F., T. W. McGuire, and E. C. Prescott, "Earnings and Employment Dynamics of Manpower Trainees: An Exploratory Econometric Analysis," in Farrell E. Bloch (ed.), Evaluating Manpower Training Programs, Greenwich, Connecticut: JAI Press, 1979, pp. 119-147.

Dickinson, Katherine P., "Supported Work for Ex-Addicts, "Journal of Human Resources, Fall 1981, pp. 551-599.

Dickinson, Katherine P., Terry R. Johnson, and Richard W. West, "An Analysis of the Impact of CETA Programs on Participants' Earnings," report prepared for DOL under Contract No. 20-06-82-21, SRI International, Menlo Park, California, November 1984a.

Dickinson, Katherine P., Terry R. Johnson, and Richard W. West, "An Analysis of the Impact of CETA Programs on Participants' Earnings," Journal of Human Resources, forthcoming Winter 1986.

Dickinson, Katherine P., Terry R. Johnson, and Richard W. West, "An Analysis of the Impact of CETA Programs on Components of Earnings and Other Outcomes," report prepared for DOL under Contract No. 20-06-82-22, SRI International, Menlo Park, California, November 1984b.

Dickinson, Katherine P., Terry R. Johnson, and Richard W. West, "Summary of an Analysis of the Impact of CETA Programs on Participants' Earnings and Implications for Evaluating JTPA," prepared for JTLS Technical Advisory Panel, January 1985.

Duncan, Otis D., Introduction to Structural Equation Models, New York, NY: Academic Press, 1975.

Finney, D.J., Probit Analysis, Cambridge University Press, 1964.

Fraker, Thomas and Rebecca Maynard, "An Assessment of Alternative Comparison Group Methodologies for Evaluating Employment and Training Programs," report prepared for DOL under Contract No. 20-11-82-15, Mathematica Policy Research Inc., Princeton, New Jersey, June 1984.

Gay, Robert S., and Michael E. Borus, "Validating Performance Indicators for Employment and Training Programs," Journal of Human Resources, Winter 1980, pp. 29-48.

Geraci, Vincent J., "Short-Term Indicators of Job Training Program Effects on Long-Term Participant Earnings," report prepared for DOL under Contract No. 20-48-82-16, August 1984.

Goldberger, Arthur S. and Otis D. Duncan (Eds.), Structural Equation Models in the Social Sciences, New York, NY: Seminar Press, 1973.

Heckman, James J., "Sample Selection Bias as a Specification Error," Econometrica, January 1979, pp. 153-161.

Heckman, James J., and Richard Robb, "The Longitudinal Analysis of Earnings," unpublished manuscript, June 1982.

Heckman, James J., and Richard Robb, "Alternative Methods for Evaluating the Impact of Interventions: An Overview," unpublished manuscript, 1985.

Johnson, Terry, R., Katherine P. Dickinson, and Richard W. West, "An Evaluation of the Impact of ES Referrals on Applicant Earnings," Journal of Human Resources, Winter 1985, pp. 117-137.

Joreskog, Karl, "A General Method for Estimating a Linear Structural Equation System," in A.S. Goldberger and O.D. Duncan (Eds.), Structural Equati Models in the Social Sciences, New York, NY: Seminar Press, 1973.

Kiefer, Nicholas M., "Federally Subsidized Occupational Training and the Employment and Earnings of Male Trainees," Journal of Econometrics, Augu 1978, pp. 111-125.

_________, "The Economic Benefits from Four Government Training Programs," in Farrell E. Bloch (ed.), Evaluating Manpower Training Program Greenwich, Connecticut: JAI Press, 1979, pp. 159-187.

Kirschner Associates, Inc., Evaluation of Five Concentrated Employment Programs in Regions I and II, New York: Kirschner Associates, Inc., 1969.

Lillard, Lee, and Robert Willis, "Dynamic Aspects of Earning Mobility," Econometrica, September 1978, pp. 985-1012.

Maddala, G., and L. F. Lee, "Recursive Models with Qualitative Endogenous Variables," Annals of Economic and Social Measurement, December 197 pp. 525-545.

Mallar, Charles, et al., Evaluation of the Economic Impact of the Job Corps Program: Third Follow-up Report, Mathematica Policy Research, Inc. Princeton, New Jersey, September 1982.

Masters, Stanley H., "The Effects of Supported Work on the AFDC Target Group," Journal of Human Resources, Fall 1981, pp. 600-636.

Rao, Potluri, and Roger L. Miller, Applied Econometrics, Belmont, CA: Wadsworth Publishing Company, Inc., 1971.

Rubin, D. B., "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies," Journal of the Americ Statistical Association, 1979, pp. 318-328.

Schiller, Bradley, "Lessons from WIN: A Manpower Evaluation," Journal of Human Resources, Fall 1978, pp. 502-523.

Somers, Gerald G., and Ernst W. Stromsdorfer, A Cost-Effectiveness Study of the In-School and Summer Neighborhood Youth Corps, Industrial Relatio Research Institute, University of Wisconsin, 1970.

Stafford, Frank P., "Optimal Sample Sizes for Evaluating JTPA," unpublished manuscript, May 1985.

West, Richard W. and Katherine P. Dickinson, "Development of Adjustment Model for PY 85 JTPA Performance Standards," report prepared for DOL und Contract No. 99-3104-75-101-01, SRI International, Menlo Park, Californi April 1985.

Westat, Inc., "Net Impact Report Number 1 (Supplement Number 1): The Impact of CETA on 1978 Earnings," report prepared for DOL under Contra No. 23-24-75-07, Westat, Inc., Rockville, Maryland, July 1982.

Westat, Inc., "Summary of Net Impact Results," report prepared for DOL under Contract No. 23-24-75-07, Westat, Inc., Rockville, Maryland, April 1984.

APPENDIX A

WASHINGTON STATE APPLICATION FORMS FOR
JTPA AND THE EMPLOYMENT SERVICE

EXHIBIT A-1

Washington State JTPA Application Form

APPLICATION FORM

Job Training Partnership Act Program

SUB CODE ____________

SSN ____ - ___ - _____ - __

NAME
Last ____ First ____ M __

ADDRESS
Street ____ Apt # ____

City ____ Cnty ____ St ____ Zip ____

Phone (___) ___ - ____ Message Phone (___) ___ - ____

Date of Birth

__/__/__
M M D D Y Y

Age ___

Sex __

F Female
M Male

Race/Ethnic Group __

1 White
2 Black
3 Hispanic
4 American Indian/ Alaskan Native
5 Asian/Pacific Islander

SSN └─┴─┴─┘ - └─┴─┘ - └─┴─┴─┴─┘ - └─┘ Name ______________________________

Application Date └─┴─┘/└─┴─┘/└─┴─┘

SELECTIVE SERVICE └─┘
1 Registered
2 Not registered
3 Not applicable

CITIZENSHIP └─┘
1 US Citizen
2 Eligible Non-Citizen
3 Non-citizen

IF ELIGIBLE NON-CITIZEN, ALIEN CARD NUMBER
____

FOSTER CHILD └─┘
1 Yes
2 No

HANDICAPPED └─┘
1 Physical
2 Mental
3 Not Applicable

SSI └─┘
1 Yes
2 No

RECEIVING PUBLIC ASSISTANCE

AFDC └─┘ CASE NO. └─┴─┴─┴─┴─┴─┴─┴─┴─┘ - └─┘ GRANT AMOUNT $ └─┴─┴─┴─┴─┘
START DATE ___/___/___
1 Yes
2 No

REFUGEE ASSISTANCE └─┘ CASE NO. └─┴─┴─┴─┴─┴─┴─┴─┴─┘ - └─┘ GRANT AMOUNT $ └─┴─┴─┴─┴─┘
START DATE ___/___/___
1 Yes
2 No

GENERAL ASSISTANCE └─┘ CASE NO. └─┴─┴─┴─┴─┴─┴─┴─┴─┘ - └─┘ GRANT AMOUNT $ └─┴─┴─┴─┴─┘
START DATE ___/___/___
TYPE ______________________________
1 Yes
2 No

FOOD STAMPS (ONLY) └─┘ CASE NO. └─┴─┴─┴─┴─┴─┴─┴─┴─┘ - └─┘
START DATE ___/___/___
1 Yes
2 No

WIN REGISTRANT └─┘
1 Yes
2 No

UNEMPLOYMENT COMPENSATION STATUS └─┘
1 Eligible Claimant
2 U.C. Exhaustee
3 Not Applicable

WEEKLY BENEFIT AMOUNT $ └─┴─┴─┘

OFFICE USE ONLY

Family Size └─┴─┘

Annualized Family Income $ └─┴─┴─┴─┴─┘

Economically Disadvantaged ___
1 Yes
2 No

Reason ______________________________

10% Window Reason ______________________________

SSN [ ][ ][ ] - [ ][ ] - [ ][ ][ ][ ] - [ ]   NAME ____________________

EDUCATION STATUS [ ]
1 School Dropout
2 Student (high school or less)
3 High School or Equivalent
4 Post High School Attendee

LIMITED ENGLISH [ ]
1 Yes
2 No

DISPLACED HOMEMAKER [ ]
1 Yes
2 No

DISPLACED WORKER [ ]
1 Yes
2 No

MEMBER OF A SEASONAL/ MIGRANT FARM FAMILY [ ]
1 Yes
2 No

SINGLE PARENT [ ]
1 Dependent Children Under Six Years
2 Dependent Children Over Six Years
3 Not Applicable

TEENAGE PARENT [ ]
1 Yes
2 No

OFFENDER [ ]
1 Yes
2 No

ALCOHOLIC/ ADDICT [ ]
1 Yes
2 No

VETERAN [ ]
1 Yes
2 No

RECENTLY SEPARATED [ ]
1 Yes
2 No

VIETNAM ERA VETERAN [ ]
1 Yes
2 No

DISABLED VETERAN [ ]
1 Yes
2 No

DATES OF MILITARY SERVICE
Entry ____/____/____
Exit ____/____/____

PRIOR JTPA PARTICIPATION [ ]
1 Yes
2 No

LABOR STATUS [ ]
1 Employed
2 Unemployed
3 Not in Labor Force

LAST JOB TITLE [ ]   DOT [ ][ ][ ]   LAST HRLY WAGE $ [ ][ ][ ].[ ][ ]   HOURS PER WEEK [ ][ ]

RECEIVED LAYOFF NOTICE [ ]
1 Plant Closure
2 Job Eliminated
3 Other ____________
4 No Notice

DATE NOTICE WAS ISSUED
____/____/____
M M D D Y Y

OPPORTUNITY FOR RE-EMPLOYMENT [ ]
1 Good   3 Poor
2 Fair   4 N/A

NUMBER WEEKS UNEMPLOYED IN LAST 26 WEEKS [ ][ ]

NUMBER WEEKS EMPLOYED IN LAST 13 WEEKS [ ][ ]

CERTIFICATION

I certify that the information provided is true to the best of my knowledge. I am also aware that the information I have provided is subject to review and verification and I may have to provide document to support this application. I am also aware that I am subject to immediate termination if I am found ineligible after enrollment and may be prosecuted for fraud and/or perjury if I intentionally supplied inaccurate or misleading information. I allow release of this information for verification purposes and understand that it will be used to determine eligibility. I have been advised of equal opportunity and appeal rights and the Privacy Act of 1974.

________________________________ Application Date ____/____/____
Signature of Applicant   Month Day Year

________________________________ Date ____________
Signature of Parent, Guardian or Responsible Adult   Month Day Year

________________________________ Date ____________
Signature of Interviewer   Month Day Year

Sub Code [ ][ ][ ][ ][ ]   Counselor [ ][ ][ ]

EXHIBIT A-2 Washington State Employment Security Application For Service

PLEASE PRINT

WORK HISTORY (Minimum last 3 years)

PRESS HARD

MOST RECENT OR PRESENT EMPLOYER	JOB SUMMARY (does what, using what, to what)
ADDRESS / CITY / STATE / ZIP	
JOB LOCATION if different / JOB TITLE / SALARY	
DATE STARTED / DATE LEFT / MONTHS ON JOB / PAY UNIT 1__HOURS 3__WEEK 5__YEAR 2__DAY 4__MONTH 6__OTHER	
REASON FOR SEPARATION 1____VOLUNTARY QUIT 3____LABOR DISPUTE 5____LACK OF WORK 2____DISCHARGE 4____ILLNESS 6____STILL EMPLOYED	
EMPLOYER	JOB SUMMARY (does what, using what, to what)
ADDRESS / CITY / STATE / ZIP	
JOB LOCATION if different / JOB TITLE / SALARY	
DATE STARTED / DATE LEFT / MONTHS ON JOB / PAY UNIT 1__HOURS 3__WEEK 5__YEAR 2__DAY 4__MONTH 6__OTHER	
REASON FOR SEPARATION 1____VOLUNTARY QUIT 3____LABOR DISPUTE 5____LACK OF WORK 2____DISCHARGE 4____ILLNESS	
EMPLOYER	JOB SUMMARY (does what, using what, to what)
ADDRESS / CITY / STATE / ZIP	
JOB LOCATION if different / JOB TITLE / SALARY	
DATE STARTED / DATE LEFT / MONTHS ON JOB / PAY UNIT 1__HOURS 3__WEEK 5__YEAR 2__DAY 4__MONTH 6__OTHER	
REASON FOR SEPARATION 1____VOLUNTARY QUIT 3____LABOR DISPUTE 5____LACK OF WORK 2____DISCHARGE 4____ILLNESS	

13. I HEREBY REGISTER FOR WORK AND/OR REQUEST AN INITIAL DETERMINATION OF BENEFITS POTENTIALLY PAYABLE TO ME UNDER THE WASHINGTON EMPLOYMENT SECURITY ACT AND/OR THE FEDERAL UNEMPLOYMENT COMPENSATION ACT. THE BASE YEAR AND POTENTIAL BENEFITS HAVE BEEN EXPLAINED TO ME AND I CHOSE TO FILE ON THIS DATE. I HEREBY CERTIFY THE INFORMATION I HAVE PROVIDED ON THIS FORM IS ACCURATE.

EMS 5327/511 (1-86) -436-

FULL SIGNATURE OF APPLICANT ______________________ DATE ____________

3. NAME: LAST FIRST M.I.

4. ADDRESS

5. CITY 6. STATE 7. ZIP

8. OTHER NAME(S) used during last 2 years

9. PHONE/MESSAGE PHONE

10. HIGHEST GRADE COMPLETED
Grade School High School/GED
0 1 2 3 4 5 6 7 8 9 10 11 12
College
13 14 15 16 17 18 19

11. BIRTHDATE MONTH DAY YEAR

12. SEX 1__MALE 2__FEMALE

13. WOULD YOU RELOCATE? 1____ YES 2____ NO

14. DATES OF ACTIVE MILITARY SERVICE Entered______ Released______

15. I have served in the armed services of the United States during the period shown. I was NOT dishonorably discharged.

APPLICANT'S INITIAL:______

30. FIPS CODE
31. LOP
32. ID#
33. ISSUE 1__ Yes 2__ No
34.
35. RACE/ETHNIC 1__WNH 3__HISP 5__ASN/PAC IS. 2__BNH 4__AI/AN 6__NA
36. EFFECTIVE DATE OF CLAIM
37. PROGRAM BATCH 1__TTD 3__UCX 5__W 2__CWC 4__FE 6__UC
38. Referral Union Member? 1__ Full 2__ Qual 3__No
39. Earnings since effective date? 1__ Yes 2__ No
40. Retirement/Pension Deduction? 1__ Yes 2__ No
41. Legally entitled to work in U.S.? 1__ Yes 2__ No
42. Work Search Directive Number______
43. Alpha Work Search Indicator______
From: WWYY______ Thru: WWYY______
44. VETERAN/OTHER ELIGIBLE 0__No 1__Vietnam Era 2__Other Vet
45. RECENTLY SEPARATED VET 0__No 1__Yes
46. ERI
47. CK School Wages
48. Other
49. MLFA Override____
50. MLFA PBR 0__No 1__Yes
51. EMS 650 Trigger____
60. REGISTRATION DATE
61. PA CASE #
62. DATES OF SERVICE
52. APPLICATION TYPE 1__BOTH 2__ES 3__UI
53. SELECTIVE SERVICE REGISTRATION 0__No 1__Yes 3__N/A
54. HDCP/DISABLED 0__No 1__Hdcp 3__Spec. Dis. Vet 2__Disabled Vet
55. FOOD STAMP 0__No 1__Yes
56. ECON. DISAD. 0__No 1__Yes
57. PA 0__No WIN NON-WIN
58. MSFW 0__No 1__Seasonal Farm Wrkr
59. SUMMER YOUTH 1__Yes

BEST COPY AVAILABLE