National Job Corps Study: Assessing Program Effects on Earnings for Students Achieving Key Program Milestones

June 2001

Prepared for:

U.S. Department of Labor
Employment and Training Administration
Office of Policy and Research
Room N-5637
200 Constitution Ave., NW
Washington, DC 20210

Prepared by:

R. Mark Gritz
Terry Johnson

Battelle Memorial Institute
P.O. Box 5395
4500 Sandpoint Way, NE
Seatle, WA 98105

Prime Contract No.: K-4279-3-00-80-30
Battelle Reference No. FG303039

National Job Corps Study: Assessing Program Effects on Earnings for Students Achieving Key Program Milestones

June 2001

Submitted to:

U.S. Department of Labor
Employment and Training Administration
Office of Policy and Research
Room N-5637
200 Constitution Ave., NW
Washington, DC 20210

Submitted by:

Battelle Memorial Institute
4500 Sand Point Way NE
Suite 100
Seattle, WA 98105-3949

In conjunction with:

Mathematica Policy Research, Inc.
(Prime Contractor)
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

Decision Information Resources, Inc.
(Subcontractor)
2600 Southwest Freeway
Suite 900
Houston, TX 77098

Project Officer:
Daniel Ryan

Project Director:
John Burghardt

Principal Investigators:
Terry Johnson
Charles Metcalf
Peter Schochet

This report has been produced under Contract Number K-4279-3-00-80-30 with the U.S. Department of Labor, Employment and Training Administration. The contents of the report do not necessarily reflect the views or policies of the Department of Labor, nor does mention of trade names, commercial products or organization simply endorsement of these by the U.S. Government.

ACKNOWLEDGMENTS

We would like to acknowledge the efforts of the many individuals from numerous organizations who made valuable contributions to this report. We also owe our gratitude to the numerous individuals who contributed to the earlier report, "National Job Corps Study: Job Corps Applicants' Programmatic Experiences," as this report builds on the analysis presented there.

The authors are especially indebted to the multitude of Job Corps staff who provided key data on the practices of outreach and admissions counselors and center characteristics that were critical to the development of the non-experimental estimates examined in this report. In particular, we would like to thank the nearly 500 OA counselors that completed a detailed telephone survey on the practices they follow in recruiting and screening students for the program and the 110 center directors nationwide, along with their staff, who contributed to completing a comprehensive mail survey of center characteristics. We are also grateful for the assistance of the Job Corps Data Center for providing student program records used in the analysis. Finally, we would like to recognize the many people who assisted in the design and data collection for the study.

The authors also wish to acknowledge the U.S. Department of Labor and the Job Corps national office staff who reviewed preliminary results from this analysis and helped shape the focus and format of this report. We appreciate the support and guidance of Dan Ryan throughout the study, as well as his careful review of an earlier draft of the report. We would also like to thank Richard Trigg, Mary Silva, Pete Rell, and Jenny Gallo for their time in meeting with the research team on several occasions to review preliminary findings and for their insights on the types of analyses that would be of most policy interest to the program.

Within the research team, we are particularly grateful to John Burghardt, Peter Schochet and Steve Glazerman for their thorough and thoughtful reviews of preliminary versions of this report. Their comments were extremely valuable in creating a revised document that we hope is of interest to Job Corps policy makers. We would also like to thank Fred Dong for expert data processing and analysis assistance. Finally, we would like to thank Kate Lynch for her excellent help in producing this report.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Job Corps plays a central role in federal efforts to provide employment assistance to disadvantaged youths ages 16 to 24. The program's goal is to help these individuals become "more responsible, employable, and productive citizens" by providing them with comprehensive services that include basic education, vocational skills training, counseling, and residential support. Each year, Job Corps serves more than 60,000 new enrollees at a cost of more than $1 billion. The National Job Corps Study is expected to provide Congress and program managers with the information they need to assess how well Job Corps is attaining its goal.

This report is one of a series presenting findings from the National Job Corps Study. The main impact analysis results based on the experimental design are presented in a companion report. The impact analysis report focuses on the average impacts of the program on post-program earnings and other outcomes. The analysis reported here goes beyond simple average impacts in order to provide program operators and others with information about how specific programmatic achievements contribute to observed average impacts. We apply non-experimental statistical methods to estimate the impacts on quarterly earnings during the 48-month period after application for eligible applicants who attain key program milestones, as well as for those who do not achieve those milestones. The specific milestones we examine include completion of a vocational training program and attainment of a GED while enrolled in Job Corps. The results derived from an examination of the achievement of these two milestone address key policy relevant questions because the Workforce Investment Act of 1998 directs Job Corps to focus on the outcomes of graduates, and defines graduation as either completion of a vocational training program or attainment of a GED.

To estimate the impacts of Job Corps for participants who complete vocational training or earn a GED (as well as for those participants who do not attain these milestones), one needs a way to determine what the earnings would have been for similar individuals who did not attain these milestones. The findings summarized below are based on comparisons with youths who were part of the study's randomly assigned control group of eligible applicants who were not permitted to enroll in Job Corps. We used several different econometric models and different matching approaches to develop comparison groups from within the study's control group whose experiences can serve as a benchmark for measuring impacts. An extensive literature has applied econometric models to derive non-experimental impact estimates for many programs similar to Job Corps. However, these types of models consistently failed traditional specification checks designed to test whether key underlying assumptions were met. Consequently, the findings summarized below are based entirely on matching methods, which have recently become the methodology of choice for the estimation of impacts in non-experimental settings. Matching methods have often been criticized for not being able to develop comparison groups that are matched to program group members on observed and unobserved characteristics. However, a control group created by random assignment assures that the pool of individuals from which matches will be selected include individuals with similar observed and unobserved characteristics to those of vocational completers and GED recipients, which is a key advantage in this analysis.

The results summarized below use propensity scores as a basis for developing kernel matches. A kernel matching process was performed separately by gender to ensure that the groups were equivalent on this key characteristic. Essentially, the kernel matching method identified the control group member(s) who were best matched to each program group member based on characteristics associated with the likelihood of achieving key program milestones. Individuals in the control group with propensity scores more similar to the propensity scores of each program group member were given a greater weight in establishing the benchmark for that program group member and those in the control group with dissimilar propensity scores were given lesser weight. The outcomes of the control group members matched to each program group member in this way provides our measure of the earnings each program group member would have earned if the program had not been available.

Using the kernel matches we estimated the impacts on long-term earnings for Job Corps participants who did and did not achieve two key milestones: (1) vocational completion; and (2) GED attainment. In addition, to help assess the fit of the matching methods, we also present impact estimates for program participants and members of the program group who did not participate in Job Corps. The highlights of our findings are summarized below:

- The overall impact estimates developed from the kernel matches closely track the experimental impacts of approximately $15-$20 per week for all applicants for quarters 11-16 after random assignment. Despite this similarity to the experimental impact estimates, it should not be interpreted as strong evidence of the validity of the matching methodology since the comparison group is drawn from the entire control group and essentially involves only a re-weighting of the outcomes of control group members.

- The kernel matches yield impact estimates for those who enroll in the program (i.e., participants) that are slightly lower than the experimental results. Specifically, the experimental findings correspond to a $20-$25 per week impact for participants over these quarters, as compared to an estimated $15-$20 per week impact using the kernel matches. This indicates that the overall impact estimate based on the kernel matches for non-participants is approximately $5 per week.

- Among participants, we find that nearly all of the positive program impacts on earnings accrue to those who accomplish one of the two major milestones – completing a vocation or attaining a GED. In contrast, students who participate but fail to complete a vocation or earn a GED derive no benefit from Job Corps. This important finding lends support to the recent emphasis the Job Corps program has placed on ensuring that students graduate, in response to the Workforce Investment Act of 1998. At the same time, however, the finding that students who enroll in Job Corps but do not complete their vocational training have smaller impacts than non-participants raises questions about the reliability of the estimates based on the kernel matches.

- The estimated impact for students who complete their vocational programs becomes positive after the sixth quarter, reaches $40 per week by quarter 11 and remains between $40-$50 per week through quarter 16. Students who do not complete their vocations are estimated to have slightly lower earnings than their matched comparison group during the period they are most likely enrolled in Job Corps and then have earnings that are nearly identical to their matched comparison group throughout the remainder of the observation period.

- Similar to the findings for vocational completion, nearly all of the positive impacts for students who did not have a GED at entry are estimated to accrue to participants who earned a GED. Among students without a GED at entry, non-recipients are estimated to have an initial negative impact during the period they are most likely enrolled in Job Corps and then have earnings that are nearly identical to their matched comparison group. Among students without a GED at entry, the estimated impact for GED recipients becomes positive in the fifth quarter, reaches about $60 per week during quarter 11 and remains between $60-$70 per quarter through quarter 16.

- We also estimated the impacts of achieving specific program milestones separately by age and found that the general conclusions described above hold for each of the age groups (16-17, 18-19, 20+). Although this result might appear to be inconsistent with the differences across age groups found in the experimental impact estimates for participants, the consistency of the patterns across age groups of the estimated earnings impacts for participants who achieve a program milestone provides some additional confidence in these non-experimental findings. Specifically, combining the non-experimental estimates for non-participants and participants who do not achieve a milestone with the estimates for those achieving the milestone yield the same age pattern in overall earnings impacts as obtained from the experimental design.

In interpreting the policy implications of these findings, it is important to recognize the questions these findings address and those they do not. For example, although the results indicate no impacts for non-graduates, this should not be interpreted as evidence that Job Corps should not serve students who do not complete the program. This is because of the inherent difficulty of determining a priori which students will complete the program and graduate and which students will not. For example, although the propensity score models help distinguish participants that achieve program milestones from those who do not, these models are not well suited to identifying whether a specific individual student will succeed or fail in the program. In addition, although we believe the findings provide reasonable evidence of the effects of Job Corps for those students who completed their vocational training and those who did not, they cannot be interpreted as representing what would happen if more students were turned from non-completers into completers.

In understanding the policy implications, it is also important to recognize the uncertainty surrounding the specific impact estimates. Impact findings based on non-experimental methods – such as those necessary to measure impacts for students who did or did not achieve key milestones – are forced to rely on inherently untestable assumptions about the relationships of observed and unobserved factors to program participation and post-program earnings. This inherent shortcoming of non-experimental methods always raises the possibility that the findings may not present an accurate or reliable estimate of a program's impact because key assumptions underlying these methods may or may not be satisfied.

It is also important to recognize that these findings do not disentangle the variety of mechanisms through which Job Corps can improve the outcomes for participants. Specifically, because students who complete key milestones typically remain in the program for a long time and receive extensive residential services (including social skills training), the impacts we have attributed to completing a vocation or receiving a GED may also simply result from more time in the program and greater exposure to the other experiences that Job Corps offers. Also, by estimating impacts separately for vocational completion and for GED attainment, we have not examined the effects of achieving one milestone but not the other or the effects of achieving both. Yet, our inability to fully disentangle the effects of completing the program from the effects of greater exposure to the program does not materially affect the importance of the main finding: Job Corps program practices that promote longer retention to facilitate achieving completion of vocational training or attainment of a GED or high school diploma are likely to be beneficial.

I. INTRODUCTION

Job Corps plays a central role in federal efforts to provide employment assistance to disadvantaged youths ages 16 to 24. The program's goal is to help these individuals become "more responsible, employable, and productive citizens" by providing them with comprehensive services that include basic education, vocational skills training, counseling, and residential support. Each year, Job Corps serves more than 60,000 new enrollees at a cost of more than $1 billion.

The National Job Corps Study, funded by the U.S. Department of Labor (DOL), was designed to provide information about the effectiveness of Job Corps in attaining its goal.[1] The central feature of the study was the random assignment of all youths found eligible for Job Corps to either a program group or a control group. Program group members were permitted to enroll in Job Corps, and control group members were not (although they could enroll in other training or education programs). The research sample for the study consists of approximately 9,400 program group members and 6,000 control group members randomly selected from among nearly 81,000 eligible applicants nationwide. Sample intake occurred between November 1994 and February 1996.

The national study consists of three major components: (1) an impact analysis, (2) a process analysis, and (3) a benefit-cost analysis. To estimate the overall impact of Job Corps, the main impact analysis exploits the random assignment design and calculates the mean difference in earnings and other outcome measures between the program group and the control group. Although the average impact of the program on earnings is a critical component of the overall evaluation and of the benefit-cost analysis,

[1] The study is being conducted by Mathematica Policy Research, Inc. (MPR) and its subcontractors, Battelle Memorial Institute and Decision Information Resources, Inc.

it provides no information to program operators regarding the specific programmatic elements that are responsible for the impacts. To address this gap, in this report, we apply various non-experimental statistical methods to shed light on the effects of Job Corps for applicants who achieve specific program milestones. In particular, we estimate the impacts on earnings for applicants that achieve one (or more) of two specific milestones: (1) complete a vocational training program; or (2) obtain a GED or high school diploma while enrolled in Job Corps. Results for these two milestones address key policy-relevant questions because the Workforce Investment Act of 1998 directs Job Corps to focus on the outcomes of graduates, and defines graduation as either completion of a vocational training program or GED attainment.

In the remainder of this chapter, we provide additional background information concerning the Job Corps program and the National Job Corps Study. We also provide additional details concerning how this report fits into the overall study and the specific objectives and policy issues that we address in this report. The chapter concludes with a description of the organization of the report.

A. OVERVIEW OF JOB CORPS

The Job Corps program, established by the Economic Opportunity Act of 1964, currently operates under provisions of the Workforce Investment Act of 1998. The operational structure of Job Corps is complex, with multiple levels of administrative accountability, several distinct program components, and numerous contractors and subcontractors. DOL administers Job Corps through a national office and nine regional offices. The national office establishes policy and requirements, develops curricula, and oversees major program initiatives. The regional offices procure and administer contracts and perform oversight activities, such as reviews of center performance.

Through its regional offices, DOL uses a competitive bidding process to contract out operations of the three main program elements: recruiting and screening of new students, center operations, and placement of students into jobs and other educational opportunities after they leave the program. At the time of the study, 80 centers were operated under such contracts. In addition, the U.S. Departments of Agriculture and of the Interior operated 30 centers, called Civilian Conservation Centers (CCCs), under interagency agreements with DOL.[2] Next, we briefly outline the roles of the three main program elements as they operated at the time of the study.

1. Outreach and Admissions

Recruitment and screening for Job Corps are conducted by outreach and admissions (OA) agencies, which include private nonprofit firms, private for-profit firms, state employment agencies, and Job Corps centers. These agencies provide information to the public through outreach activities (for example, by placing advertisements and making presentations at schools), screen youth to ensure that they meet the eligibility criteria, assign eligible youth to centers (when the regional office delegates this function), and arrange for their transportation to centers.

2. Job Corps Center Services

Centers are the cornerstone of Job Corps as they provide a comprehensive and intensive set of program services. The major services provided by centers include basic education, vocational

[2] Currently, 91 contract centers and 28 CCCs provide Job Corps training.

training, residential living (including training in social skills), health care and education, counseling, and job placement assistance. Services in each of these components are tailored to meet the needs of individual students.

Education. The goal of the education component is to enable students to learn as fast as their individual abilities permit. Education programs in Job Corps are individualized and self-paced and operate on an open-entry and open-exit basis. The programs include remedial education (emphasizing reading and mathematics), world of work (including consumer education), driver education, home and family living, health education, classes designed for those whose primary language is not English, and a General Educational Development (GED) program of high school equivalency for students who are academically qualified. In addition, about one-fourth of the centers can grant state-recognized high school diplomas.

Vocational Training. As with the education component, the vocational training programs are individualized, self-paced, and operate on an open-entry and open-exit basis. Each Job Corps center offers training in several vocations, typically including business and clerical occupations, health occupations, construction, culinary arts, and building and apartment maintenance. Instruction is provided by staff with occupational experience that are hired by the center, as well as by national labor and business organizations under national training contracts at many centers.

Residential Living. Residential living is the component that distinguishes Job Corps from most other publicly funded employment and training programs. From its inception in 1964, residential living has been considered a key element of the program because most students come from disadvantaged environments and it is believed they require new and more supportive surroundings to derive the maximum benefits from education and vocational training. A key part of residential living consists of formal social skills training in which all students must participate, including nonresidential students. The residential living component also includes meals, dormitory life, entertainment, sports and recreation, center government, and other related activities.

Health Care and Education. Job Corps centers also provide comprehensive health services to both residential and nonresidential students. Services include medical examinations and treatment; biochemical tests for drug use, sexually transmitted diseases, and pregnancy; immunizations; dental examinations and treatment; counseling for emotional and other mental health problems; and instruction in basic hygiene, preventive medicine, and self-care.

Counseling and Other Ancillary Services. Job Corps centers offer students a range of other supportive services including providing counselors and residential advisers. These staff help students plan their educational and vocational curricula, offer motivation, and create a supportive environment. Support services are also provided during recruitment, placement, and the transition to self-sufficiency and employment.

3. Placement

The final step in the Job Corps program is placement. This component of the program helps students find jobs in training-related occupations with prospects for long-term employment and advancement. Placement contractors may be state employment offices or private contractors, and sometimes the centers themselves perform placement activities. Placement agencies help students find jobs by providing assistance with interviewing and resume writing and services for job development and referral. They are also responsible for distributing the readjustment allowance; a stipend students receive after leaving Job Corps.

B. OVERVIEW OF THE NATIONAL JOB CORPS STUDY

The National Job Corps Study addresses six major research questions:

1. How effective is Job Corps overall at improving the employability of disadvantaged youth?

2. Does the effectiveness of Job Corps differ for youths with different personal characteristics or experiences before application to Job Corps? Do impacts vary by gender, age, the presence of children, education level, race and ethnicity, or arrest history?

3. Do program impacts differ for centers with different characteristics? Do impacts vary by CCC or center contractor type, center size, center performance level, or region?

4. Do program impacts differ for enrollees with different program experiences? Do impacts differ by residential status? Do impacts differ by programmatic accomplishments?

5. What is the Job Corps program "model," and how is this model implemented in practice?

6. Is Job Corps cost-effective?

To address these questions, the study consists of an impact analysis (Questions 1 to 4), a process analysis (Question 5), and a benefit-cost analysis (Question 6).

In this report, we focus on the second aspect of the fourth research question addressed by the impact analysis. Specifically, as stated above, this report assesses the impact of the Job Corps experience on the participants who achieve one, or both, of the key program milestones of completing a vocational program and attaining a high school credential (most commonly a GED). Findings examining the estimated impacts for residential and non-residential students, along with the analyses addressing the first two questions is reported in Schochet et al (2001) and findings addressing the third research question are presented in Burghardt et al (2001). For a description of the process analysis design and findings, see Johnson et al (1999); for a description of the benefit-cost analysis, see McConnell et al (2001).

To address the first two impact analysis questions, Schochet et al (2001) exploits the random assignment feature of the experimental study design and calculates simple differences in mean outcomes between program group members and control group members overall and for key subgroups of youths. In addition, to estimate the impacts of the Job Corps residential component, Schochet et al (2001) compare the mean outcomes of program group members who, before random assignment, were expected by their OA counselor to be assigned to a residential slot with the mean outcomes of control group members who, before random assignment, were similarly expected to be assigned by their OA counselor to a residential slot. The same approach is used by Schochet et al (2001) to estimate the impacts of the nonresidential component and by Burghardt et al (2001) to estimate the impact of attending Job Corps centers with specific characteristics.

In contrast to the estimation of impacts for the residential or nonresidential components and for centers with different sets of characteristics, the estimation of impacts for applicants who achieve different program milestones cannot rely solely on the random assignment design. The evaluation design called for the random assignment to program and control group status to be made at the time of eligibility determination and all subsequent programmatic outcomes of program group members, such as enrollment at a center and completion of a vocational training program, partly reflect the specific choices of the student, as well as program and other factors. As a result, to address this research question we must take into account the process by which applicants move through the different program components. Additional information about the specific questions we examine in this report and the types of methods used are described in the next section.

C. OBJECTIVES OF THE ANALYSIS

As described earlier, the research question motivating the analysis and results presented in this report concerns whether program impacts differ for applicants who achieve different key program milestones. Answering questions such as this is widely recognized as presenting a formidable challenge because it requires an extensive understanding of the processes determining the program experiences and accomplishments of individuals, as well as the processes that determine these individuals' labor market and other related outcomes. Specifically, to answer this question analysts must separate out the effects of the factors that determine which individuals have a particular set of program experiences from the effects of these experiences on outcomes. As such, we separate the underlying question into two research objectives:

- How do applicant personal characteristics and program operational features/practices impact key programmatic experiences or accomplishments? Specifically, what factors affect an applicant's decision to enroll in Job Corps, his/her choice of vocational programs, and programmatic achievements such as GED attainment, completion of vocational training and length of stay in the program?

- Do program effects differ for students who achieve different program milestones? Specifically, do impacts vary by vocational program completion status and GED attainment status?

In a previous report (Johnson et al 2000), we presented the results of statistical models designed to address the first of these two research objectives. Specifically, in that report, we presented results that summarized the relationships among personal characteristics, program practices, and different program milestones, including enrollment in Job Corps, completion of a vocational training program and receipt of a GED or high school diploma, among others. Although understanding which applicants attain specific program milestones provides insights into program operations, it does not address the question of whether and how the program makes a difference in the lives of the young men and women who reach these milestones. In this report, building on the results of the previous analyses and using data through 48 months after application, we use several different non-experimental statistical methods to estimate the impacts on quarterly earnings for students who achieve specific program milestones.

The analysis reported herein focuses on two specific program milestones: (1) completion of a vocational training program; and (2) receipt of a GED or high school diploma while in Job Corps (among those without a GED or high school credential at entry). These two experiences are the major achievements that define program graduation status under the Workforce Investment Act of 1998. In this report, we examine the effects of these key program milestones separately. Consequently, our findings do not address the effects of achieving one milestone but not the other, or the effects of achieving both.

The fundamental problem in estimating impacts for applicants who achieve specific program milestones is that we can observe who achieves each of these milestones only for the program group who had the option to enroll. We cannot observe them for the control group who were ineligible to enroll in Job Corps. If we knew which control group members would have accomplished these same milestones if they had been given the opportunity to participate in the program, then we could use the same experimental methods to estimate these impacts as are used to estimate the impacts for all eligible applicants to Job Corps. Because control group accomplishments cannot be known, we must rely on non-experimental approaches – econometric models and matching methods – to construct appropriate comparison groups for applicants who achieved specific programmatic milestones. We describe these non-experimental methods in detail in Chapter II.

The main findings from our analysis are that nearly all of the positive program impacts accrue to those participants who accomplish one of the two major program milestones – complete vocational training or attain a GED or high school diploma in Job Corps. As such, program practices that promote longer retention to facilitate achieving completion of vocational training or attainment of a GED or high school diploma are likely to be beneficial. Although these results have intuitive appeal and are consistent with the focus of the Workforce Investment Act of 1998, it is important to keep certain caveats in mind. First, these impact estimates correspond to the impacts for those who reached these programmatic achievements at the time they were enrolled. It is not the impact for the average participant, and does not necessarily reflect the impact we could expect from taking a person who did not quite reach the milestone and then providing additional assistance that enabled him/her to become a vocational completer or a GED recipient. Second, it is important to recognize that since the impact results reported here are based on non-experimental methods, they will be subject to extensive scrutiny and review and will not be consistently accepted by the evaluation community.

D. OVERVIEW OF THE REPORT

The remainder of this report is organized as follows. In Chapter II, we describe the non-experimental statistical approaches used to estimate the effects of the program on long-term earnings for applicants who achieve key program milestones. In that discussion, we also clarify the policy questions that each approach addresses, as well as describe the specification tests that each approach must meet in order for the impact results to be considered as valid. In Chapter III, we describe the various data sources underlying the analysis and present descriptive information on the participant samples. In Chapter IV, we present the overall estimated impacts on earnings for participants who complete their vocational training program and attain their GED, as well as for students who do not achieve these key milestones. Finally, Chapter V summarizes our findings and discusses implications for program operations.

II. METHODOLOGY FOR ESTIMATING IMPACTS FOR PARTICIPANTS WHO ACHIEVE KEY PROGRAM MILESTONES

One of the primary objectives of the National Job Corps Study is to estimate the impact of Job Corps on applicants who achieve key program milestones, including completion of a vocational program, and receipt of an academic credential (primarily the GED). Such information is essential to understand which aspects of program experiences contribute to the differences in outcomes or impacts for the young men and women who apply to Job Corps. An understanding of the potential sources of the observed impacts of the program is also needed by program operators to carry out the program's philosophy of continuous program improvement. For example, the Workforce Investment Act of 1998 increased the program's focus on vocational program completion and receipt of an academic credential. While this focus has a great deal of intuitive appeal, this policy focus was adopted without any direct evidence that such a focus would improve outcomes for program participants. Providing program operators with such evidence will permit them to modify the program to enhance the program experiences that are most likely to improve outcomes for the youth who participate in the program.

Assessing the impacts of the program for students who achieve specific program milestones on labor market earnings, as well as other outcomes, presents a number of challenges that do not arise in the other elements of the impact analysis. Whereas estimates of the overall impact of offering Job Corps to the eligible population of young men and women can rely on the random assignment of eligible applicants, estimation of the impacts of different programmatic experiences cannot rely solely on the experimental design because random assignment was made at the time of eligibility determination. Hence, all subsequent programmatic experiences reflect the specific choices of individuals assigned to the program group, as well as the specific features of the program. Moreover, the individual characteristics that determine these choices are also likely to affect the labor market and other outcomes of these individuals after the time of eligibility determination. As a result, to address this objective of the evaluation we must rely on non-experimental statistical methods to estimate the impacts of key program milestones on the labor market outcomes of Job Corps applicants.

The fundamental challenge in estimating impacts for eligible applicants with specific programmatic experiences arises because it is only possible to observe the experiences of the eligible applicants who are assigned to the program group. That is, it is impossible to observe the programmatic experiences of the eligible applicants who are assigned to the control group because they were embargoed from participating in the program. If it were possible to ascertain the programmatic experiences that all eligible applicants assigned to the control group would have had if they were not embargoed from Job Corps, standard experimental methods could be used to estimate the impact of specific programmatic experiences, such as the achievement of significant program milestones. For example, if the members of the control group who would have completed a vocation in Job Corps were identifiable, the difference in the mean outcome between the program group members who are observed to complete a vocation in Job Corps and the subset of control group members who would have completed a vocation provides a consistent estimate of the impact of vocational completion. However, without this information alternative non-experimental methods must be used to estimate these types of impacts.

The inability to ascertain the programmatic experiences control group members would have had is analogous to missing data problems. The problems arising from missing data are generally dealt with by a variety of methods including: (1) imposing assumptions about the relationship between missing information and observed data; (2) using proxies for the missing information; and, (3) using statistical or econometric methods to estimate the missing information. All three of these approaches to dealing with missing information are being used in the National Job Corps Study.

The first approach to dealing with missing information is being used in the main impact analysis to estimate the impact of Job Corps on participants. Specifically, the analysis is imposing the assumption that the observed outcomes for the program group members who do not enroll in Job Corps are equal to the outcomes of the control group members who would not have enrolled in Job Corps if they were not embargoed from the program. This approach is equivalent to assuming that the option of enrolling in Job Corps has no impact on the outcomes for those who do not participate in the program. While this assumption is plausible for no-shows, it is unreasonable to impose such an assumption for those who participate in the program but fail to complete a vocation or attain an academic credential in Job Corps. As such, this approach cannot be used to estimate the impacts of interest in this analysis.

The second approach to missing data is being used in the National Job Corps Study to estimate separate impacts of offering Job Corps to residential and non-residential eligible applicants and to estimate impacts for subsets of Job Corps centers. To estimate the impact of offering Job Corps to residential students, the outreach and admissions counselors were asked to identify the eligible applicants who they thought would likely participate in Job Corps as residential students and those who would likely participate as non-residential students prior to random assignment to the program or control group. Similarly, outreach and admissions counselors were also asked to identify the specific Job Corps center an applicant was most likely to attend prior to random assignment. In essence this approach substitutes proxy measures (i.e., the judgments of outreach and admissions counselors) for observed data (i.e., the experiences of participants in the program group), as well as the missing data for control group members and program group members who do not enroll. In principle, this approach could have been used prior to random assignment to identify program and control group members who would be most likely to achieve specific program milestones. However, because outreach and admissions counselors were unable to consistently distinguish between applicants who were going to arrive at a Job Corps center and those who would end up not participating in the program, this approach cannot be used to estimate the impacts of enrollment or any other subsequent program experiences.[3]

The third approach to dealing with missing data encompasses what are commonly referred to as non-experimental methods for deriving estimates of program impacts. There is an extensive literature examining the use of non-experimental methods to estimate program impacts and over the last 20 years a significant body of research has developed comparing experimental and non-experimental methods. This literature has focused almost exclusively on the replication of experimentally derived impact estimates by non-experimental methods rather than the application of non-experimental methods within an experimental setting, such as is the case here. However, many of the lessons learned over the last 20 years apply to this analysis. Specifically, this literature has clarified the specific evaluation questions that are answered by alternative non-experimental approaches and identified several specification checks that can be used to assess the reliability of non-experimental methods.

[3] For example, OA counselors identified approximately 99 percent of all eligible applicants as being either very likely or likely to participate in Job Corps.

This chapter describes the non-experimental methods we examined to estimate the impact for students who achieve key program milestones in Job Corps. The next section briefly discusses the various evaluation questions that can be answered with non-experimental methods and describes the different methods used to estimate impacts for vocational completers and for those who receive an academic credential (GED or high school diploma) in Job Corps. The next two sections describe the empirical specifications used to derive the various non-experimental impact estimates. The chapter concludes with a discussion of various specification checks that are used to gauge the reliability of the estimates derived from these different non-experimental methods.

A. ALTERNATIVE NON-EXPERIMENTAL METHODS AND THE EVALUATION QUESTIONS THEY ADDRESS

An extensive body of research has developed a number of alternative methods to estimate program impacts that encompass a variety of approaches other than random assignment experiments. Because of the strong connection with random assignment experiments and the evaluation of alternative program services or treatments, this body of research has typically used the jargon of "treatment groups" and phrased evaluation questions in terms of the "impacts of the treatment." Thus, for convenience of presentation and consistency with the literature, in the following sections we use "treatment group" to represent the randomly assigned program group of eligible applicants.

This body of research includes an extensive examination of the benefits and shortcomings of using non-experimental and experimental methods in different settings and it has clarified the different evaluation questions answered by the different types of non-experimental methods.[4] Among the wide-range of possible evaluation questions, this literature has focused on three measures of the direct impact of a program that are relevant to this analysis. These three different measures are:

1. The average (mean) impact of the program on those who receive program services, which is often referred to as the impact of "treatment on the treated,"

2. The average impact on a randomly selected person with specific observed characteristics from placing him/her into the program and compelling this person to receive program services; and,

3. The average impact of switching people who are on the margin (i.e., indifferent between participating and not participating) from nonparticipation to participation in the program, which is referred to as the "local average treatment effect" (LATE).

It should be noted that the National Job Corps Study does not directly answer any of these three evaluation questions. Rather, it directly measures the impact of offering the opportunity to participate in the program to disadvantaged youth who apply and are found eligible for Job Corps.[5]

All non-experimental methods are based on making comparisons between the individuals who have a particular programmatic experience and other individuals, the comparison group, who do not have the specific experience. The composition of the comparison group affects the type of non-experimental estimation methodology that can provide answers to one or more of the evaluation questions. There are two potential sources from which to draw comparison groups to estimate the impacts of the programmatic achievement of interest. One source is the control group and the other source is the members of the treatment group who do not have a particular program experience, such as no-shows.

---

[4] A more extensive discussion of the issues discussed in this section can be found in Heckman, LaLonde and Smith (1999) and the references cited therein.

[5] This impact is often referred to as the impact of the "intent to treat." As noted above, the experimental impact can be adjusted to estimate the impact of treatment on the treated by assuming the offer of the program has no effect on non-participants.

A comparison group drawn from the individuals randomly assigned to the control group naturally leads to the use of the intuitively appealing method of matching. A comparison group drawn from treatment group members who do not have a particular programmatic experience requires the use of econometric models, including selection bias correction methods and instrumental variable approaches, to obtain impact estimates for programmatic achievements.

Matching methods provide a non-experimental answer to the first evaluation question. Specifically, matching each treatment group member who has a particular experience to one or more control group members and taking the difference in the average outcomes across the treatment group and the matched controls provides a straightforward estimate of the mean impact of the programmatic experience for those who have the experience. As described in the next section, we examined two different methods of matching treatment group members with corresponding control group members to estimate the impacts for vocational completers and youth who receive an educational credential in Job Corps.

Although econometric models can be used to answer a wider range of evaluation questions, models based on selection bias correction methods have generally been used to answer the second evaluation question and instrumental variable techniques have been used to derive LATE estimates. Selection bias correction methods are primarily applied in observational studies that do not include an experimental design. Within an experimental setting, these models use only information for the treatment group to develop comparison groups and the control group is not needed to serve this role. For example, no-shows from the treatment group are used as the comparison group in using selection correction models in estimating the impact of participation in Job Corps. These methods provide answers to the second evaluation question above; that is, they provide an estimate of taking a randomly selected eligible applicant and compelling him or her to have a particular programmatic experience, such as enrolling in the program. In contrast, instrumental variable techniques applied with the same comparison group yield estimates of the impact of moving an eligible applicant who is on the margin from not having the experience to having the experience. For example, instrumental variable techniques with no-shows comprising the comparison group provide an estimate of the impact of moving eligible applicants who were on the verge of enrolling at a Job Corps center but decided not to enroll, and compelling them to enroll. As described in Section C, we examined both selection bias correction models and instrumental variable techniques to estimate the impact of Job Corps on participants.

B. THE METHOD OF MATCHING

The method of matching provides an intuitively appealing approach to estimate impacts of programmatic experiences and achievements, particularly in the context of an experiment where a comparison group can be drawn from a randomly assigned control group. Specifically, one of the major shortcomings of matching methods as a non-experimental approach for estimating program impacts is the potential that the comparison group does not share the characteristics of the program participants. In an experimental setting, however, the control group, by construction, has the same characteristics as the treatment group, which eliminates this shortcoming of the method of matching. For example, a comparison group drawn from the general population of disadvantaged youth might not include anyone who would have completed a vocational training program in Job Corps if they had participated in the program. In contrast, as a result of random assignment the control group consists of individuals who are virtually identical to the treatment group, including similar individuals who would have completed a vocational training program if they were not embargoed from the program.

To formalize the method of matching in this context, assume that each treatment group member has an identical control group member who would have the same programmatic milestones or experiences, let Y_i represent an outcome for treatment group member *i* and let Y_i^C represent the outcome for the control group member who is identical to this individual. Further, let $i \in E$ represent the subset of treatment group members who have a particular programmatic experience and let N_E represent the number of treatment group members who have this experience. The experimental estimate of the impact of programmatic experience E is given by

$$\Delta_E = \frac{1}{N_E} \sum_{i \in E} \left(Y_i - Y_i^C \right),$$

which represents the mean impact of treatment on the treated for those who experience E.

In practice, the quantity Y_i^C is not observable and the method of matching replaces this unobservable quantity with an estimate based on a group of comparable persons to obtain estimates of the impact of a programmatic experience. Although in principle a different group of comparable persons can be used for each treatment group member, generally all members of the comparison group are included in deriving an estimate for Y_i^C, which will be denoted by $\overline{Y}_i^C$. Such an estimation approach can be represented by a weighted average of the outcomes of all control group members. To illustrate this approach, let Y_j^C represent the observed outcome for person j in the control group and let $W(i, j)$ be the weight placed on control group member j in forming the matched comparison outcome for person *i* from the treatment group. Using this notation, $\overline{Y}_i^C$ can be written as $\overline{Y}_i^C = \sum_{j=1}^{N_C} W(i, j) Y_j^C$, where N_C represents the number of control group members, $0 \le W(i, j) \le 1$, and $\sum_{j=1}^{N_C} W(i, j) = 1$.

Substituting this estimate for the unobserved Y_i^C yields an estimate of the impact of programmatic experience E of

$$\Delta_E = \frac{1}{N_E}\sum_{i\in E}\left(Y_i - \overline{Y}_i^C\right) = \frac{1}{N_E}\sum_{i\in E}\left(Y_i - \sum_{j=1}^{N_c} W(i,j)Y_j^C\right).$$

There are a number of alternative matching schemes to estimate $\overline{Y}_i^C$ that use different approaches to calculating $W(i, j)$.[6] Two of the most widely-used methods are nearest-neighbor matching and kernel matching. The nearest-neighbor matching estimator sets $W(i,j)=1$ for the one control group member who is the most similar to treatment group member i and $W(i,j)=0$ for all other j.[7] Kernel matching typically uses a standard distribution function to weight each control group member differently for each treatment group member. To illustrate the kernel matching approach, let $D(i,j)$ represent the difference between treatment group member i and control group member j such that $D(i,j)=0$ when the two individuals are identical. The kernel matching method sets

$$W(i,j) = \frac{K[D(i,j)]}{\sum_{j=1}^{N_C} K[D(i,j)]},$$

where K is a kernel function. This approach represents a smooth method that reuses and weights the control group members' outcomes for all treatment group members. In the results presented below, K is a normal distribution function with a very small variance so that control group members who are very similar to a treatment group member receive a relatively large weight and control group members that are different receive a very small weight.

---

6 For a survey of alternative matching schemes see Heckman, Ichimura and Todd (1997).

The final element of the matching method is the specification of the distance function $D(i, j)$. In principle, the distance function can incorporate all of the observed characteristics of treatment and control group members at the time of random assignment. However, such an all encompassing approach creates a problem of such large dimensionality that the matching process can become very inefficient. One approach to reducing the dimensionality of the matching process that has been widely used is to specify the distance function in terms of propensity scores or the probability of experiencing a particular event. For example, in the analysis below the propensity scores are the predicted probability of enrolling in Job Corps and, conditional on enrollment, the predicted probability of completing a vocation and the predicted probability of receiving a high school credential in Job Corps.

To further clarify the specification used in this analysis, let $P_i^T(E)$ represent the predicted probability that treatment group member *i* will have experience *E* and let $P_j^C(E)$ represent the predicted probability that control group member *j* would have had experience *E* if he/she were not embargoed from participating in Job Corps. These predicted probabilities or propensity scores depend on a large number of personal characteristics of the eligible applicants in both the treatment and control group as measured at the time of random assignment.[8] Using this notation, the nearest-neighbor matching method weights are defined by

$$W(i,j) = \begin{cases} 1 \text{ if } j = \underset{j \in (1,\ldots,N_C)}{Min} \left| P_i^T(E) - P_j^C(E) \right| \\ 0 \quad \text{Otherwise} \end{cases}$$

and for the kernel matches the weights are defined by

---

[7] Nearest-neighbor matches are implemented with replacement so an individual control group member can be matched to more than one treatment group member.

[8] The specification of these propensity scores is presented in Johnson et al (2000).

$$W(i,j) = \exp\left[-\left(P_j^C(E) - P_i^T(E)\right)^2 \Big/ 2s^2 \right],$$

which consists of the kernel of a normal probability density function with mean $P_i^T(E)$ and variance s^2. As noted above, in the analysis presented below s^2 is set to a very small number that differs for each of the three propensity scores to place more weight on control group members with propensity scores closest to treatment group member *i* and essentially setting the weight to zero for control group members with vastly different propensity scores. Finally, the matching methods used for both the nearest-neighbor and kernel matches stratify both the treatment and control groups into separate strata based on gender before calculating the weights.

To develop the nearest-neighbor and kernel matches we first calculated three different propensity scores: (1) the propensity to enroll in Job Corps; (2) the propensity to complete a vocational training program conditional on enrollment; and (3) the propensity to attain a GED or a high school diploma conditional on enrollment. These propensity scores are then used in the weight functions for the nearest-neighbor and kernel matches. Estimates for the impact of vocational completion use the propensity score for completing a vocation in these formulas. As described in the next section, this propensity score is calculated as the product of the predicted probability of enrollment and the predicted probability of completing a vocation conditional on enrollment. Similarly, the propensity scores used for estimating the impact of receiving an educational credential in the program are also calculated as the product of the predicted probability of enrollment and the predicted probability of receiving an educational credential in Job Corps conditional on enrollment.[9] Two different sets of impact estimates (i.e., Δ_E's) based off of the nearest-neighbor and kernel matching methods described above are calculated to estimate the impact for vocational completers and for recipients of an educational credential in Job Corps.

C. ECONOMETRIC MODELS

In addition to the method of matching, two alternative econometric modeling approaches were examined to estimate the impact of Job Corps on vocational graduates and recipients of high school credentials. The first approach applies standard selection bias correction models such as those formulated by Heckman and the second applies instrumental variable methods such as those formulated by Lee and others. As noted above, the first approach provides an answer to the evaluation question of the impact of Job Corps on a randomly selected eligible applicant and the second provides an estimate of the impact of the program on applicants who are on the margin of either enrolling, completing a vocation or receiving a high school credential in the program.[10]

Conventional econometric selection bias correction models assume that outcomes for those who have a programmatic experience and those who do not can be represented by distinct relationships between the outcome and individuals' characteristics and other variables. To formalize this approach, consider the case of two possible regimes: one characterizing outcomes under the regime where individuals participate in a program; and, the second summarizing outcomes under the regime where

9 Although not reported here, results from using a kernel match with weights that averaged the difference between separate propensity scores for enrollment and propensity scores for vocational completion conditional on enrollment yielded virtually identical results to those using only the combined propensity score.

10 Sample selection models can also be used to estimate the impact of treatment on the treated and LATE.

individuals do not participate in a program. Let Y^1 represent an outcome under the regime where individuals participate in Job Corps, Y^0 represent an outcome under the regime where individuals do not participate in Job Corps, and X represent the observed characteristics of individuals that in part determine the outcome under both regimes. In its simplest form, the two equations are specified as:

$$Y^1 = X\beta_1 + u_1 \text{ and } Y^0 = X\beta_0 + u_0 ,$$

where for simplicity it is assumed that $E(u_1 \mid X) = 0$ and $E(u_0 \mid X) = 0$.

Obviously Y^1 and Y^0 cannot be observed for the same person. Further, it is assumed that individuals self select into either regime 1 or regime 0. In addition, the selection process is postulated to depend upon a set of measured variables represented by Z, which can include individual characteristics and program characteristics, and the selection equation is given by

$$IN = Z\gamma + v ,$$

where $E(v \mid Z) = 0$, v is independent of Z and v follows a known distribution function. This index function is used to define an indicator for the regime an individual selects, represented by D, such that $D = 1$ if $IN \geq 0$ and $D = 0$ if $IN < 0$.

The final element of the selection bias model is the assumption that u_1, u_0 and v are dependent, or related to each other. In its simplest form the selection bias correction model assumes that these variables are linearly related such that that $u_1 = \alpha_1 v$ and $u_0 = \alpha_0 v$. This assumption implies that $E(u_1 \mid X, D) \neq 0$ and $E(u_0 \mid X, D) \neq 0$.

In this conventional framework, the prototypical selection correction regression model for the observed outcomes is given by

$$Y = DY^1 + (1 - D)Y^0 = X\beta + \delta D + Du_1 + (1 - D)u_0 ,$$

where it is assumed that $\beta_1 = \beta_0$ except for the coefficient on the intercept. Bias arises in estimating this equation for two reasons. First, the composite error term $Du_1 + (1-D)u_0$ is correlated with one of the right hand side variables (*D*). Second, the composite error term does not have an expected value of zero conditional on *X* and *D* (i.e., $E[Du_1 + (1-D)u_0 \mid X, D] \neq 0$). These sources of bias can be removed by incorporating selection terms into a regression model of the form

$$Y = X\beta + \delta D + DE(u_1 \mid X, D=1) + (1-D)E(u_0 \mid X, D=0) + e,$$

where *e* denotes an error term that is uncorrelated with the right hand side variables and $E[e \mid X, D] = 0$. Under the assumption that v follows a known distribution and that v is linearly related to u_1 and u_0, the expected values in the above equation can be shown to depend upon the index function *IN*.[11] Replacing the unknown parameters in the index function with estimated quantities, the above regression model will yield estimates of β and δ that are unbiased as long as all of the assumptions specified above are satisfied. Estimates of δ based on the assumption that v has a logistic distribution are presented in the appendix. As noted above, these estimates represent the impact of Job Corps on a randomly selected eligible applicant who is compelled to enroll in the program.[12]

Conventional instrumental variable models also estimate the regression model given by

$$Y = DY^1 + (1-D)Y^0 = X\beta + \delta D + Du_1 + (1-D)u_0.$$

Instead of imposing specific functional forms for the distribution of v and the relationship between u_1, u_0 and v, this approach uses the variables included in *Z*, or functions of these variables, as instrumental

[11] These models are also referred to as “index sufficient” methods.

[12] An estimate of the impact of treatment on the treated can be derived by adding an estimate of $E(u_1 - u_0 \mid X, D=1)$ to the estimate of δ derived from the regression model. For details of this calculation see Heckman, LaLonde and Smith (1999).

variables for D and requires the assumption that *Z* is independent of u_0 and $u_1 - u_0$ given *X* and *D*. In addition, this method requires at least one variable included in *Z* must be excluded from *X* and that the index function *IN* must be a non-trivial function of these excluded variables.

Applying standard instrumental variable methods to the above equation yields estimates of d. However, the interpretation of this estimate varies depending on the choice of the instruments included in *Z* and whether u_1 and u_0 are independent of *D* given *X* and *Z*. If changes in *D* (i.e., individuals switching from "0" to "1") are a consequence of a change in a *Z* variable that measures characteristics of the program, d can be interpreted as a local area treatment effect of a marginal change in program operations on those eligible applicants who are induced to switch from being a non-participant to a participant as a result of the policy change. Alternatively, if the variables included in *Z* that are excluded from X remove all of the dependence between *D* and u_1 and u_0, the application of this method yields an estimate of the

effect of the program on a randomly chosen person with characteristics *X*. This latter value corresponds to the impact estimate resulting from the application of the selection bias correction model described above under the assumption that the program has the same impact on all individuals. Finally, if the variables in *Z* do not affect the relationship between $u_1 - u_0$ and the variables in (*X*, *D*), under additional conditions this method yields an estimate of the impact of treatment on the treated.

The application of instrumental variables reported in the appendix follows the approach first specified by Lee (1983) and uses the predicted probability of arrival at a Job Corps center for all treatment group members as the primary instrumental variable to estimate the impact of programmatic experiences. The key variables used in developing the predicted probabilities that are excluded from the variables included in X consist of a variety of program characteristics, which are primarily related to outreach and admissions practices. As such, the instrumental variable results most closely estimate the LATE impact of changes in the outreach and admissions component for the eligible applicants who are on the margin of enrolling at a center.

Finally, in contrast to the matching methods described above that use the control group to develop comparison groups, both econometric modeling approaches rely on comparison groups drawn from the program group members who do not have a specific program experience. Whereas the existence of a control group circumvented many of the shortcomings of matching methods that arise in their application outside of an experimental setting, the availability of the control group does not alleviate any of the potential shortcomings of the econometric models. Specifically, the selection bias correction models are still dependent upon the validity of the functional form assumptions required to estimate these models and the instrumental variable models must impose the same exclusion restrictions required to identify the model. As such, the potential biases in the estimated impacts that have dominated the literature comparing experimental and non-experimental methods apply to the use of these methods here and it is essential that these models be subjected to specification checks to assess the extent to which the models meet the conditions required to yield unbiased estimates of program impacts.

D. SPECIFICATION CHECKS FOR ALTERNATIVE METHODS

In addition to investigating the extent to which non-experimental methods can replicate experimental results, another primary focus of program evaluation researchers over the last 20 years has been in identifying methods to determine when non-experimental methods can be applied to obtain valid impact estimates. Much of this research has described various specification checks that can be used to assess the extent to which non-experimental methods yield accurate answers to one or more of the evaluation questions discussed above. Much of this literature has focused on specification checks for selection correction bias models and instrumental variable approaches. To date, very little attention has been given the development of specification checks for the method of matching. Moreover, this literature almost solely addresses specification checks in the context of observational studies and does not examine these types of checks within an experimental setting. This section briefly describes the different specification checks used to examine the appropriateness of the matching, selection bias correction and instrumental variable methods used to estimate the impacts of programmatic achievements.

Assessments of the matching methods have primarily focused on examining the extent to which the observed characteristics of the treatment or program group overlap with the observed characteristics of the comparison group. As noted above, the availability of the control group circumvents this primary concern about this method. Specifically, the distributions of the observed characteristics of the two groups are identical because of random assignment at the time of eligibility determination. With the availability of an identical, at least in the statistical sense, individual for each treatment group member the key element of matching methods within an experimental setting is in identifying this individual from the entire control group. Without a well-specified statistical test, it necessary to rely on less formal checks of the reasonableness of the matches.

Three specification checks are used to examine the reasonableness of the nearest-neighbor and kernel matches based on the propensity score for enrollment at a Job Corps center and an additional check is used for the nearest-neighbor matches. The first specification check used for both types of matches compares the estimated impact on eligible applicants by comparing the experimental estimate to the estimate derived from each of the matching methods for the entire program group. Although this is a weak specification check, if the different estimates are virtually identical it provides evidence that the weights used in each of the matching methods yield a reasonable combination of information from the control group to estimate $\overline{Y}_i^C$. The second specification check compares the estimated impact for participants derived from the "no-show adjusted" experimental impact estimate and the matching method estimates using only the program group members who enrolled at a Job Corps center. Again, to the extent the estimates from the two matching methods replicate the experimental estimates there is evidence that the matches are weighting the appropriate control group members in estimating $\overline{Y}_i^C$. The third specification check, which is very closely related to the second, estimates the impact of offering the opportunity to enroll in Job Corps for the program group members who did not take advantage of this opportunity. Under the assumption that random assignment did not alter the application process for Job Corps and that being embargoed from Job Corps did not change the behavior of the control group members who would not have enrolled in the program if they had the option, the estimated impact for the program group who did not enroll should be zero. If these conditions hold, then deviations of the matching method estimated impacts for no-shows away from zero would indicate that the weights were not selecting the appropriate control group members in constructing $\overline{Y}_i^C$. A fourth specification check that only applies to the nearest-neighbor matches and is closely related to the first check is the number of times that each control group member is matched to a program group member. The nearest-neighbor method does not exclude a control group member once he or she is matched and because the control group has fewer members than the program group we expect each control group member to be matched to more than once. However, the extent to which control group members are never matched or that they are selected a very large number of times would indicate that the nearest-neighbor matching approach was not selecting the appropriate controls.[13]

The results from applying these four specification checks for the two matching methods are presented in the appendix. As described there, the results show a slight preference for the kernel method over the nearest-neighbor approach. In summary, the application of the first three specification checks yield very similar results with both methods tracking the experimental estimates very closely. However, as expected, the kernel approach provides estimates of $\overline{Y}_i^C$ that are less variable than the nearest-neighbor approach that relies on a single match. Based on this, all of the results presented in Chapter IV are based on the kernel matching method.

The specification tests developed for selection bias correction models and instrumental variable methods in observational studies provide the foundation for specification checks of these approaches within an experimental setting. Whereas the literature has focused on rather complex statistical tests of the assumptions required for selection bias correction and instrumental variable models to yield unbiased estimates, the availability of a control group facilitates the types of specification checks that can be performed for these two econometric modeling approaches in an experimental context. Specifically, the fundamental assumption underlying these econometric models is that the equation characterizing outcomes under the regime where an individual does not participate in a program applies regardless of whether individuals had the option to participate or not. That is, the structural outcome equation for Y^0 applies equally to program group members who decided not enroll in the program and the entire control group that was embargoed from participation.

13 Although not directly applicable to the kernel method, the selection of a very small number for σ^2 results in this check also providing some indirect evidence for this method.

To formalize this specification check, consider the regression model for regime "0" separately for the control group and the no-show program group. For the control group the structural outcome equation

$$Y^0 = X\beta_0 + u_0$$

applies directly because the conditions that u_0 is independent of X and $E(u_0 \mid X) = 0$ hold for this group. For the program group no-shows, this structural equation must take into account the fact that $E(u_0 \mid X, D = 0) \neq 0$ because of the selection process determining the composition of this group. Taking this fact into account, the structural outcome equation can be rewritten as

$$Y^0 = X\beta_0 + E(u_0 \mid X, D = 0) + \varepsilon ,$$

where ε is uncorrelated with X and D and $E(\varepsilon \mid X, D = 0) = 0$. Combining these two equations the structural outcome equation for the combined control and program no-show group can be written as

$$Y^0 = X\beta_0 + TE(u_0 \mid X, D = 0) + (1 - T)u_0 + T\varepsilon ,$$

where T is an indicator variable representing random assignment to the treatment group.

Introducing a parameter to measure any systematic difference between controls and no-shows (denoted by a), a regression model corresponding to this structural equation is given by

$$Y = Xb + aT + TE(u_0 \mid X, D = 0) + e ,$$

where the composite error term has a mean of 0 and is uncorrelated with X, $TE(u_0 \mid X, D = 0)$, and--because of random assignment—with T. If the assumptions underlying the selection correction model are appropriate, this regression should yield estimates of a that are insignificantly different from 0. In other words, after accounting for the selection process, treatment group members who do not enroll in the program should on average have the same outcomes as the entire control group.

A similar specification check is also applied to the instrumental variable approach used to estimate the impact of Job Corps on participants. To formalize this specification check, we rewrite the structural outcome equation specified above as

$$Y^0 = X\boldsymbol{b}_0 + (1-T)u_0 + T\left[E(u_0 \mid X, D=0) + \boldsymbol{e}\right] = X\boldsymbol{b}_0 + \boldsymbol{w}$$

and the regression equation as

$$Y = Xb + aT + w\,.$$

Although eligible applicants were randomly assigned to the program and control group, T is correlated with the composite error term in this equation because only the subset of program group members with $D = 0$ are included in the regression. If the conditions used in estimating the instrumental variable impacts described in Section C are met, instrumental variable methods applied to this equation using the same instrument for the treatment group no-shows should account for this correlation. Hence, estimates of a resulting from the application of instrumental variables to this equation with observations for all control group members and program group members who did not enroll in Job Corps that are insignificantly different from 0 provides evidence supporting this specification.

Results from the application of these specification checks for the two alternative econometric models are also presented in the appendix. As shown in the appendix, estimates of the parameter a for both the selection correction bias models and the instrumental variable approach suggest that neither model adequately accounts for the selection process determining participation in Job Corps. In short, the estimates of a presented in the appendix are very unstable and at times are quite large. Although many of the estimates of this parameter are not statistically different from 0 because of large standard errors, the magnitude and variability of the estimates raises serious doubts about the appropriateness of these methods to yield reasonable estimates of the impacts of enrollment in Job Corps, not to mention the subsequent achievements of vocational completion and receipt of an academic credential. As such, the impact estimates derived from the application of these methods are presented only in the appendix and are not discussed below in Chapter IV.

III. DATA SOURCES AND OUTCOME MEASURES

Our analysis of the impacts on post-program outcomes for students who achieve specific program milestones draws on multiple data sources. In this chapter we briefly describe the outcome measures, analysis samples and data sources used in the analysis. The chapter concludes with a description of the characteristics of the participant samples that is helpful in understanding the results of the impact results presented in Chapter IV.

A. OUTCOME MEASURES

The overall impact analysis was designed to examine five major types of outcome measures: (1) employment and earnings; (2) education and training; (3) dependence on welfare and other public transfers; (4) antisocial behavior, such as arrests, crimes committed by and against sample members, and alcohol and drug use; and (5) family formation and childbearing. The primary source of data for these outcome measures are interviews conducted with sample members at intake (as soon as possible after random assignment), and again at 12, 30 and 48 months after intake. Interviews are conducted by telephone, with in-person follow-up of sample members who could not be interviewed by telephone.

In this report, to examine the impacts for students who achieve specific program milestones and to improve our understanding of the strengths and weaknesses of alternative non-experimental statistical methodologies, we focus entirely on a single outcome measure, average weekly earnings. We focus on earnings because it is a key summary indicator of the quality of an applicant's post-program employment experiences and because the vocational training and educational accomplishments obtained within Job Corps are intended to improve a student's subsequent earnings. We also focus on earnings because it is a critical element of the overall benefit-cost analysis. Finally, we focus on earnings because it was the primary outcome measure available to previous evaluations of alternative non-experimental impact methods.

To provide as comprehensive a picture of the earnings patterns of all applicants as possible, the analysis examines an average weekly earnings series for the first 16 quarters following random assignment. The earnings values are constructed from employment history data obtained through the interviews with sample members, with the exact same earnings series used for this analysis as in the main impact analysis. For a description of how the quarterly earnings measures were constructed, including methods used to impute missing data, see Schochet (2001).

B. ANALYSIS SAMPLES

To be included in the analysis reported here, only two criteria come into play. First, all of the results are restricted to the subset of sample members who completed a 48-month interview. The analysis incorporates weights to adjust both for the initial sample design and for potential differences in non-response to the 48-month interview.[14]

For most of the analyses reported here, the restriction to having completed the 48-month interview is the only applicable sample restriction. However, in our analysis of the impacts of obtaining a GED or high school diploma while in Job Corps, it was necessary to restrict the analysis to those applicants that did not have either of these academic credentials at application. To implement this restriction, we reviewed both the baseline interview data and the program

[14] For a description of the weights, see Schochet (2001).

administrative data, and excluded from the analysis all sample members who reported that they had a GED or high school credential at entry in either data source.

C. DATA SOURCES

The data used in this report were obtained from three primary sources. First, as indicated above, the average weekly earnings measures were calculated from the follow-up survey data. In addition, the personal background characteristics that were included as independent variables in the selection models were obtained from the baseline survey. Second, the programmatic achievements of interest – enrollment in Job Corps, completion of a vocational training program, and receipt of a GED while in Job Corps – were obtained from the Job Corps Student Pay and Allotment Management Information System (SPAMIS). Finally, the propensity scores used to match program group members who reached certain program milestones with control group members who would likely have reached the same milestones if given the opportunity to enroll in Job Corps, were developed from the statistical models of program experiences reported in Johnson et al (2000). Below, we briefly provide additional details on how these propensity scores were derived.

The propensity scores were derived from statistical models of the likelihood of enrolling in a Job Corps center, the likelihood of enrolling and completing a vocational program, and the likelihood of enrolling and completing a GED while in Job Corps among those who did not have a high school credential at the time of application. Simple binary logit models were used to estimate the likelihood of enrolling in Job Corps and, conditional on enrollment, the likelihood of completing a vocational program and of attainment of a GED. The models were estimated using the treatment group only, separately for each of the three main applicant age groups (16-17, 18-19, and 20-24). The models included a wide range of independent variables, including applicant personal background characteristics, OA practices, OA counselor characteristics and center characteristics.[15] The coefficients from these models were then used to calculate the propensity scores for treatment and control group members using information on the intended center. For additional details, see Johnson et al (2000).

D. PARTICIPANT SAMPLE CHARACTERISTICS

The impact results described in Chapter IV are presented separately for vocational completers and for GED recipients. As such, the results do not address the effects of achieving one milestone, but not the other, or of achieving both milestones. Moreover, because students who complete their vocations and/or attain a GED typically remain in the program for a long time, it is difficult to disentangle the effects of reaching specific milestones from the effects of greater exposure to the program. In this section, we present simple descriptive information on the characteristics of the samples of students that achieve certain milestones to provide insights concerning the potential overlap of these samples.

In Table III-1, we provide some background information on the characteristics of the Job Corps participants that achieved specific program milestones. This table is organized into six columns based on whether the participant had a GED or high school diploma at application and the possible program milestones that can be achieved for each GED status. In the first row of Table III-1, we show the composition of the total sample. For example, this shows that 21.5% of

15 Information on OA practices and OA counselor characteristics were obtained from a telephone survey of OA counselors in all OA agency offices nationwide that were operating at the time of sample intake for the study. Center characteristics and center operating practices that might affect students' programmatic achievements were obtained from a mail survey of all Job Corps centers that were in operation at the time of sample intake. Additional information on these two surveys can be found in Johnson et al (1999).

TABLE III-1

Characteristics of Job Corps Enrollees who Achieve Specific Program Milestones

	Students Without GED or High School Diploma at Entry				Students With GED or High School Diploma at Entry	
Program Milestone	GED and Voc. Comp.	GED Only	Voc. Comp. Only	Neither GED or Voc. Comp.	Voc. Comp.	Not Voc. Comp.
With Achievement	18.5%	4.9%	15.3%	39.8%	8.4%	13.1%
Background Characteristic	Characteristics Among Those With Achievement					
Age 16-17	48.5%	53.4%	56.8%	58.4%	4.8%	4.9%
Age 18-19	32.1%	30.7%	26.2%	27.4%	37.1%	40.1%
Male	59.9%	62.1%	60.6%	61.8%	50.9%	51.8%
Black	39.3%	40.8%	58.6%	53.5%	41.4%	41.7%
Hispanic	20.6%	16.0%	15.8%	15.8%	16.9%	15.5%
White	32.9%	38.2%	16.7%	23.1%	34.3%	36.9%
Highest Grade Completed-11th	26.2%	22.9%	18.6%	18.0%	3.8%	3.9%
Bad Health	9.8%	12.8%	12.2%	15.1%	10.5%	9.8%
Received Welfare	52.5%	49.9%	58.7%	62.9%	54.5%	51.2%
Lived in Public Housing	18.0%	16.0%	24.7%	23.6%	17.0%	13.2%
Ever Arrested	24.1%	30.8%	20.7%	28.6%	15.5%	18.4%
Prior Drug Use	33.6%	44.1%	25.5%	36.0%	23.1%	25.1%
Prior Work Experience	81.2%	83.9%	69.1%	75.0%	90.5%	92.5%
Employed 9-12 Months in Prior Yr.	17.0%	15.0%	11.7%	11.5%	25.3%	24.2%
Had Child	12.6%	16.7%	13.9%	16.3%	19.2%	20.2%
Average Paid Days on Center	336	124	309	69	320	63

all participants had a GED or high school application at entry and that about 40% of those (8.4/21.5) were vocational completers. Of the 78.5% of all participants who did not have a GED or high school diploma at entry, we see that approximately one-half (39.8/78.5) neither completed their vocational training program nor received a GED. More importantly, this table reveals the significant overlap in the groups of students that achieve these key milestones. For example, of the 23.4% of all students who attained a GED or a high school diploma while in Job Corps nearly 8 out of 10 also completed their vocational training course. In addition, of the 42.2% of all students who completed their vocational training, over 4 out of every 10 also received a GED while in Job Corps. This significant overlap in the groups of students who achieve these two key milestones suggests we need to exert caution in interpreting the impact findings as resulting from the completion of the individual milestone.

The last row in the table shows the average length of stay in terms of paid days for each of these six student groups. These data further reinforce the overlap of certain milestone groups, as well as revealing the strong confounding factor that program length of stay plays in the evaluation process. That is, all student groups that complete their vocational training stay in the program for about 10-11 months (in paid days) on average. This compares to only about 2 months for those students who do not complete their vocational training program (and do not obtain a GED for those without an academic credential at application), and to about 4 months for those students who receive their GED only.

The remainder of Table III-1 provides information on the background characteristics of each of these student groups. The results in these other rows generally follow the pattern that students without a GED at application who reach both milestones and those who only attain a GED are typically less disadvantaged. For example, relative to students without a GED at application who either only complete their vocation or who do not reach any milestones, individuals in these groups are more likely to have completed 11th grade, less likely to have received welfare or have lived in public housing, less likely to be black and more likely to have prior work experience. Among those with a GED or high school diploma at application, there are few differences in the characteristics of students between vocational completers and non-completers.

IV. RESULTS

In this chapter we present our estimates of the effects on earnings for students that achieve key program milestones using comparison groups developed from the kernel matching methods described in Chapter II. We first present the overall impacts on earnings for all applicants, and for the subgroup that enroll in Job Corps. We then present estimated impacts for participants who did and did not complete a vocational training program, and for those who did and did not obtain a GED or high school diploma while enrolled in Job Corps.

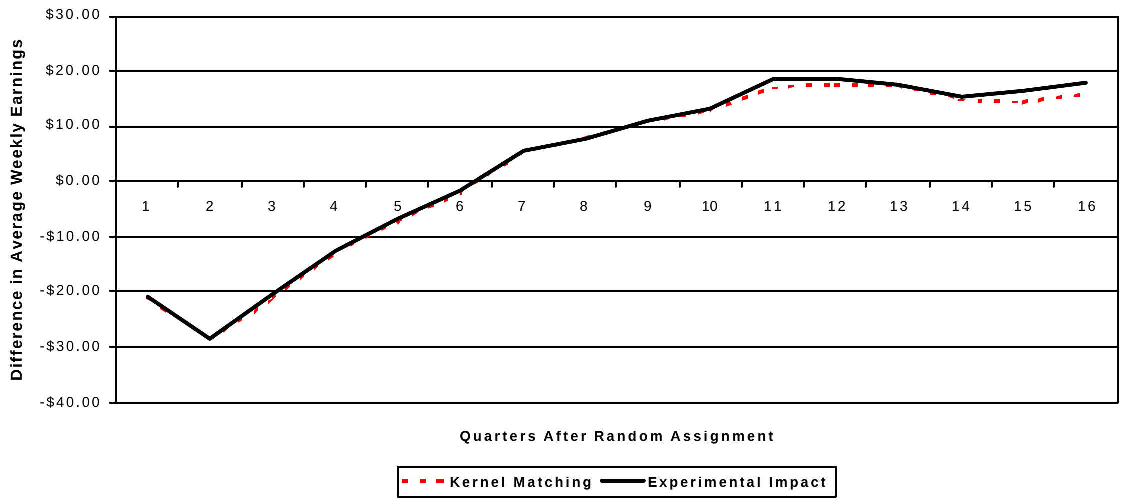
A. IMPACTS FOR APPLICANTS AND PROGRAM PARTICIPANTS

We begin by comparing the impact estimates based on the kernel matches with the estimated impacts obtained using the entire control group created from the random assignment design. Because the kernel matches are drawn from the control group and essentially involve a re-weighting of control group outcomes, we expect the impact estimates from the kernel matching technique will be very close to the experimental impact estimates for the program group as a whole. Only if the matching method created a real distortion of the control group would the impact results from the kernel matches differ from the experimental findings. Although the similarity of the impact estimates and the experimental results is not a very strong test of the kernel matching method, it is a useful check of the general approach.

In Figure IV-1, we compare the impact estimates on average weekly earnings for the 16 quarters after random assignment based on the experimental design with the impact estimates derived from the kernel matching methods for all Job Corps applicants. As this figure shows, the

Figure IV-1

Comparison of Impact Estimates Derived From Experimental and Kernel Matching Methods: Age 16-24



experimental impact estimates are negative for the first 6 quarters after random assignment. This is the period when program group members are most likely to be enrolled in Job Corps, and control group members are relatively more likely to be working. Following the sixth quarter, the overall experimental impact estimates are consistently positive and grow to average about $15-$20 per week during quarters 11-16.

As can be seen in Figure IV-1, the impact estimates based on the kernel matching methods are nearly identical to the experimental estimates throughout the 16 quarters following random assignment. Although not reported in this figure, this consistency of the earnings impact estimates across the two methods also holds when the analysis is conducted separately for each of the three applicant age groups (16-17, 18-19, and 20+). This gives us some assurance that the kernel matching methods are not introducing any major distortions in the overall control group.

Figure IV-2

Comparison of Impact Estimates for Participants Derived From Experimental and Kernel Matching Methods: Age 16-24

In Figure IV-2, we provide a second benchmark for assessing the reasonableness of the kernel matching technique. Specifically, we compare the participant earnings impact estimates based on the kernel matches with the experimental impact estimate adjusted for no-shows. The participant impact estimates reported using the experimental design are based on the standard assumption that the impact of Job Corps on the program group members who choose not to enroll in the program is zero. In other words, we assume that the option of being able to enroll in Job Corps has no effect on the subsequent labor market outcomes for those who do not enroll.

The results in Figure IV-2 indicate that the kernel matching method for participants tracks the adjusted experimental impacts reasonably well, but are slightly lower in every quarter. Specifically, the kernel matches yield estimates of program impacts on participants of about $15-$20 per week in quarters 11-16. A comparison of Figures IV-1 and IV-2 indicates that the average impact for participants using the experimental design adjusted for no-shows follows the same general pattern as the impact for all applicants except it is somewhat magnified in absolute value. In particular, the experimental findings for participants are negative up through quarter 6 and correspond to a $20-$25 per week impact for quarters 11-16. Taken together, this is consistent with a relatively stable but positive non-experimental impact for no-shows of about $5 per week. These results indicate that the kernel matches provide reasonable estimates of net impacts, although they may be biased slightly downward.

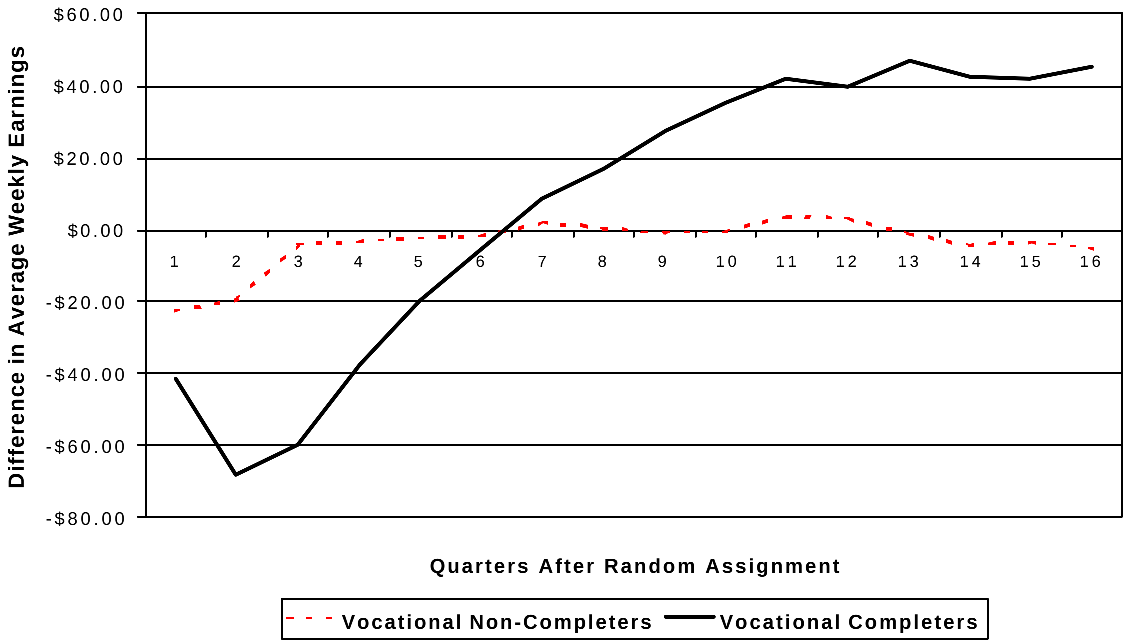
B. IMPACTS FOR VOCATIONAL COMPLETERS

We now present the estimates of impacts for students who achieve key program milestones based on the kernel matching methods. Because these results are obtained from among those who participate in the program, it is important to recognize that there are no counterparts that can be developed from the experimental design to serve as benchmarks of the reasonableness of the estimates. Moreover, the reliability of the impact estimates for vocational completion are not only affected by the reliability of the propensity scores for completing a vocation, but also affected by the reliability of the propensity scores for enrollment. It is important to keep these potential caveats in mind in interpreting the results presented below.

In Figure IV-3, we present the kernel impact estimates for participants who complete a vocational training program and those who do not complete a vocational training program across all age groups. As this figure indicates, participants who do not complete their vocational training are estimated to have lower earnings than their kernel matched comparison group during the period they are most likely to be enrolled in Job Corps (i.e., the first 2-3 quarters after random assignment). However, throughout the

Figure IV-3

Non-Experimental Impact Estimates for Participants by Vocational Completion Status: Age 16-24



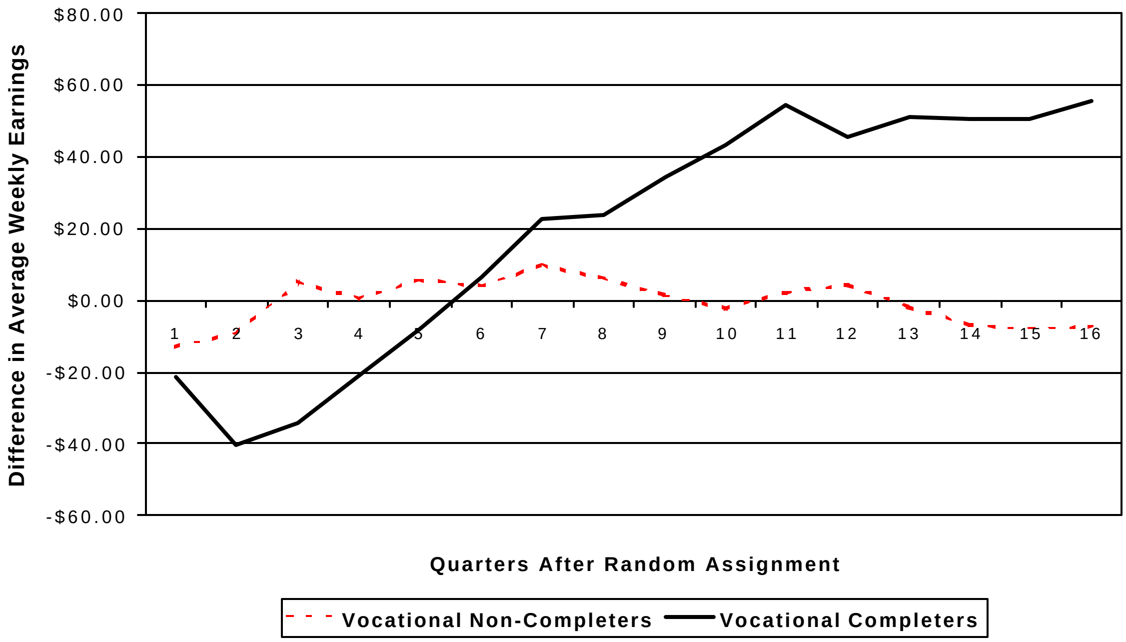
remainder of the 16-quarter observation period, the earnings of those who do not complete their vocational training are nearly identical to the earnings of their matched comparison group. This indicates that essentially all of the estimated positive program impacts for participants accrue to those who complete a vocation.

The size of the estimated impacts for vocational completion are also shown in Figure IV-3. Consistent with the overall participant results reported above, we find that the impact for those students who complete their vocational training is negative until about quarter 6, and is positive thereafter. The earnings impacts for vocational completion are quite stable -- between $40-$50 per week – in quarters 11-16.

To get some sense of the robustness of these findings, we also examined the earnings impact estimates for vocational completion status by age group. Although there is more variability in these

Figure IV-4

Non-Experimental Impact Estimates for Participants by Vocational Completion Status: Age 16-17



findings than the results for all students combined, the pattern of results by age are extremely similar to the overall findings in Figure IV-3. For 16-17 year-olds,

the earnings impact estimates for vocational completers range between $45-$55 per week in quarters 11-16, and the impacts for non-completers are close to zero in all quarters (Figure IV-4).

The results for 18-19 year-olds generally show smaller impact estimates for both vocational completers and non-completers (Figure IV-5). The impact estimates for vocational completers range between $20-$30 per week in quarters 11-16; the impact estimates for non-completers typically range between $0 and -$20. It is important to note that the lower impact estimates for 18-19 year olds is not an artifact of this matching technique; the same pattern is present in the overall main impact results based on the experimental design.

The results for 20-24 year-olds (Figure IV-6) are quite similar to the pattern observed for 16-17 year-olds. That is, the impacts for vocational completers are initially negative, become positive after quarter 6 and average between $40-$50 per week during quarters 11-16. The

Figure IV-5

Non-Experimental Impact Estimates for Participants by Vocational Completion Status: Age 18-19

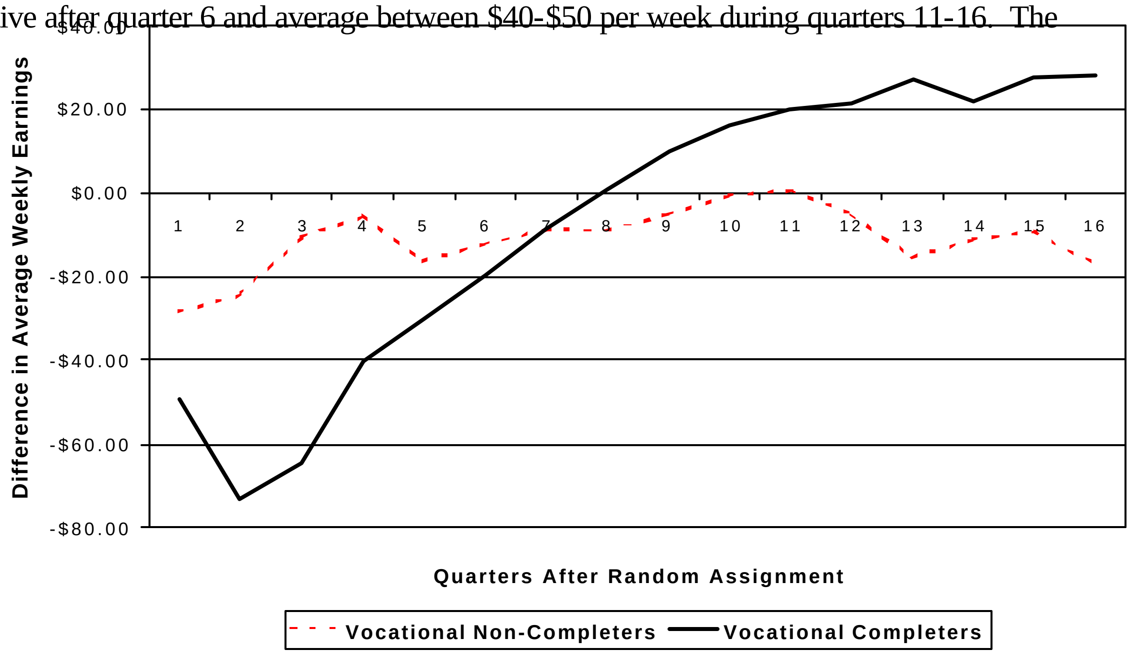
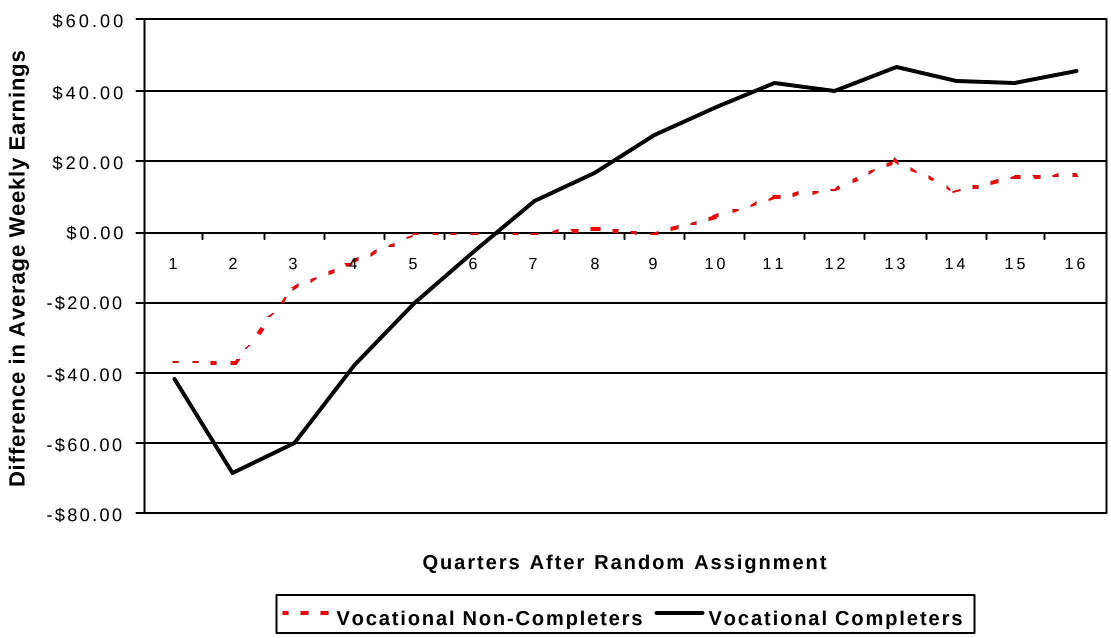


Figure IV-6

Non-Experimental Impact Estimates for Participants by Vocational Completion Status: Age 20-24



impact estimates for non-completers are also negative initially and essentially zero for the next several quarters. However, in quarters 11-16, the estimated impacts are consistently positive at about $10-$20 per week. This estimated positive impact on earnings for participants who did not complete their vocational training primarily reflects the (relatively) lower earnings of the control group as a whole for students of this age, and not the higher earnings of the program group.

Taken together, we interpret the pattern of the impacts for vocational completers on earnings over time and by age group as evidence that positive program impacts primarily accrue to students who complete a vocation.

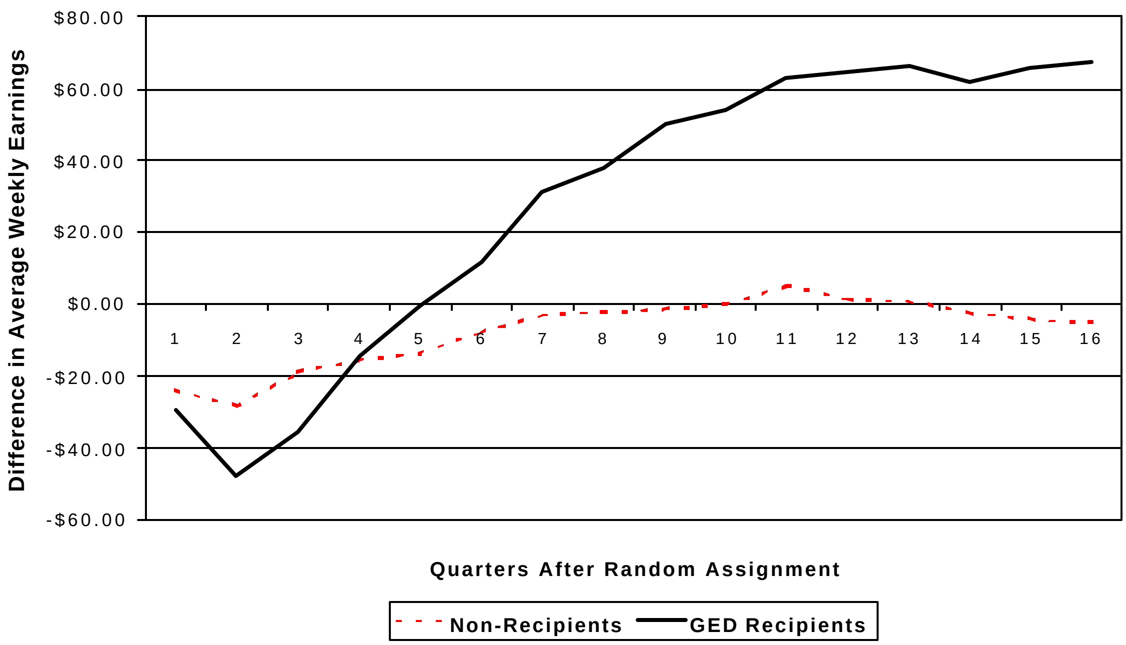
C. IMPACTS OF GED ATTAINMENT

Attainment of a GED or high school diploma in Job Corps represents achievement of another key program milestone and this section presents impact estimates for the participants who reach this milestone. However, this achievement is only attainable by the youth who enter the program without a high school credential. As such, this analysis is restricted to the subset of youth that did not have such a credential at the time they applied to Job Corps.

The earnings impact estimates for all participants who complete a GED or high school diploma while enrolled in Job Corps, and for all who do not are shown in Figure IV-7. As for vocational completers and non-completers, nearly all of the positive impacts accrue to participants who earn a GED or high school diploma and none to the non-completers. That is, participants who do not earn a GED have an initial negative earnings impact during the period

Figure IV-7

Non-Experimental Impact Estimates for Participants by GED Recipiency Status: Age 16-24



they are most likely to be enrolled in Job Corps. However, during the remainder of the observation period, the earnings of this group of participants are nearly identical to those of their matched comparison group.

In contrast, the estimated impacts for students without a GED at application but who earn a GED or high school diploma in Job Corps is negative initially, and becomes positive in the fifth quarter. The fact that GED recipients experience positive earnings impacts earlier than other participants on average is consistent with the somewhat shorter program length of stay of GED completers than of vocational completers. After quarter 5, the estimated earnings impacts rise steadily from quarter 5 to quarter 11 to about $60 per week, and remain in the $60-$70 per week range throughout quarters 11-16.

In Figures IV-8 through IV-10, we show the estimated impacts for GED completers and non-completers by participant age. In contrast to the findings for vocational completion status, the impacts for GED (or high school) recipients are essentially the same across participant age groups. More specifically, for each of the three age groups, the earnings impact estimates for GED or high school diploma recipients in quarters 11-16 is typically between $60-$80 per week. Moreover, the earnings impact estimates for participants who do not receive their GED or high school diploma are usually near zero. The consistency of the patterns of estimated earnings impacts across age group and over time provides additional confidence in these non-experimental findings.[16]

16 Although the pattern of estimated impacts for GED or high school diploma recipients across the three age groups may appear inconsistent with the findings from the experimental analysis, combining the estimated impacts for non-participants and the participants who did not reach this program milestone resolves this apparent discrepancy. Specifically, the negative estimated impacts for the 18 and 19 year old non-participants and participants who did not attain a high school credential averaged in with the positive impact for those who reach this milestone yield estimates for the entire 18 to 19 age group that are consistent with the experimental results.

Figure IV-8

Non-Experimental Impact Estimates for Participants by GED Recipiency Status: Age 16-17

Figure IV-9

Non-Experimental Impact Estimates for Participants by GED Recipiency Status: Age 18-19

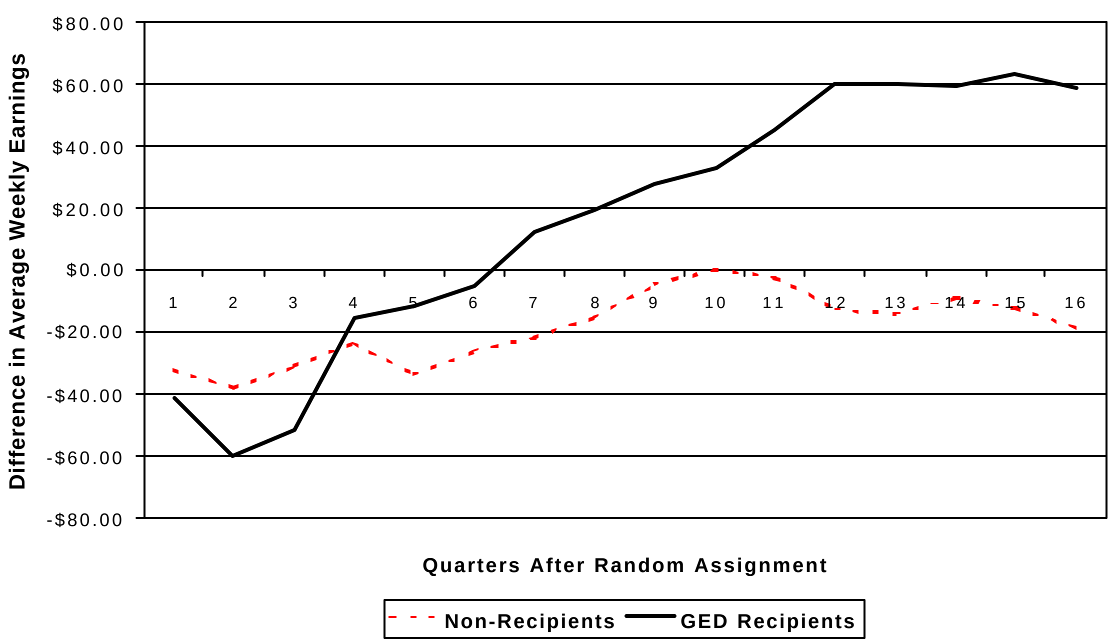
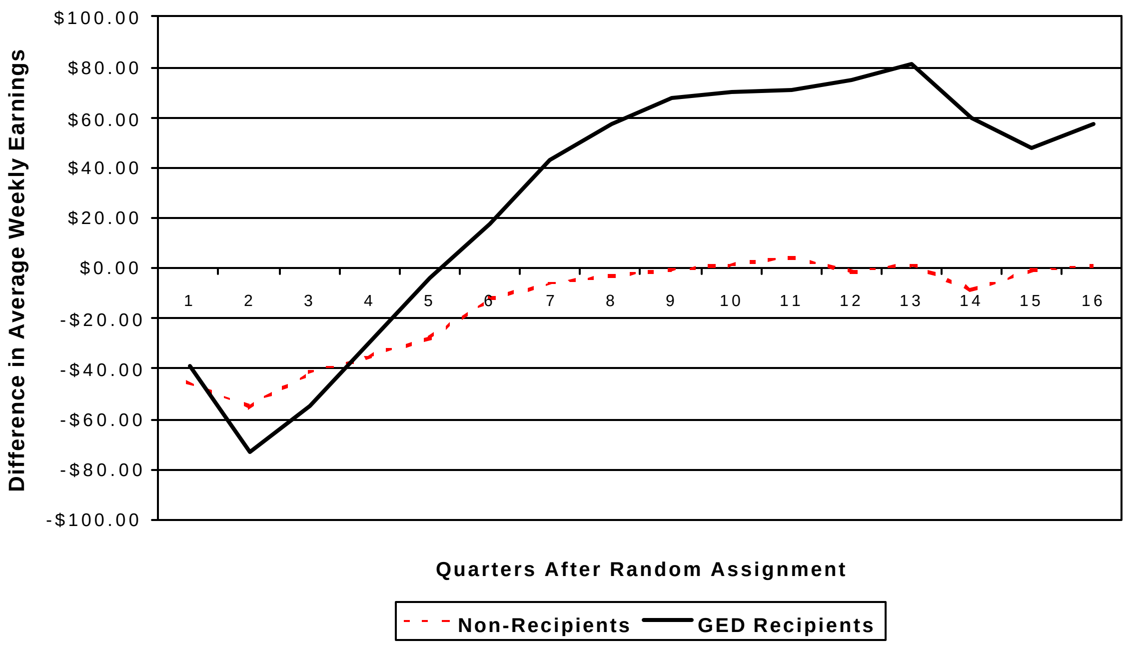


Figure IV-10

Non-Experimental Impact Estimates for Participants by GED Recipiency Status: Age 20-24



V. CONCLUSIONS

This report is one of a series presenting findings from the National Job Corps Study. The purpose of this report is to go beyond simple average impacts and provide information to program operators and others regarding the effects for students who achieve key program milestones. In particular, we apply non-experimental statistical methods to estimate the impacts on average weekly earnings over the 48 months after application for eligible applicants who achieve one (or more) of two key milestones: (1) complete a Job Corps vocational training program; or (2) obtain a GED or high school diploma while enrolled in Job Corps.

The results presented in Chapter IV are based on kernel matched comparison groups stratified by gender that were developed from the study's randomly assigned control group of eligible applicants who were not permitted to enroll in Job Corps. In addition, we also investigated using econometric models to estimate the effects on earnings for students that achieved key program milestones. Although an extensive literature has applied econometric models to derive non-experimental impact estimates for many programs similar to Job Corps, these models consistently failed our application of traditional specification checks. Given the extensive amount of individual background characteristics and program characteristics available to incorporate in these models, their failure to meet standard specification checks is an important result. As a consequence, the conclusions summarized below are based exclusively on matching methods that have recently become the methodology of choice for the estimation of non-experimental impacts.

The highlights of our findings based on the kernel matches are summarized below:

- The overall impact estimates developed from the kernel matches closely track the experimental impacts of approximately $15-$20 per week for all applicants for quarters 11-16 after random assignment. This result was expected and should not be interpreted as strong evidence of the validity of the kernel matching approach.

- However, the kernel matches yield impact estimates for those who enroll in the program (i.e., participants) that are slightly lower than the experimental results. Specifically, the experimental findings correspond to a $20-$25 per week impact for participants over these quarters, as compared to an estimated $15-$20 per week impact using the kernel matches. This indicates that the overall impact estimates for non-participants based on the kernel matches is approximately $5 per week.

- Among participants, we find that nearly all of the positive program impacts are estimated to accrue to those who accomplish one of the two major milestone achievements in the program – completing a vocation or receiving a GED.

- The estimated impact for students who complete their vocational programs becomes positive after the sixth quarter, reaches $40 per week by quarter 11 and remains between $40-$50 per week through quarter 16. Students who do not complete their vocations are estimated to have slightly lower earnings than their matched comparison group during the period they are most likely enrolled in Job Corps and then have earnings that are nearly identical to their matched comparison group throughout the remainder of the observation period.

- Similar to the findings for vocational completion, nearly all of the positive impacts for students who did not have a GED at entry are estimated to accrue to participants who receive a GED. Among students without a GED at entry, non-recipients are estimated to have an initial negative impact during the period they are most likely enrolled in Job Corps and then have earnings that are nearly identical to their matched comparison group. Among students without a GED at entry, the estimated impact for GED recipients becomes positive in the fifth quarter, reaches about $60 per week during quarter 11 and remains between $60-$70 per quarter through quarter 16.

- We also estimated the impacts of programmatic achievements separately by age and found that the general conclusions described above hold for each of the age groups (16-17, 18-19, 20+). The consistency of the patterns of estimated earnings impacts across age groups provides additional confidence in these non-experimental findings.

Job Corps has a long history of trying to promote the attainment of vocational and academic skills and credentials. The Workforce Investment Act of 1998 has further encouraged Job Corps to focus on the attainment of program credentials and graduation. The analyses reported here suggest that this focus is appropriate. The strong and consistent patterns of the impact results over time and across age groups indicate the program's positive average impacts are very likely due to the impacts realized for vocational completers and students who earn a GED in Job Corps and that little or none of the impact is for students who enrolled but did not complete the program.

In interpreting the policy implications of these findings, it is important to recognize the questions these findings address and those they do not. For example, although the results indicate no impacts for non-graduates, this should not be interpreted as evidence that Job Corps should not serve students who do not complete the program. This is because of the inherent difficulty of determining a priori which students will complete the program and graduate and which students will not. Put another way, the predictive power of the propensity score models are not very high. In addition, although we believe the findings provide reasonable evidence of the effects of Job Corps for those students who completed their vocational training and those who did not, they cannot be interpreted as representing what would happen if more students were turned from non-completers into completers.

In understanding the policy implications, it is also important to recognize the extensive uncertainty surrounding the specific impact estimates. Impact findings based on non-experimental methods – such as those necessary to measure impacts for students who did or did not achieve key milestones – are forced to rely on inherently untestable assumptions about the relationships of observed and unobserved factors to program participation and post-program earnings. This fact leaves the results open to the criticism that because key assumptions were not satisfied, the findings may not be accurate. Moreover, because students who complete key milestones typically remain in the program for a long time and receive extensive residential services (including social skills training), the impacts we have attributed to completing a vocation or receiving a GED may also simply result from more time in the program and greater exposure to the other experiences that Job Corps offers. Also, by estimating impacts separately for vocational completion and for GED attainment, we have not examined the effects of achieving one milestone but not the other or the effects of achieving both. Yet, our inability to fully disentangle the effects of completing the program from the effects of greater exposure to the program does not materially affect the importance of the main finding: Job Corps program practices that promote longer retention to facilitate achieving completion of vocational training or attainment of a GED or high school diploma are likely to be beneficial.

REFERENCES

Burghardt, J. and P. Schochet. "National Job Corps Study: Impacts by Center Characteristics." Princeton, NJ: Mathematica Policy Research, Inc., June 2001.

Heckman, J., J. Ichimura, and P. Todd. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program", *Review of Economic Studies*, 64(4): 605-654, 1997.

Heckman, J., R. Lalonde, and J. Smith. "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics, Volume 3,* Ashenfelter, A. and D. Card, eds., Amsterdam: Elsevier Science, 1999.

Johnson, T., R. M. Gritz and M. K. Dugan. "National Job Corps Study: Job Corps Applicants' Programmatic Experiences." Seattle, WA: Battelle Memorial Institute, May 2000.

Johnson, T., R. M. Gritz, R. Jackson, J. Burghardt, C. Boussy, J. Leonard and C. Orians. "National Job Corps Study: Report on the Process Analysis." Seattle, WA: Battelle Memorial Institute, February 1999.

Lee, L. "Generalized Econometric Models with Selectivity", *Econometrica*, 52(2):507-512, 1983.

McConnell, S. and S. Glazerman. "National Job Corps Study: The Benefits and Costs of Job Corps." Princeton, NJ: Mathematica Policy Research, Inc., June 2001.

Schochet, P. "National Job Corps Study: Methodological Appendixes on the 48-Month Impact Analysis." Princeton, NJ: Mathematica Policy Research, Inc., June 2001.

Schochet, P., J. Burghardt, and S. Glazerman. "National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes." Princeton, NJ: Mathematica Policy Research, Inc., June 2001.

APPENDIX

This appendix includes three sections. The first section presents the results of the specification checks described in Chapter II for both matching methods and for the econometric models. The second section presents information on the goodness of fit of these statistical models. In particular, we describe the extent to which the propensity scores based on the statistical models used to summarize the programmatic experiences of Job Corps applicants accurately distinguish between applicants who do and do not achieve a specific milestone. Finally, the third section presents additional non-experimental impact estimates for students who achieve certain program milestones. Specifically, this section presents estimates of the impact of completion of a vocational program and receipt of a GED or high school diploma in Job Corps based on the nearest-neighbor matching method. In addition, we also present estimates of the impact of participation in Job Corps derived from the econometric models described in Chapter II.

A. SPECIFICATION CHECKS FOR NON-EXPERIMENTAL METHODS

As described in Chapter II, four specification checks are used to assess the extent to which the matching methods yield reliable impact estimates and one specification check is used to examine the reliability of the two econometric modeling approaches. Specifically, for the nearest-neighbor and kernel matching methods, the extent to which the matching estimates replicate the experimental impact estimates, both overall and for participants only, and estimates of the impact for "no-shows" provide three specification checks. In addition, for the nearest-neighbor approach, counts of the number of times each control group member is matched to a program group member provides a fourth specification check that can also be used to infer the relationship for the kernel matches because a small window is used to calculate the weights for this approach. The regression models specified in Section D of Chapter II are used to examine the extent to which the two alternative econometric models yield reasonable estimates of the impacts of Job Corps on participants.

This section presents the findings from these specification checks. Results from these specification checks are presented separately for each of the three age groups (16-17, 18-19, and 20-24) because distinct models were estimated to characterize the experiences of eligible applicants in each of these age groups. Moreover, for the fourth specification check for the nearest-neighbor matching methods, results are presented separately for men and women within each of these three age groups because the matches were also stratified along this dimension.

Figures A-1 through A-3 present the comparison of the overall impact estimates based on the experimental design with the corresponding estimates derived from the nearest-neighbor and the kernel matching methods defined in Chapter II. Figure A-1 presents the comparisons for applicants who were 16 to 17 years old at the time of application, Figure A-2 presents the comparisons for applicants 18 to 19 years old at the time of application and Figure A-3 presents the comparisons for applicants 20 to 24 years old at the time of application. Overall, these three figures show that the kernel matching method more closely replicates the experimental impact estimates than the nearest-neighbor approach for all three age groups. Although there are a few exceptions, the nearest-neighbor approach generally over-estimates the impact for all age groups. In comparison, the kernel method nearly matches the experimental impact estimate for all age groups over the entire 16 quarters after random assignment. As such, this first specification check suggests that the kernel method provides more reliable impact estimates of the two matching methods.

FIGURE A-1
Comparison of Overall Impact Estimates Derived From Experimental and Matching Methods: Age 16-17

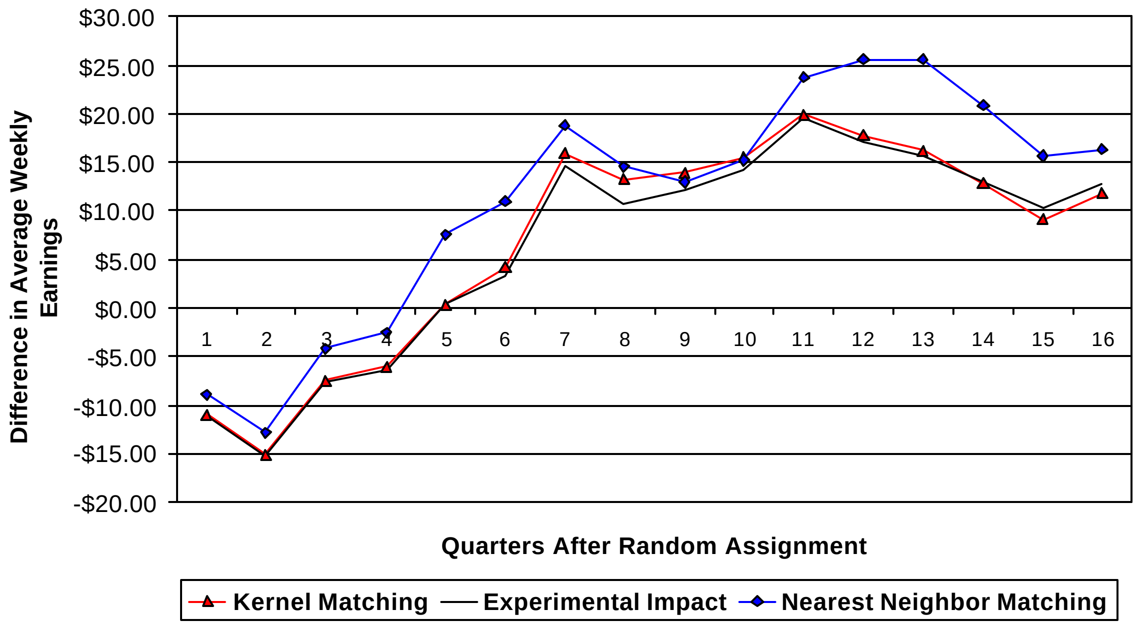


FIGURE A-2
Comparison of Overall Impact Estimates Derived From Experimental and Matching Methods: Age 18-19

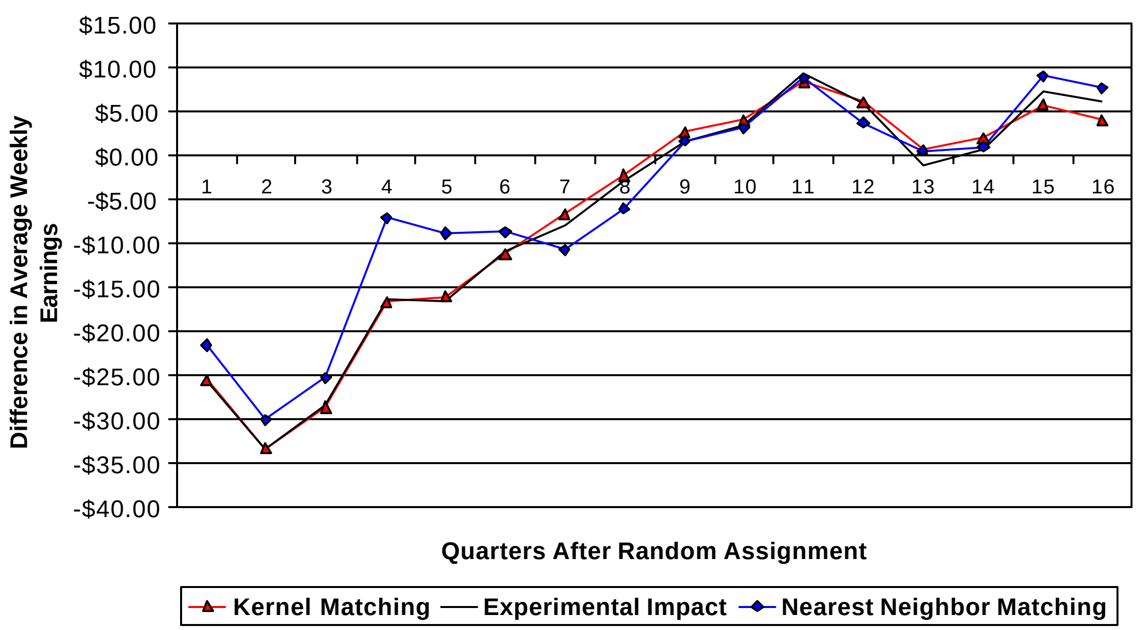
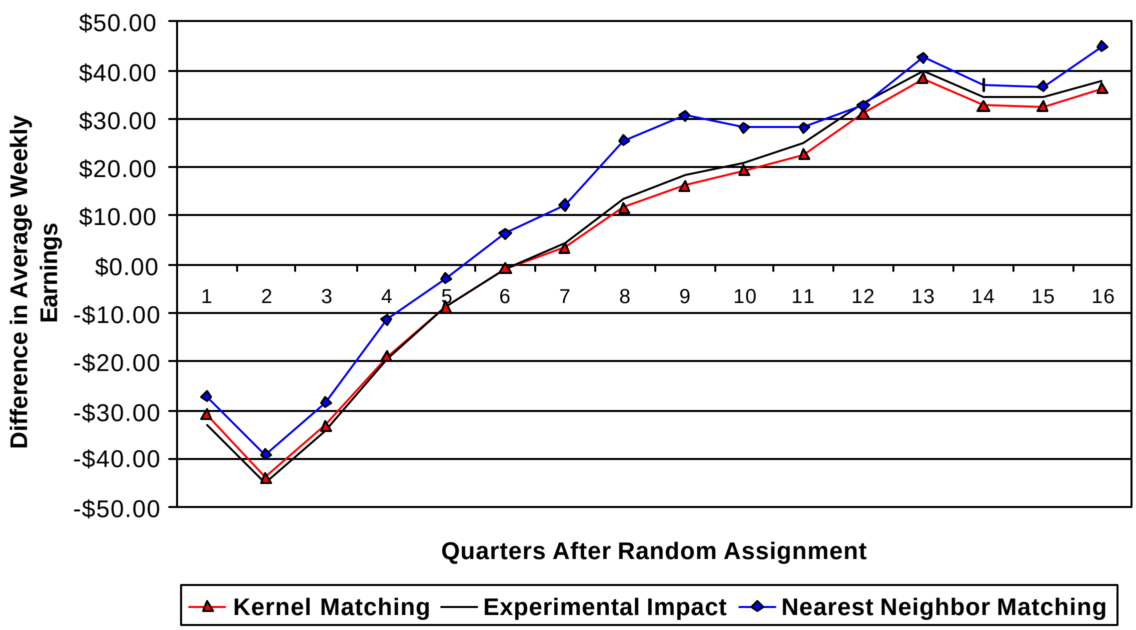


FIGURE A-3
Comparison of Overall Impact Estimates Derived From Experimental and Matching Methods: Age 20-24



Figures A-4 through A-6 present the corresponding comparisons for the estimated impact of Job Corps on participants. As shown in these figures, for the 16 to 17 and 18 to 19 age groups the kernel matches more closely track the "no-show" adjusted experimental estimate relative to the nearest neighbor method. In contrast, the results for the 20 to 24 year old age group are more mixed. Specifically, during the early part of the period, which corresponds to the period when participants are enrolled in the program, the nearest-neighbor method more closely tracks the experimental estimate. However, during the last half of the period both the nearest-neighbor and kernel method estimates deviate noticeably from the experimental estimate. The results from this specification check suggest that the kernel matches are likely to provide more reliable impact estimates for the two younger age groups and that both methods are likely to be less reliable for the oldest age group because of the noticeable difference from the experimental estimate.

FIGURE A-4
Comparison of Particpant Impact Estimates Derived From Experimental and Matching Methods: Age 16-17

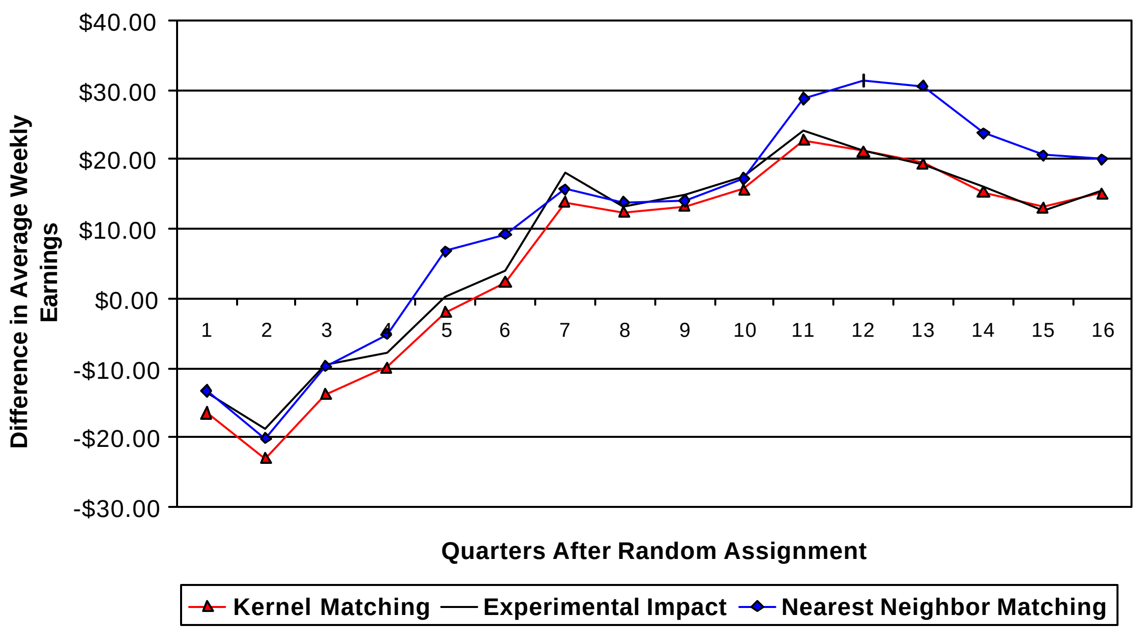


FIGURE A-5
Comparison of Particpant Impact Estimates Derived From Experimental and Matching Methods: Age 18-19

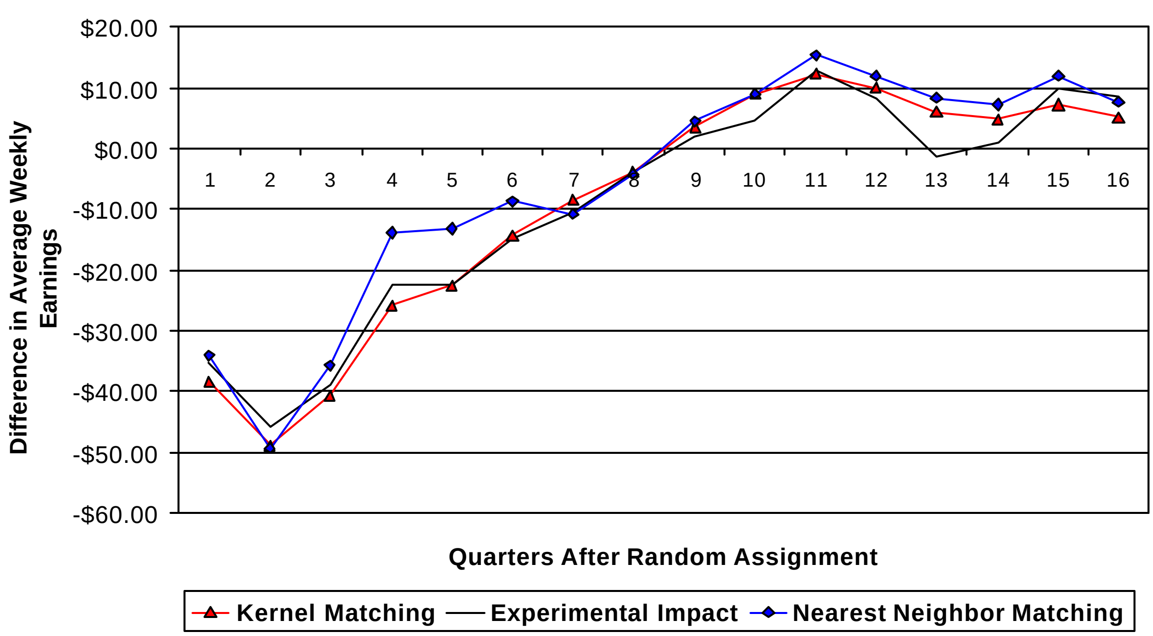
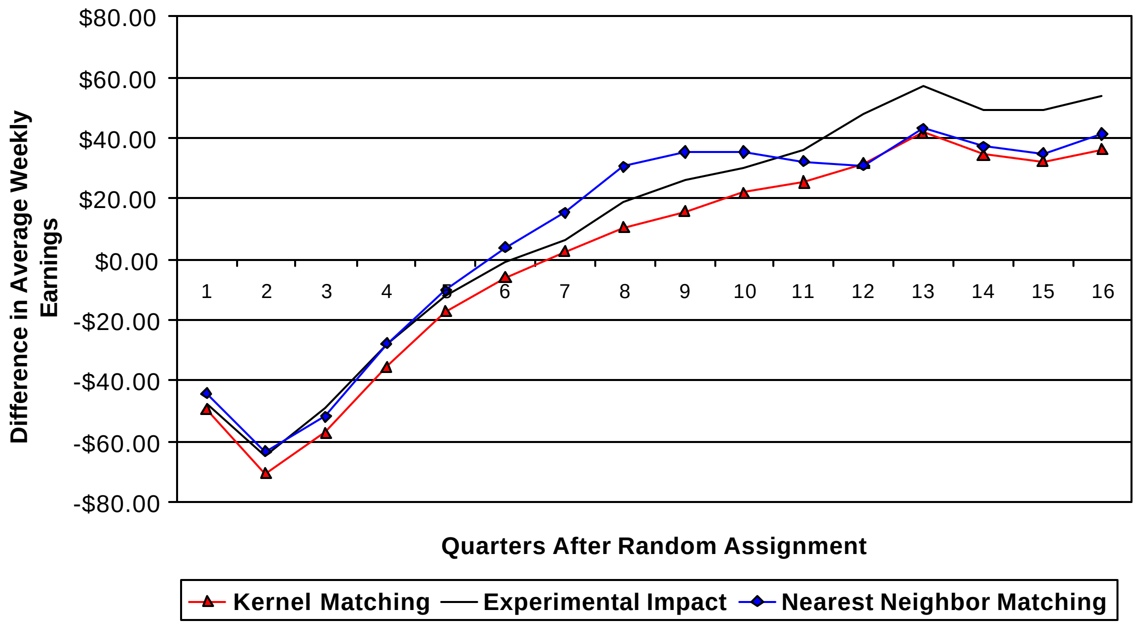
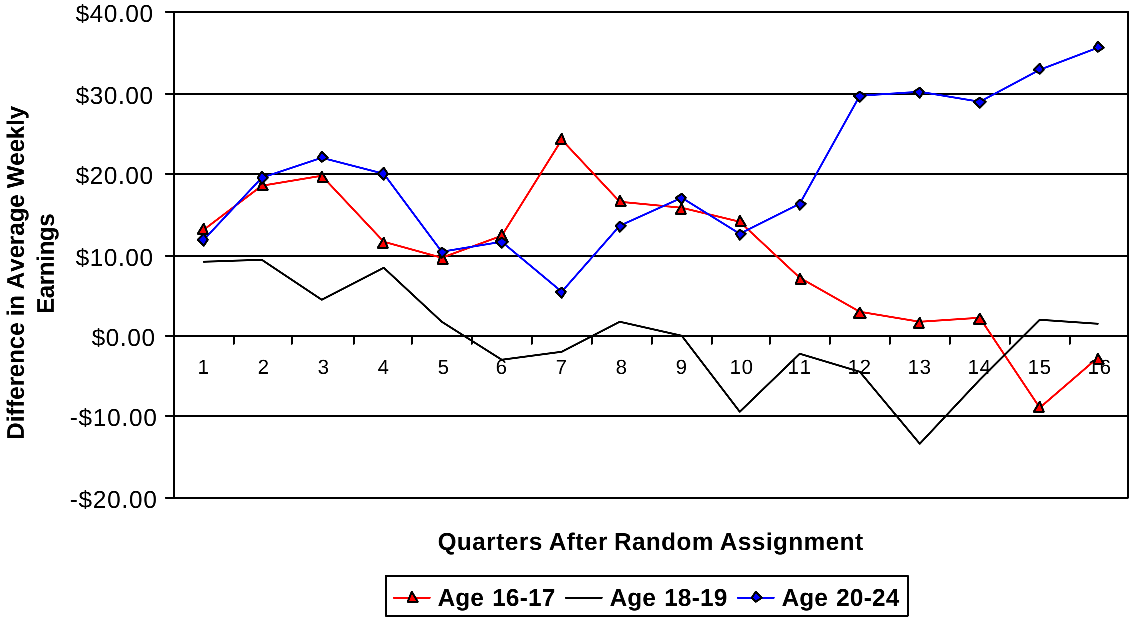


FIGURE A-6
Comparison of Particpant Impact Estimates Derived From Experimental and Matching Methods: Age 20-24



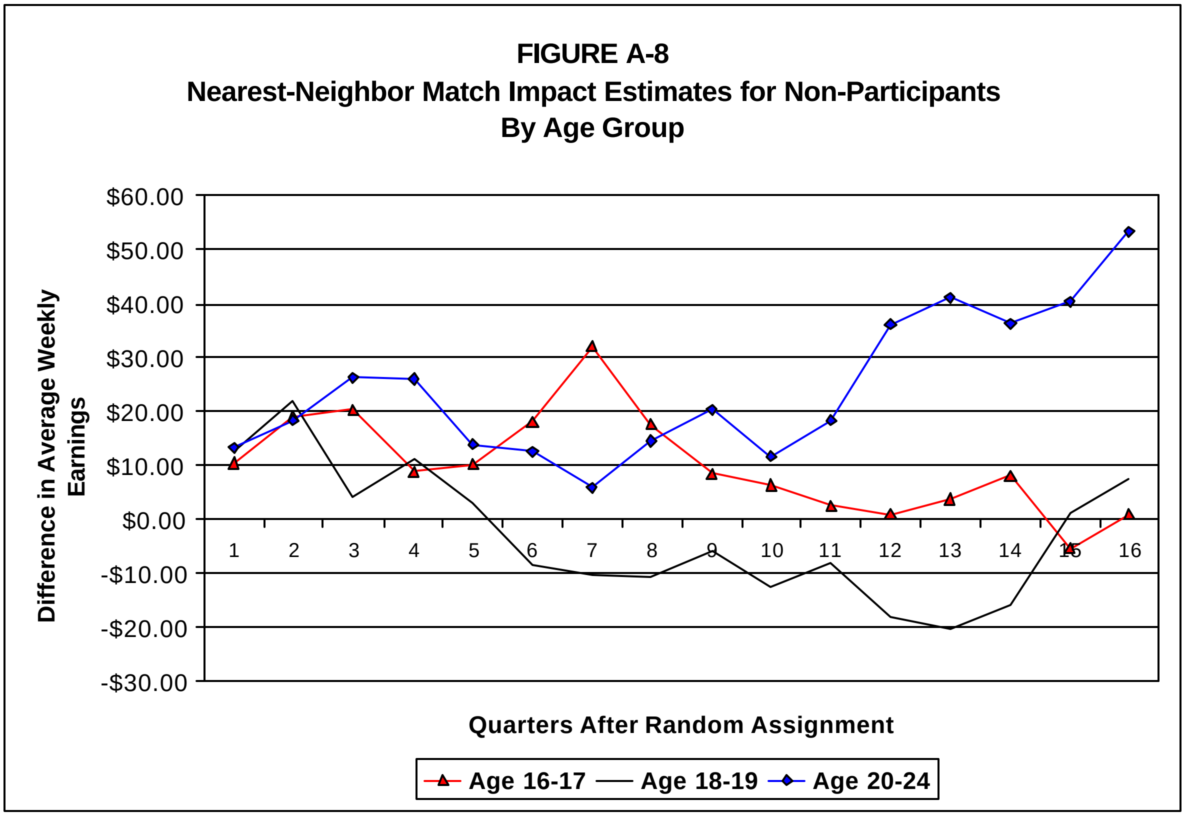
The third specification check for the two matching methods compares the estimated impact for eligible applicants in the program group who decided not to enroll in Job Corps to zero, which represents the value used in the adjustment of the experimental impact to estimate the impact for participants. Figure A-7 presents the results from this specification check for all three age groups for the kernel matches and Figure A-8 presents the corresponding results for the nearest-neighbor matches. As shown in these figures, this specification check confirms the findings and cautions raised by the previous specification checks. Comparing Figures A-7 and A-8 these findings support the general conclusion that the kernel matching approach is preferred over the nearest-neighbor method. Moreover, under the assumption that the program has no impact on applicants who do not enroll, which is the assumption used in deriving the

FIGURE A-7
Kernel Match Impact Estimates for Non-Participants
By Age Group



experimental impact estimate for participants, these findings suggest that the kernel matching methods are more likely to yield reliable estimates of the impacts for participants who achieve specific program milestones for the two younger age groups, particularly the 18 to 19 year old age group. However, the results in Figure A-7 also suggest that caution should be used in interpreting the results for all three age groups.

The final specification check for the matching methods applies only to the nearest-neighbor approach. This check examines the number of times each control group member is matched to multiple program group members, as well as the extent to which a single control group member is matched to one or more program group members who participate in Job Corps and one or more program group members who do not enroll in the program. We expect each control group member to be matched to approximately two program group members because fewer applicants were randomly assigned to the



control group. Further, because different rates were used to

assign men and women to the control group, we expect female control group members to be matched to more program group members. Finally, to the extent that the estimated propensity scores, which are used in the matching process, distinguish participants from non-participants we would expect that control group members would only be matched to participants or non-participants and not both. Hence, the more control group members are matched to only participants or non-participants the more confidence one can place in the estimates derived from the matching process.

Table A-1 presents findings from this specification check for the nearest-neighbor matches based on the arrival propensity scores. Table A-1 shows the frequency distribution of the number of times a control group member was matched to a program group member within each age and gender group

TABLE A-1
Frequency Distribution of the Number of Nearest-Neighbor Matches For Control Group Members By Age Group and Gender

Number of matches	Percentage of Control Group					
	Age 16-17		Age 18-19		Age 20-24	
	Male	Female	Male	Female	Male	Female
1	44.8	40.1	46.1	36.0	44.8	35.9
2	28.5	23.1	25.8	26.3	25.8	24.7
3	13.9	16.6	13.8	16.3	14.0	19.5
4	7.3	10.5	7.5	10.7	8.7	9.6
5	2.1	3.1	3.4	3.8	3.5	3.8
6	2.4	2.9	2.3	2.0	1.8	3.3
7	0.5	1.6	0.4	2.8	0.9	1.6
8	0.3	0.5	0.4	0.8	0.2	1.4
9+	0.3	1.5	0.4	1.3	0.2	0.3

based on the nearest-neighbor method. As expected, these results show that the vast majority of control group members were matched to one, two or three program group members and that female control group members were more frequently matched to multiple program group members compared to their male counterparts. Table A-2 shows the percentage of control group members that are matched only to participants and the percentage that are matched only to non-participants. As expected, a higher percentage of males were matched to either just participants or just non-participants compared to females. The findings in this table also show an inverse relationship between the age and the percentage matched to only participants or non-participants. Again, the results suggest that the matching methods work best for the younger age groups and raise some concerns about the oldest age group.

Taken together, these four specification checks present a mixed picture regarding the extent to

TABLE A-2
Percentage of Control Group Members Matched Only to Participants and Non-Participants
By Age Group and Gender

	Percentage of Control Group					
	Age 16-17		Age 18-19		Age 20-24	
Matched to	Male	Female	Male	Female	Male	Female
Participants only	70.7	61.9	62.9	48.7	57.7	44.1
Non-participants only	10.2	10.5	13.2	15.8	16.6	19.7
Total	80.9	72.4	76.1	64.5	74.3	63.8

which matching methods are likely to yield reliable non-experimental impact estimates. The first two specification checks by and large indicate that the kernel method is preferred to the nearest-neighbor approach and that this approach is likely to yield reasonable estimates. All four checks consistently suggest that the matching method is more reliable for the younger age groups and raise cautions about both matching approaches with respect to the oldest age group. However, the third specification check suggests that caution should be used in interpreting the results from the application of either matching method to estimate the impacts for students who achieve specific Job Corps milestones.

While the specification checks for the matching methods are not very precise, the specification checks for the econometric models described in Chapter II are very clear-cut. Specifically, the assumptions required for the selection correction and instrumental variable methods to yield unbiased estimates imply that after correcting for selection in the outcome equations no-shows should be indistinguishable from the entire control group. That is, in a regression model estimated over no-shows and the control group and that incorporates selection on both observed and unobserved variables, the estimated “impact” of assignment to the program group should be zero.

Figures A-9 through A-11 present the findings from the regression models specified in Chapter II for the entire control group and the program group no-shows. Figure A-9 presents the findings for the 16 to 17 year old age group, Figure A-10 presents the findings for the 18 to 19 year old age group and Figure A-11 presents the findings for the oldest age group at application. Each figure shows the estimated difference in average weekly earnings of the no-shows relative to the entire control group after accounting for selection in both the selection correction model and in the instrumental variables approach. If the assumptions required for these models to yield unbiased estimates of the impact of Job Corps for participants hold, then the estimated "impact" of assignment to the program group for no-shows should be zero.

The findings presented in Figures A-9 through A-11 suggest that the selection correction model does not adequately account for the selection process characterizing the participation decision of eligible applicants. For example, although the estimates for the two younger age groups are centered around zero, the parameter estimates fluctuate widely and display patterns over time that are not consistent with random variation around zero. Moreover, for the oldest age group, the selection correction model estimates are uniformly greater than zero and quite large near the end of the period.

These findings for the instrumental variable approach are somewhat more promising, although they still suggest that this approach is not fully accounting for the selection process. For example, the instrumental variable estimates for the youngest age group are generally above zero over most of the quarters after random assignment. Further, the estimates for the 18 to 19 year old age group are generally below the expected value of zero over most of the quarters after random assignment. Finally, as was the case for the selection correction model, the estimates

FIGURE A-9
Specification Check for Econometric Models: Regression of Controls and Non-Participants for Age 16 - 17

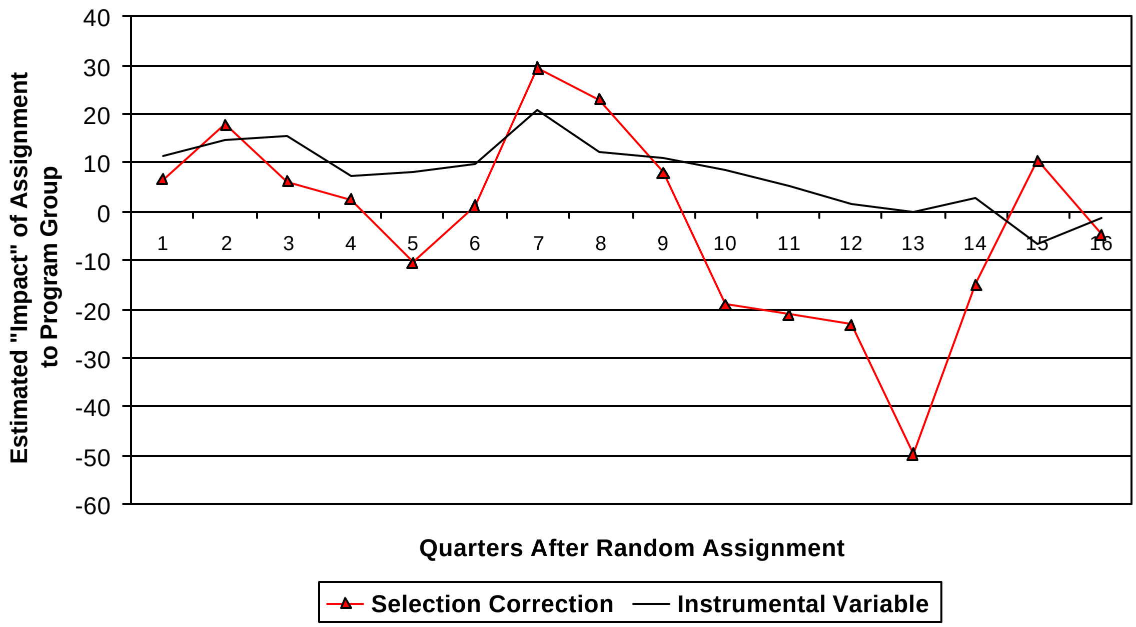


FIGURE A-10
Specification Check for Econometric Models: Regression of Controls and Non-Participants for Age 18 - 19

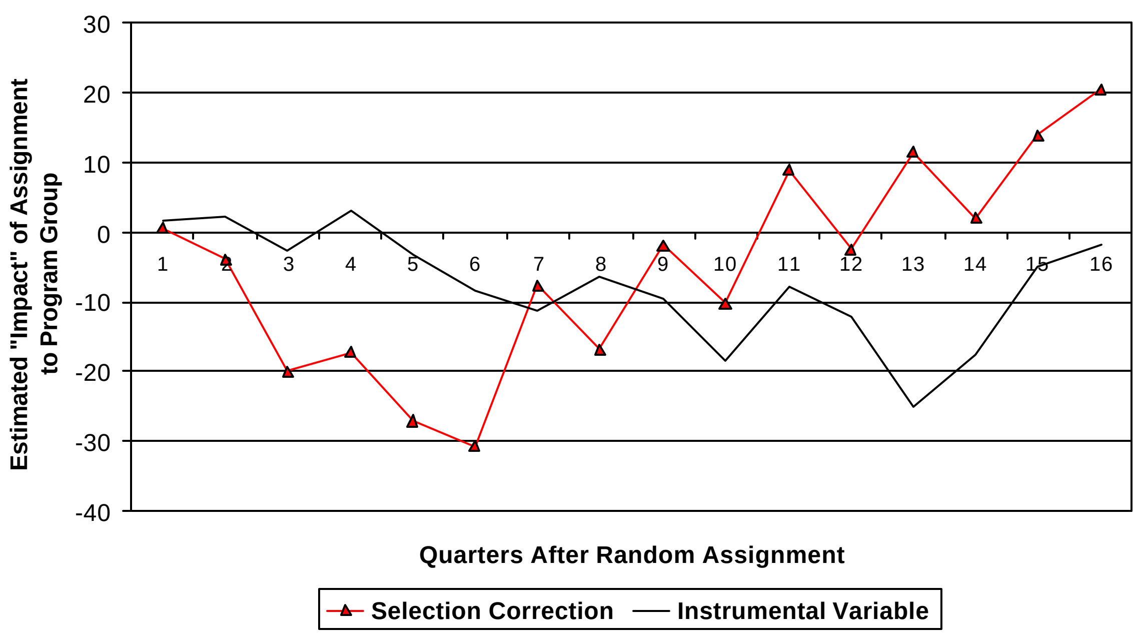
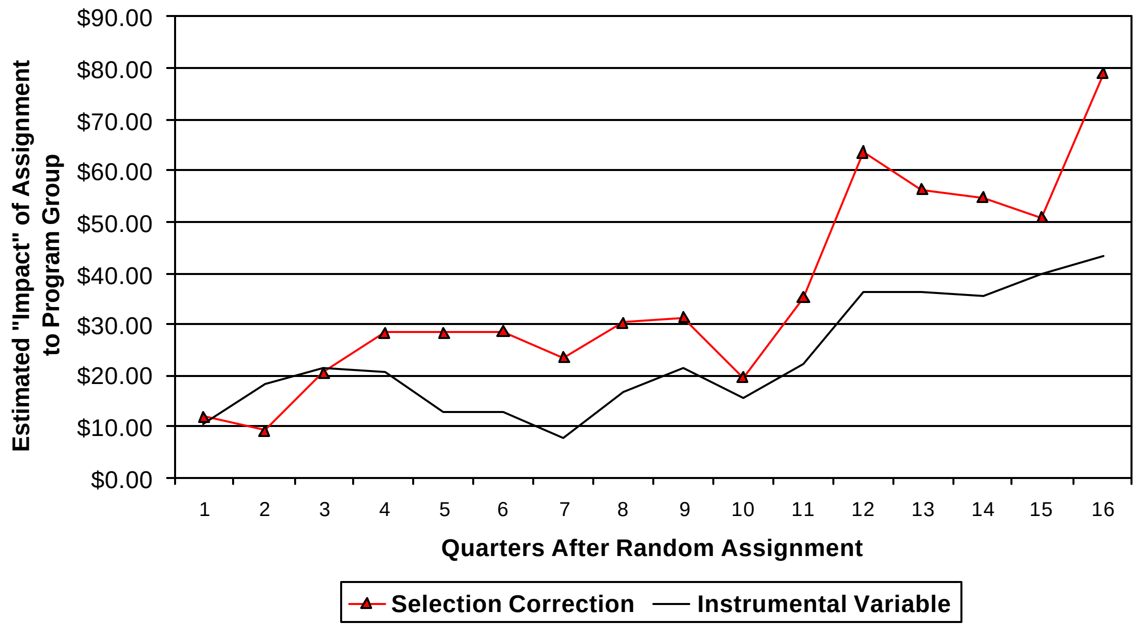


FIGURE A-11
Specification Check for Econometric Models: Regression of Controls and Non-Participants for Age 20 - 24



shown in Figure A-11 indicate that the instrumental variable approach does not account for the selection process governing participation for the oldest age group, as the estimates are positive and trending away from zero over time.

Taken together, the specification checks for the econometric models suggest that the application of either approach to gauge the impact of participation in Job Corps will yield biased estimates. As such, estimates derived from the application of these approaches are not presented in the main body of the report and are only presented below for comparison purposes.

B. GOODNESS OF FIT FOR PROPENSITY SCORES

The extent to which the propensity scores for arrival, vocational completion and receipt of a GED or high school diploma in Job Corps distinguish the eligible applicants who have the experience from those who do not plays a central role in both the matching methods and the econometric modeling approaches used to estimate the impacts of these experiences. As such, measures of the goodness of fit for these propensity scores can supplement the information regarding the reliability of the non-experimental estimates based on the specification checks described in the previous section. Unfortunately, goodness of fit measures for the qualitative choice models that underlie these propensity scores are not nearly as straightforward as the types of measures that are available for simple regression models. This section examines three alternative measures of goodness of fit for the three sets of propensity scores used in the analysis.

The first goodness of fit measure examines the within-sample predictive ability of the propensity scores. Specifically, this measure compares the predicted value for program group members who could potentially have a particular experience with the actual experiences of these individuals. For example, an eligible applicant is predicted to enroll in Job Corps if his or her propensity score is greater than or equal to 0.5 and he or she is predicted to not participate in the program if the propensity score is below 0.5. This within-sample prediction is considered correct if the individual's actual experience is the same as their predicted experience. For instance, if the arrival propensity score for a program group member is 0.7 and this individual participated in Job Corps, the propensity score resulted in a correct prediction. Alternatively, if this individual did not enroll, the propensity score resulted in an incorrect prediction. The higher the percentage of correct predictions the better the goodness of fit of the propensity scores.

Table A-3 presents the results from calculating the percentage of correct predictions based on the propensity scores for arrival, vocational completion and receipt of a GED or high school diploma in Job Corps. These findings are based on the within sample prediction among the program group members who potentially could have achieved each of these milestones. For example, while the arrival results use

TABLE A-3
Percentage of Correct Propensity Score Predictions For Program Group Members
By Age Group

Propensity Score	Percentage with Correct Prediction		
	Age 16 – 17	Age 18 – 19	Age 20 -24
Arrival	81%	76%	73%
Vocational Completion	55%	59%	64%
Receipt of GED/Diploma	75%	71%	68%

the entire program group, the vocational completion findings are based only on participants and the GED or high school diploma findings only use participants who did not have a high school credential at enrollment. Results are presented separately for each of the three age groups.

The results presented in Table A-3 suggest that the models are making more correct prediction than incorrect predictions and--except for vocational completion--the predictions are better for younger age groups. However, these results are only marginally better than a very naive prediction approach that assigned everyone, regardless of their propensity score, the modal value. That is, if more than 50 percent experience the event everyone is predicted to experience the event and if less than one-half experience the event everyone is predicted to not experience the event. For example, this very simple prediction approach would result in correct prediction percentage for arrivals among 16 and 17 year old applicants of 80 percent. Hence, these results suggest that the propensity scores are not adequately distinguishing between those who achieve a milestone and those who do not on an individual basis.

It is not uncommon for statistical models to poorly predict individual-level behavior but yet still capture systematic relationships among variables and distinguish among different groups of individuals.

The second goodness of fit measure more directly examines the extent to which the propensity scores distinguish the groups of individuals based on their actual experiences. Specifically, we examine the distribution of propensity scores for the program group members that achieve each milestone and those that do not. For example, Table A-4 summarizes the distribution of the arrival propensity scores for the program group members that did not participate in Job Corps and the program group members that participated for each of the three age groups. These tables present the mean value of the propensity score, as well as the minimum, the first quartile (i.e., 25th percentile), the median, the third quartile (i.e., 75th percentile) and the maximum values for the corresponding propensity scores.

The results in Table A-4 indicate that the propensity scores are distinguishing between those who have an experience and those who do not have the experience as a group. For example, in Table A-4 the mean arrival propensity score for participants is between 11 and 14 percentage points higher than the mean for non-participants. Moreover, the entire distribution of

TABLE A-4
Distribution of Arrival Propensity Scores for Program Group
By Age Group and Participation Status

	Age 16 - 17		Age 18 - 19		Age 20 - 24	
	Non-Participants	Participants	Non-Participants	Participants	Non-Participants	Participants
Mean	0.71	0.82	0.61	0.74	0.57	0.71
Minimum	0.10	0.18	0.06	0.14	0.01	0.12
1st Quartile	0.62	0.76	0.49	0.65	0.44	0.62
Median	0.75	0.84	0.62	0.77	0.58	0.74
3rd Quartile	0.83	0.90	0.74	0.85	0.71	0.83
Maximum	0.96	0.99	0.93	0.98	0.95	1.00

TABLE A-5
Distribution of Vocational Completion Propensity Scores for Participants
By Age Group and Completion Status

	Age 16 - 17		Age 18 - 19		Age 20 - 24	
	Non-Completers	Completers	Non-Completers	Completers	Non-Completers	Completers
Mean	0.41	0.46	0.44	0.52	0.49	0.59
Minimum	0.05	0.10	0.08	0.14	0.07	0.11
1st Quartile	0.28	0.33	0.34	0.42	0.37	0.50
Median	0.41	0.45	0.44	0.52	0.49	0.61
3rd Quartile	0.54	0.60	0.54	0.63	0.61	0.69
Maximum	0.88	0.87	0.82	0.81	0.88	0.91

the arrival propensity scores for participants dominates the distribution for non-participants. For instance, for the 16 to 17 age group, the values of the propensity scores for all five of the percentile points in the distributions that are presented in the table are higher for the participant group compared to the non-participant group. The results in Tables A-5 and A-6 for the vocational completion and GED/diploma propensity scores display the same patterns as the results in Table A-4. Specifically, the distributions of the vocational completion propensity scores and the GED/diploma receipt propensity scores for participants who achieve these program milestones also dominate the corresponding distributions for those who do not reach the relevant milestone.

TABLE A-6
Distribution of GED/Diploma Receipt Propensity Scores for Participants
By Age Group and Recipiency Status

	Age 16 - 17		Age 18 - 19		Age 20 - 24	
	Non-Recipients	Recipients	Non-Recipients	Recipients	Non-Recipients	Recipients
Mean	0.23	0.34	0.28	0.42	0.29	0.44
Minimum	0.02	0.03	0.01	0.04	0.02	0.07
1st Quartile	0.13	0.23	0.16	0.28	0.13	0.29
Median	0.21	0.32	0.25	0.41	0.24	0.44
3rd Quartile	0.31	0.44	0.38	0.55	0.41	0.60
Maximum	0.71	0.82	0.84	0.86	0.90	0.84

Although far from a definitive statement that the propensity scores are entirely distinguishing the groups of eligible applicants who achieve specific program milestones, the findings in Tables A-4 through A-6 provide some counterbalancing evidence to the conclusions drawn from Table A-3. For example, whereas the results in Table A-3 suggested that the vocational completion propensity scores were not adequately identifying the participants who were going to complete a vocational program in Job Corps, the evidence in Table A-5 suggests that the propensity scores are significantly higher for the participants who completed a vocation compared to the group of participants who did not complete a vocation. Similarly, the findings in Table A-6 also suggest that the propensity scores are distinguishing the participants who did not receive a GED or high school diploma in Job Corps, as a group, from those who did accomplish this program milestone.

A final, although non-standard, assessment of the goodness of fit of the propensity scores measures the extent to which ranking observations by the value of the propensity scores reorders observations in the same way as actual experiences. To develop a measure that captures this concept of goodness of fit, consider a ranking that perfectly ranks observations according to their actual experiences. For example, in the case of participation in Job Corps this ranking would first order the observations so that all of the eligible applicants in the program group that enrolled in the program would come first in the ordering and then all of the non-participants would follow the participants. In this case there would be a 100 percent agreement between the ranking and the actual experiences of individuals. Second consider a ranking that randomly orders observations, which in this case of binary variables will result in an agreement between the rank order and actual experiences that is equal to the sample proportion that have the experience, which we can represent by p. Finally, consider a ranking based on propensity scores that results in a x percent rate of agreement with actual experiences and define the measure of goodness of fit as the percentage improvement from the propensity score ranking over a random ordering as:

$$\text{Goodness of fit} = {(x-p)}\big/{(100-p)}.$$

In general, this goodness of fit measure will lie between zero and one because the propensity scores will generally improve the ability to distinguish participants from non-participants over a random process. However, it is possible for this measure to take on a negative value.

To illustrate the calculation of this measure, consider the participation of eligible applicants who were 16 to 17 years old at the time of application. Among this group 80 percent enrolled in Job Corp, in which case p=80. Among the 80 percent of this group that has the highest arrival propensity scores, 86 percent enrolled in the program. Hence, this measure of goodness of fit would equal 30 percent (i.e., (86-80)/(100-80)=0.30). In other words, the propensity scores improved the goodness of fit for this group by 30 percent over a simple random ordering. Table A-7 presents the calculation of this

TABLE A-7
Percentage Improvement from Propensity Score Ranking Relative to Random Ordering
By Age Group

Propensity Score	Percentage Improvement in Agreement		
	Age 16 – 17	Age 18 – 19	Age 20 –24
Arrival	30%	33%	39%
Vocational Completion	7%	19%	29%
Receipt of GED/Diploma	28%	40%	50%

measure of goodness of fit for the arrival, vocational completion and GED/high school diploma recipiency propensity scores for each of the three age groups. While the results presented in this table are generally consistent with the findings in Tables A-4 through A-6, they are somewhat at odds with the findings presented in Table A-3. For example, the findings in Table A-7 indicate that vocational completion is the experience where the propensity scores result in the least amount of improvement, which is consistent with the findings in Table A-5 that show the least difference between the propensity score of non-completers and vocational completers. However, in contrast to the findings in Table A-3, the findings in Table A-7 suggest that the most improvement occurs for the oldest age group rather than the younger groups.

Overall, the findings presented in this section suggest that the propensity scores improve the identification of eligible applicants who will participate in Job Corps, as well as those that are likely to achieve the major program milestones. However, the findings also reinforce the conclusion drawn from the various specification checks described in the previous section. Specifically, these findings suggest that caution should be used in interpreting the findings derived from the application of non-experimental methods to estimate the impact of Job Corps for participants, vocational completers and GED or high school diploma recipients in Job Corps.

C. ADDITIONAL NON-EXPERIMENTAL ESTIMATES

This section presents additional estimates of the impact of Job Corps for participants, vocational completers and recipients of a GED or high school diploma in Job Corps. Although the specification checks examined above suggest that extreme caution should be used in interpreting the findings from the application of the econometric models to estimate impacts, we present these estimates below for comparison purposes. In addition, this section also presents the estimated impacts of Job Corps for participants who achieve the two major program milestones derived from the nearest-neighbor matching method.

Figures A-12 through A-14 present the estimated impacts of Job Corps for a typical eligible applicant from both the selection correction models and the instrumental variables approach. Figure A-12 presents the estimated impacts for eligible applicants who were 16 to 17 years of age at application, Figure A-13 presents the estimates for 18 to 19 year old eligible applicants and estimates for 20 to 24 year old eligible applicants are presented in Figure A-14. For comparison purposes, these figures also include the experimental estimates based on the difference between the program group and control group means for average weekly earnings in each of the 16 quarters following random assignment.

The results presented in these figures bear out the concerns about the reliability of non-experimental estimates derived from the application of the econometric models described in Chapter II. Moreover, these findings are also consistent with the implications of the specification checks described above. Although the estimated impacts from both the selection correction and instrumental variable models are centered around the experimental impact for the two youngest age groups, the impact estimates fluctuate considerably. Moreover, the shortcomings of these methods for the oldest age group are confirmed by the findings presented in Figure A-14. Overall, these findings reinforce the bona fide concerns about the application of these models in this study.

Results from the application of the nearest-neighbor matching method are presented in Figures A-15 through A-20 separately for each of the three age groups. Figures A-15 through A-17 present the nearest neighbor estimates for vocational completion and for participants that did not complete a Job Corps vocational program. Figures A-18 through A-20 present the nearest-neighbor estimates for the receipt of a GED or high school diploma in Job Corps among the participants who did not possess a high school credential at the time of application to the program.

The findings in these figures generally mirror the findings presented in Chapter IV based on the kernel matching approach. Overall, the estimates based on the nearest-neighbor method closely match the findings based on the kernel matching approach. However, whereas the kernel method implies that the impact for 18 to 19 year old vocational completers remains above $20 per week during the last four quarters of the follow-up period, the findings in Figure A-16 suggest the impact declines markedly during this period. Despite this one noticeable difference, the general conclusions presented in the main body of the report are supported by the results derived from the nearest-neighbor matching methods.

FIGURE A-12
Comparison of Overall Impact Estimates Derived From Experimental and Econometric Modeling Methods: Age 16 - 17

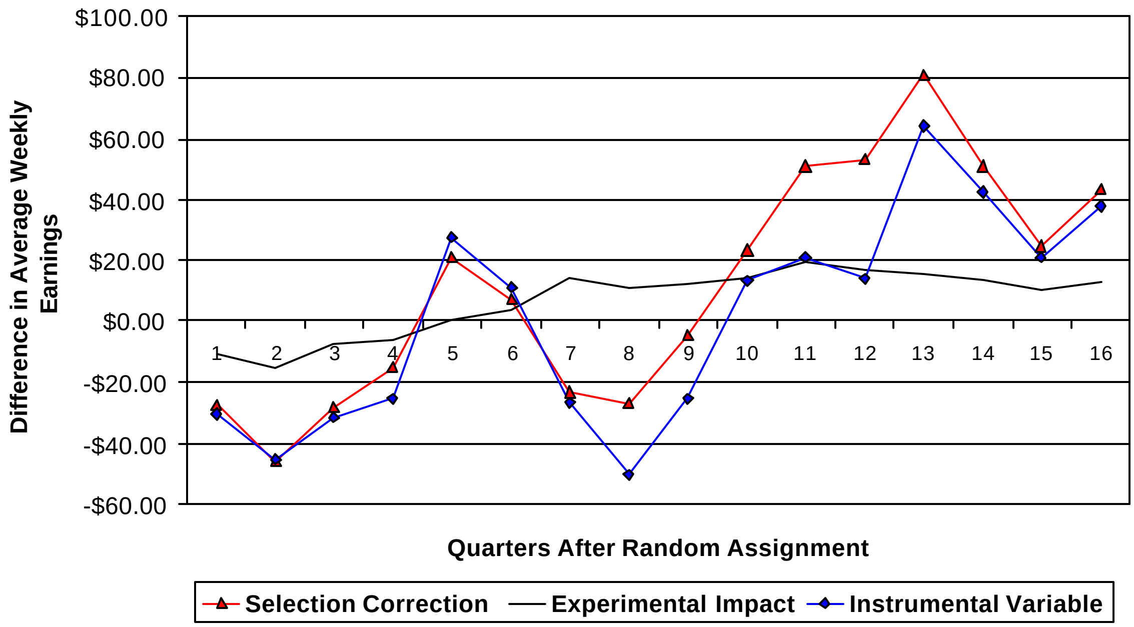


FIGURE A-13
Comparison of Overall Impact Estimates Derived From Experimental and Econometric Modeling Methods: Age 18 - 19

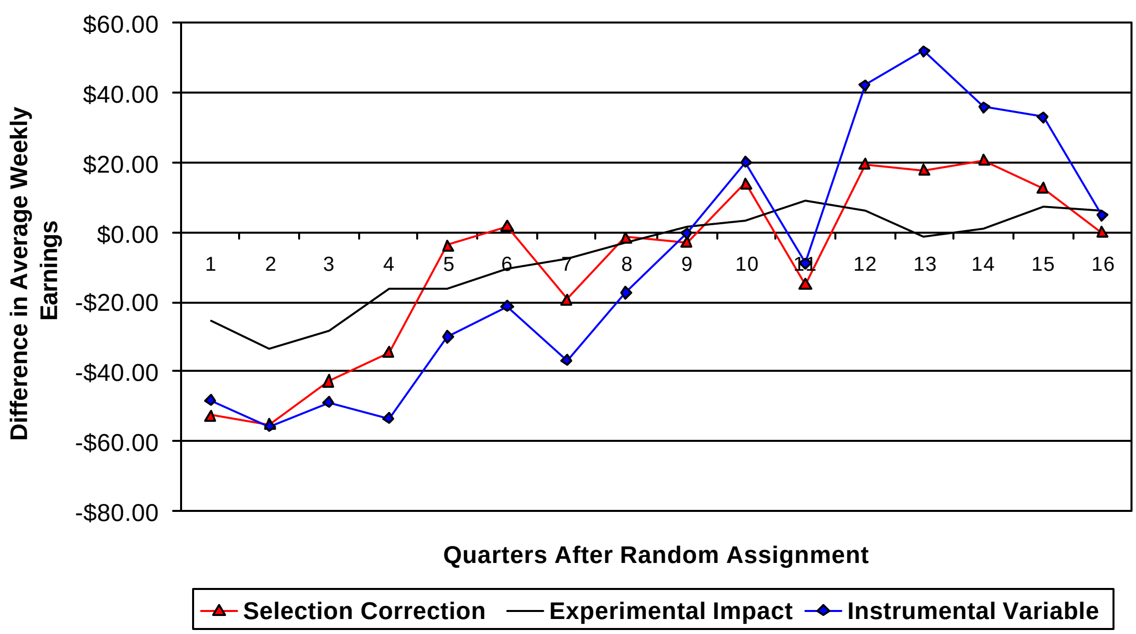


FIGURE A-14
Comparison of Overall Impact Estimates Derived From Experimental and Econometric Modeling Methods: Age 20 - 24

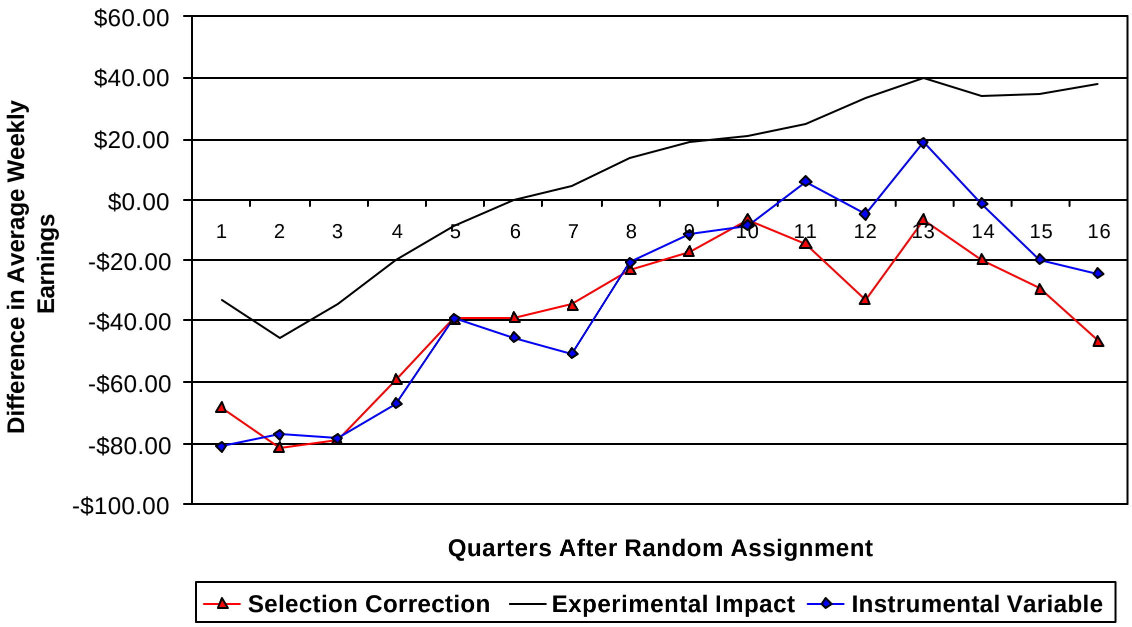


FIGURE A-15
Nearest-Neighbor Matching Impact Estimates for Vocational Completion: Age 16 - 17

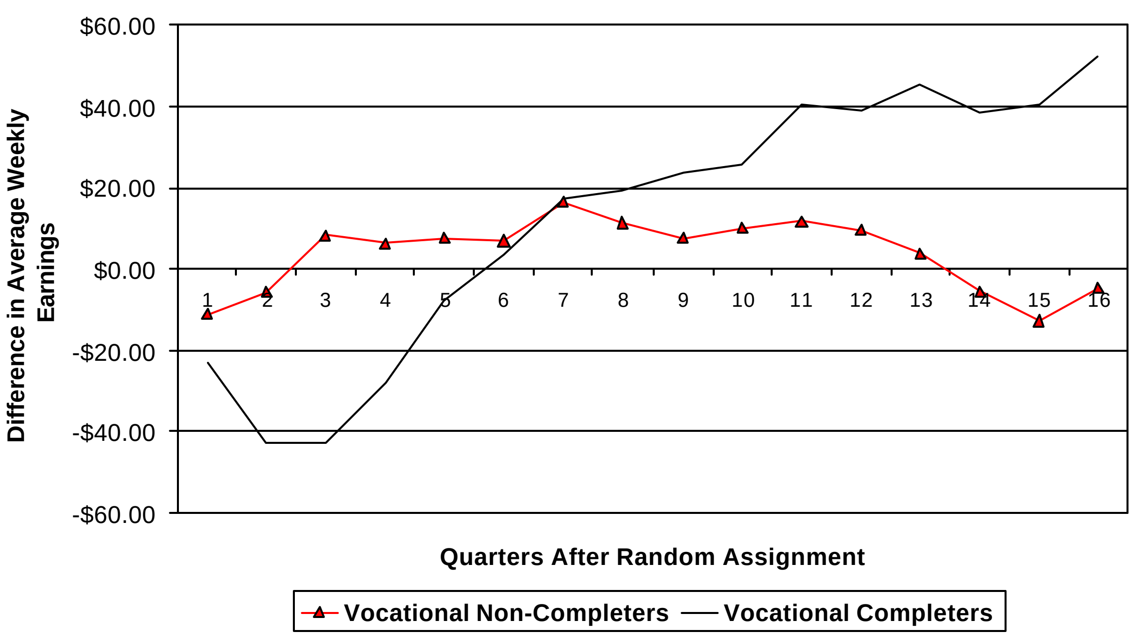


FIGURE A-16
Nearest-Neighbor Matching Impact Estimates for Vocational Completion: Age 18 - 19

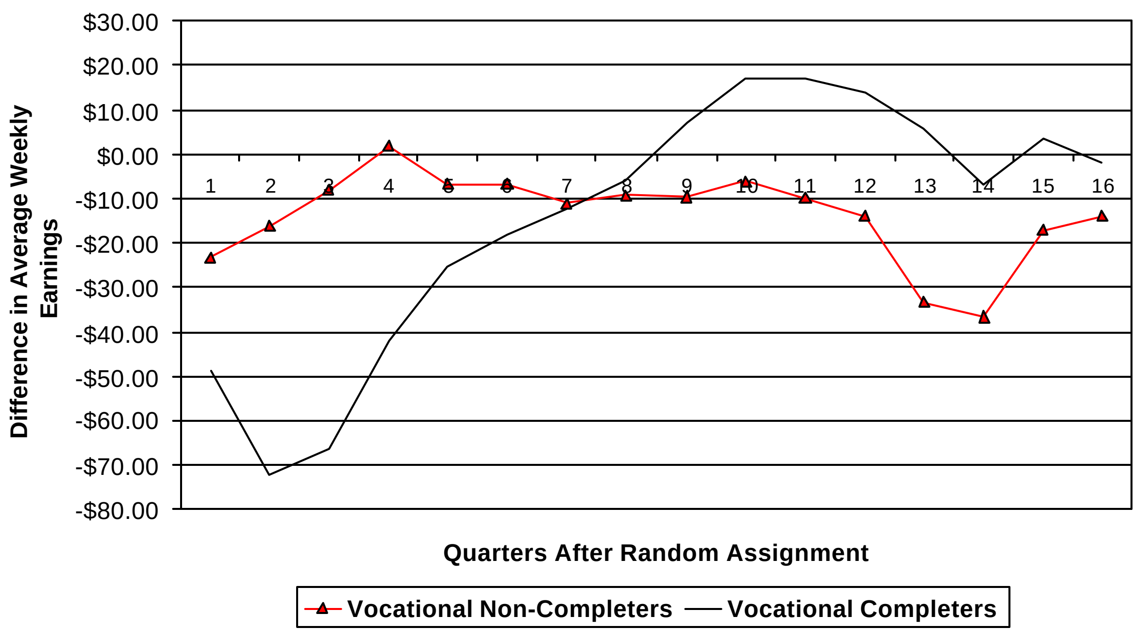


FIGURE A-17
Nearest-Neighbor Matching Impact Estimates for Vocational Completion: Age 20 - 24

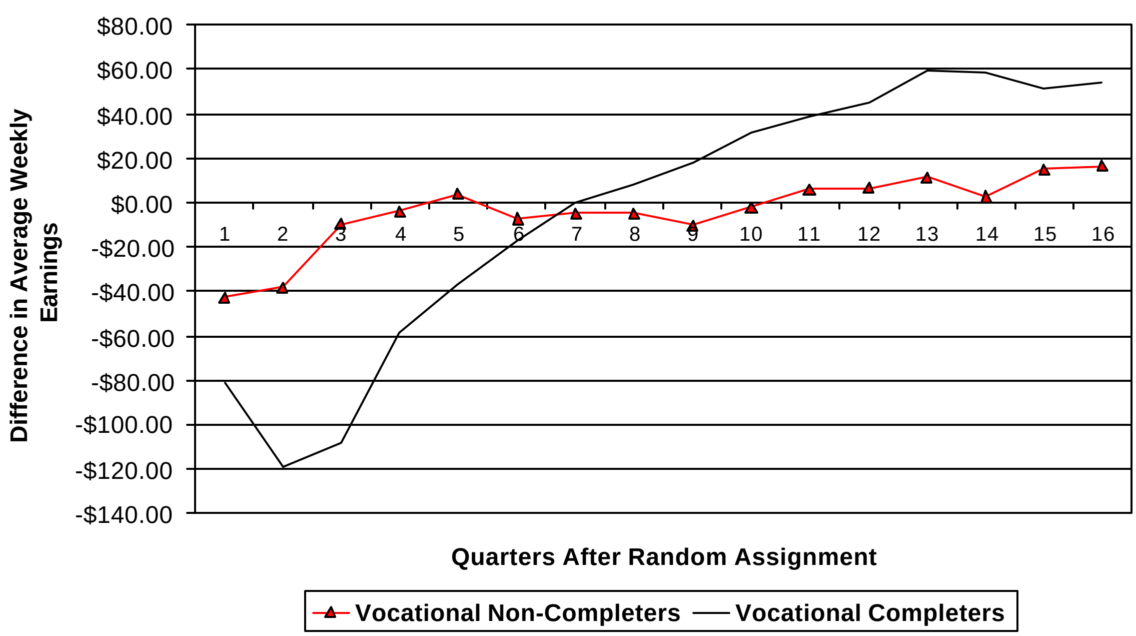


FIGURE A-18
Nearest-Neighbor Matching Impact Estimates for GED/Diploma Receipt: Age 16 - 17

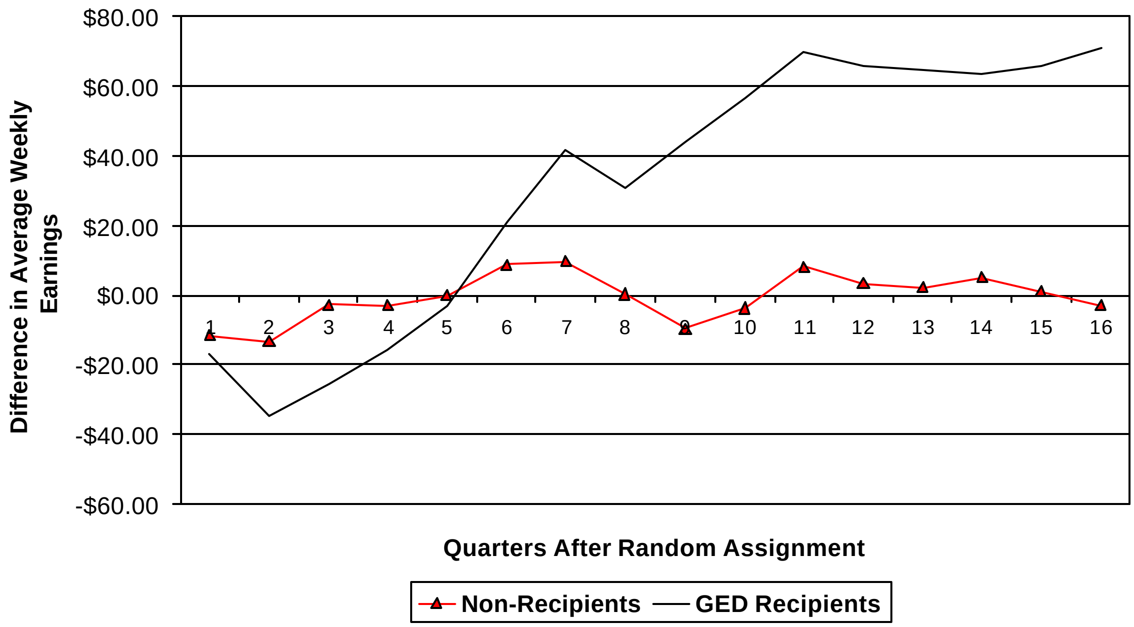


FIGURE A-19
Nearest-Neighbor Matching Impact Estimates for GED/Diploma Receipt: Age 18 - 19

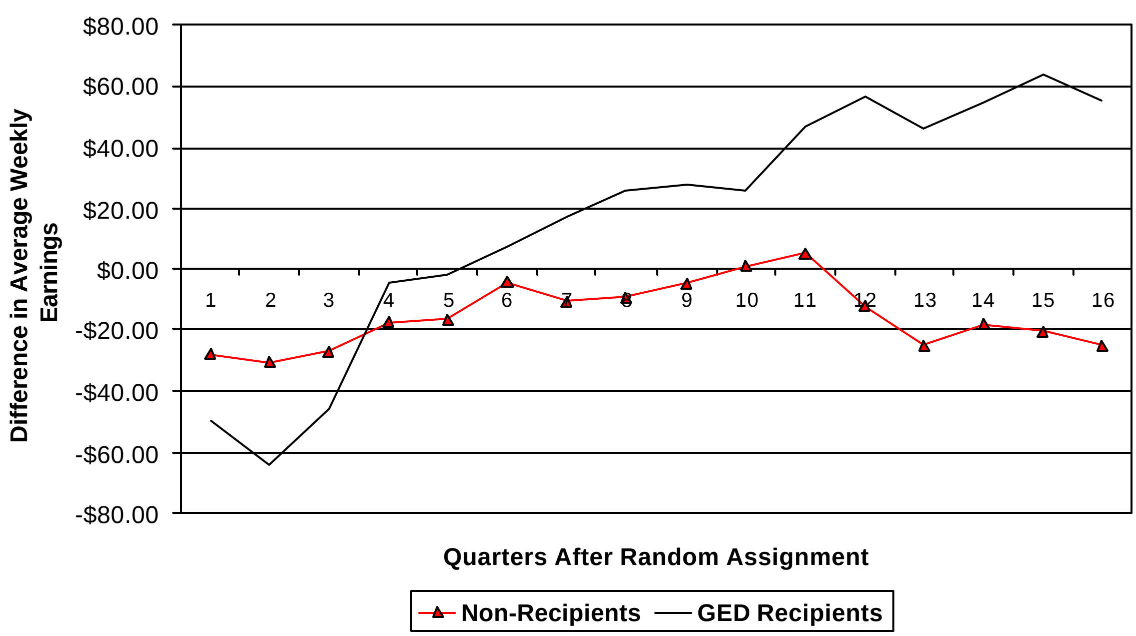


FIGURE A-20
Nearest-Neighbor Matching Impact Estimates for GED/Diploma Receipt: Age 20 - 24