1996

The Performance of Performance Standards: The Effects of JTPA Performance Standards on Efficiency, Equity and Participant Outcomes

James J. Heckman

Jeffrey Smith

POLITICAL
ECONOMY
RESEARCH
GROUP

PAPERS IN POLITICAL ECONOMY

Paper No. 73

"The Performance of Performance Standards: The Effects of JTPA Performance Standards on Efficiency, Equity and Participant Outcomes"

James J. Heckman and
Jeffrey Smith

The UNIVERSITY *of* WESTERN ONTARIO

The Political Economy Research Group was established in the faculty of Social Science at the University of Western Ontario in 1988. Its purpose is to foster scholarship, teaching and interdisciplinary research in political economy, with a focus on:

1. the application of economic models and methods to the study of political processes and institutions,
2. the economic impact of political processes and institutions,
3. the influence of economic factors on the formation of public policy and on institutional change,
4. the politics of economic policy making,
5. the political, social, and economic effects of public policy.

For further information:
Political Economy Research Group,
Department of Economics,
Social Science Centre,
London, Ontario, Canada N6A 5C2
phone: (519) 661-3877
fax: (519) 661-3292

ISSN: 1181-7151
ISBN: 0-7714-1893-0

The Performance of Performance Standards
The Effects of JTPA Performance Standards on Efficiency, Equity and Participant Outcomes

James J. Heckman

and

Jeffrey Smith

July 1995

Heckman is Henry Schultz Distinguished Service Professor in the Department of Economics and Director of the Center for Social Program Evaluation at the Irving B. Harris Graduate School of Public Policy Studies at the University of Chicago and a Fellow of the American Bar Foundation. Smith is an Assistant Professor in the Department of Economics at the University of Western Ontario. This research was supported by grants from the American Bar Foundation, the National Science Foundation (SBR-93-4115), the Russell Sage Foundation and the W.E. Upjohn Institute for Employment Research. We thank Karen Conneely and Carla VanBeselaere for excellent research assistance.

Abstract

The Performance of Performance Standards
The Effects of JTPA Performance Standards on Efficiency, Equity, and Participant Outcomes

Performance standards systems designed to direct bureaucratic behavior are widely viewed as a solution to the problem of inefficiency in government. Such systems aim to motivate bureaucrats to achieve program goals by rewarding them on the basis of observable short-run outcomes. This paper examines the JTPA performance standards system, which is widely regarded as a prototype for other government programs.

The greatest reservation about such reward systems is that they promote cream-skimming, the practice of serving the least disadvantaged among a program's target population. This paper clarifies the multiple definitions of cream-skimming used in the literature and shows that cream-skimming may be socially efficient. Evidence is presented that the cost of equitable allocation of program services is low because the gains from participation are close to uniform across skill groups. We also discuss the problem of using short term targets to motivate the achievement of long-run gains. We show that the short-run targets used in the JTPA system to monitor performance are only weakly, and sometimes perversely, related to longer-run impact measures. This represents a fundamental weakness of the JTPA performance management system and reveals a serious challenge to designers of performance standards systems.

JEL Numbers: L32, H40, I28

I. Introduction

Concerns about inefficiency in government have led to calls for bureaucratic reform from both ends of the political spectrum. When Senators Ted Kennedy and Dan Quayle redesigned and decentralized the federal job training bureaucracy with the Job Training Partnership Act (JTPA) of 1982, they incorporated a system of performance standards. These standards provide budgetary rewards and punishments to local job-training bureaucracies based on the measured labor market outcomes of the people they train. The hope was that this system would motivate local program administrators to use their newly acquired flexibility to raise participant earnings and employment in the long run.

More than a decade later, the JTPA performance standards system is widely viewed as a success, and as a prototype for incentive schemes in other government programs. The concept of rewarding government workers, administrators or offices on the basis of direct measures of bureaucratic output dominates current discussions of organizational reform in government. The "Reinventing Government" initiative of Vice President Albert Gore (1993), which culminated in the Performance Standards Act of 1993 (U.S. Congress, 1993), embraces the use of performance standards systems similar to that adopted in JTPA in many other government programs.

Despite the widespread popularity of performance standards systems in policy circles, almost nothing is actually known about their effects on bureaucratic behavior or on the gains achieved by program participants. In this paper, we use data from the recent National JTPA Study (see Larry Orr, et al., 1995) to examine the effects of the JTPA performance standards system along three important dimensions. First, we consider the effect of the system on the efficiency with which JTPA services are allocated. We address the question of how performance standards affect the extent to which JTPA services are allocated among program eligibles in ways that maximize the earnings gain per dollar spent. Second, we examine the distributional effects of the standards. In particular, we investigate their effect on the allocation of JTPA services to eligible persons with different levels of skill or economic disadvantage. Finally, we analyze how well, if at all, the JTPA performance standards system acts to motivate training centers to achieve the program's primary goal of increasing participants' long-term earnings.

Three ingredients are essential to an effective performance standards systems for job training programs, educational interventions or other activities expected to have long-term effects on participant outcomes. First, the performance measures or goals used in the system should be clearly defined and

susceptible to objective measurement at low cost. Since the aim of performance management systems is to motivate behavior in the short run, these goals or standards must inevitably be stated in terms of short-run targets. Thus, the designers of the JTPA performance standards systems face a difficult task. The goal of the program is to produce long-term increases in the earnings of participants relative to what they would have earned in the absence of training. Measuring long-term earnings gains takes time. However, in order to motivate behavior in the short run, a performance standards system must provide quick feedback. As a result, performance standards systems must rely on proxy measures to estimate the long-run gains to participation. The same is true of educational performance systems. In the case of JTPA and other similar programs, the proxy measures adopted consist of labor market outcomes such as employment and wage rates measured at or shortly after termination from the program. Put differently, the performance standards system uses short-term outcome levels in place of the long-term changes in outcomes that are the program's real objective.

Measuring performance by the outcome levels attained by trainees at, or shortly after, termination from the JTPA program raises the issue of cream-skimming. Cream-skimming occurs when JTPA program staff select persons into the program who are likely to do well in the labor market whether or not JTPA training benefits them. Selecting persons into the program based on short-run expected outcomes, or on the basis of observable characteristics that predict those outcomes, will increase training center performance as measured by the short-run performance of its trainees. In this paper, we show that the distributional and efficiency effects of cream skimming, and hence of a performance standards system that induces it, depend critically on the relationship between trainees' expected labor market outcomes in the absence of training and their expected benefits from participating in JTPA. If outcome levels and program benefits are positively related, and the costs of treatment are the same across skill levels, then cream-skimming increases efficiency by encouraging bureaucrats to focus resources on those who benefit most from JTPA training. The distributional effects of cream-skimming depend on this relationship as well, for it indicates the efficiency cost (or benefit) to concentrating JTPA services on the least well off -- called the "hard-to-serve" by JTPA administrators -- within the population eligible for JTPA.

Using experimental data from the National JTPA Study, we present evidence on the link between the gains to JTPA participation and the labor market outcomes that would occur in the absence of participation. For most groups, the gain appears to be roughly constant over much of the skill or advantage distribution within the current population of participants. Furthermore, only at the lowest end of the skill distribution does there appear to be an efficiency loss from targeting program resources away from the most employable. Thus, cream-skimming appears to produce modest efficiency gains.

The second ingredient required for a successful performance standards system is that it must motivate training centers to attain the standards. Pascal Courty and Gerald Marschke (1995a,b) and Carolyn Heinrich (1995) demonstrate that the JTPA performance system successfully motivates the bureaucrats governed by it. However, attainment of short-run targets does not ensure that the long-run goals of the program are achieved. If the short-run targets in a performance standards system do not predict the desired long-run impacts, then motivating program staff to meet them will not increase the program's attainment of social objectives. A performance standards system that adopts short-run performance measures unrelated to or negatively related to the true objectives of the program will eventually be called into question. Thus, the third key ingredient to an effective performance standards system is that the proxy measures used to gauge bureaucratic performance must be related to the long run goals of the program. For example, the use of test scores as a short-run performance measure in public schools appears to effectively motivate teachers and school systems to improve student test scores. Test scores clearly possess the first two ingredients just described. However, the success of educational testing as a motivational device has led to concerns that test score gains may result from teachers and administrators learning how to "teach to the test" rather than reflecting real gains in long-run student achievement (see Koretz, 1994). To the extent that teaching to the test takes up time and other educational resources that might otherwise be used to generate real gains in student achievement, "successful" educational performance systems based on test scores may actually reduce the sustained increases in student achievement that should be the schools' real goal.

In this paper, we present empirical evidence on the link between the short-term outcome measures set up as goals under the JTPA performance standards system and longer-run changes in participant earnings and employment. Our results indicate that there is little relationship between the two. In many cases, short-run outcomes appear to be negatively related to longer-run gains. These results have very strong implications for the effectiveness of the JTPA performance standards system. They suggest that too much attention has been paid to the issue of cream-skimming, and too little attention to the more important question of whether the bureaucratic activity induced by the performance standards system serves the end of increasing the earnings and employment of program participants in the long run.

The plan of the paper is as follows: In the next two sections, we describe the institutional structure of the JTPA performance standards system and the data from the National JTPA Study that we use in our analyses. In Section IV, we consider the efficiency and equity effects of the JTPA performance standards system, while in Section V we present evidence on the relationship between the short-run outcome measures used in the performance standards system and the long-run earnings and

employment gains experienced by program participants. In the concluding section, we summarize our results and indicate their implications for the effectiveness of the JTPA performance standards system and of other similar systems currently being proposed for application in many other government programs.

II. The JTPA Performance Standards System

In this section, we describe the structure of the JTPA performance standards system. This description provides an institutional context for the empirical work that follows.[1] The federal government, the states, and the more than 600 local JTPA training centers all play distinct roles in the system. The federal government defines a set of core outcome measures that serve as a basis for the state performance systems. These core measures have evolved somewhat over time, but consistently include employment rates, either at termination from JTPA or 13 weeks after, and average wage rates among those trainees who find employment. In the early years of the program (including the period when our data were collected), there was a cost standard, which measured average training center costs per trainee who obtained a positive outcome such as employment or enrollment in further school or training. This standard is no longer used. Each program year, the federal government defines target levels, or standards, for each outcome measure, and provides a regression adjustment model that allows states to adjust the target performance levels to reflect differences in economic conditions and participant characteristics across centers.

The individual states can adopt the federally-defined standards as offered, or else they can modify and augment them within fairly broad limits. Many states add additional standards that provide incentives to their training centers to serve particular sub-groups within the JTPA-eligible population. States also have discretion, again within limits, to define the "award function", the rule that determines the budgetary rewards each training center receives as a function of its performance relative to the standards. These functions vary widely across states along many dimensions, including the extent to which there are threshold levels below which there is no reward, the extent to which performance above the target levels leads to additional rewards, and the extent to which centers compete against each other for the awards. All of the state systems share the characteristic that a center is never worse off for increasing the average employment or wage rate among its trainees. For this reason, and because the employment and wage

[1] Longer descriptions of the JTPA performance standards system appear in Courty and Marschke (1995c), Katherine Dickinson, et al., (1988) and Janet Johnston (1987).

rate standards typically receive the greatest weight in determining awards, we focus primarily on these standards in our analysis.

The individual centers keep track of the labor market outcomes of their trainees, subject to state and federal reporting rules. At the end of each program year, states gather the performance measures for their training centers and determine the budgetary rewards they receive. Depending on the state, these rewards can vary from a few percent of a center's budget up to 20 or 30 percent. Centers value these funds both because they allow additional persons to be served but also because they can be used more flexibly than the centers' regular budget allocations. The administrators and staff of training centers that do well under the performance standards system receive additional benefits in the form of professional recognition within the JTPA community and political recognition in their local areas.

III. Data

The data we use were gathered as part of the National JTPA Study, a recent experimental evaluation of the JTPA program.[2] The experiment was conducted at 16 of the more than 600 training centers in the JTPA system. At these centers, persons who applied to and were accepted into the program were randomly assigned to either a treatment group allowed access to JTPA services or to a control group denied access to JTPA services for the next 18 months. Background information including demographic variables, educational attainment, work histories, indicators of previous training and of participation in government transfer programs and family income and composition were collected at the time of random assignment. Survey information on employment and earnings was collected around 18 months after random assignment and again for a sub-sample of the experimental group at around 30 months after random assignment.

IV. The Implications of Cream-Skimming for Efficiency and Equity in JTPA

The most frequent complaint directed against the JTPA performance standards system is that it leads to cream-skimming. In most cases, such charges are unaccompanied by a clear definition of what cream-skimming means, or by more than a cursory discussion of why cream-skimming is bad. Cream-

[2] The design of the experiment is described in Fred Doolittle and Linda Traeger (1990). The experimental impact estimates appear in Howard Bloom, et al., (1993) and Orr, et al., (1994).

skimming induced by the performance standards systems can affect both the efficiency with which JTPA program services are allocated and how they are distributed among persons with different levels of skill or advantage within the JTPA population.

In this section, we distinguish and make precise the different definitions of cream-skimming in the literature. We link these definitions to the competing goals embodied in the JTPA legislation and demonstrate the potential importance of cream-skimming for the attainment of these goals. Judging the importance of cream-skimming is an empirical issue. We establish the empirical relationship between the benefits from JTPA services and the location of JTPA clients in the overall skill distribution of current program participants. Our evidence indicates that the empirical importance of cream-skimming may be small relative to the attention it has received in the literature, because the benefits from JTPA participation do not differ much by skill level, except possibly at the lowest skill levels, where the impacts are lower.

1. Definitions of Cream-Skimming

The basic notion behind all the discussions of cream-skimming in the literature is that when cream-skimming occurs, "less deserving" persons receive JTPA services than would have received them in the absence of cream-skimming. In practice, "less deserving" often means persons who are less "hard-to-serve" because they possess some or all of the skills necessary to obtain employment at good wages in the absence of assistance from JTPA. It is important to note that all of the definitions of cream-skimming in the literature are conditional on eligibility for JTPA. As a result, even the "cream" -- those who are most advantaged within the eligible population -- live in families that either have low incomes, or receive government transfers, or both.

Three logically distinct concerns are combined and confused under the general rubric of "cream-skimming". The first is a concern about efficiency. This concern is motivated by the belief that the least advantaged within the eligible population are the most likely to benefit from JTPA services. If so, cream-skimming that reduces the participation of such persons in the program acts to reduce the net social benefit obtained from the services provided.[3]

The second concern is about equity. In this view, there is a normative preference for providing program services to the least advantaged, even within the already disadvantaged population eligible for

[3] This assumes that the costs of providing JTPA services are uniform across skill levels.

JTPA. This concern is independent of any efficiency concerns.

The third concern centers on discriminatory practices by program administrators in the admissions process that may lead to the under-representation of certain groups among JTPA participants, irrespective of where they lie in the distribution of skill or advantage. This concern is about fairness, or nondiscriminatory behavior. Regardless of whether efficiency or redistribution is adopted as the goal of JTPA, the concern is that some persons might be excluded from the program on the basis of irrelevant personal and group characteristics because of the wide latitude allowed to JTPA managers and case workers in making enrollment and placement decisions.

Stated in this way, it is clear that the importance of each of these concerns is fundamentally an empirical issue. In the first case, the efficiency effects of cream-skimming depend on the relationship between the benefits of JTPA services and the position of trainees within the distribution of skill or advantage of the JTPA-eligible population. If the (net) benefits of JTPA are positively related to skill or advantage, then cream-skimming increases efficiency, while if they are negatively related, then cream-skimming decreases efficiency. The equity concern depends indirectly on the same relationship, which indicates the efficiency costs of indulging normative preferences for serving those at the low end of the skill or advantage distribution. Finally, the third concern depends on patterns of participation conditional on eligibility for and application to JTPA. Those patterns are discussed at length in James Heckman and Jeffrey Smith (1995a,b).

2. The Conflicting Aims of JTPA

Like most acts of Congress, the legislation creating the Job Training Partnership Act (1982) expresses apparently contradictory goals. Some of the prose in the Act states that training sponsored under it should be "targeted toward the economically disadvantaged best able to use it." Other prose states that training should be given to "the most difficult to employ among the economically disadvantaged" (see Robert Guskind, 1988, Sar Levitan and Raymond Gallo, 1988, and Kathryn Anderson, Richard Burkhauser, Jennie Raymond and Clifford Russell, 1991). At issue are the two major arguments supporting government intervention in any area: first, that government programs should promote efficiency and maximize the social surplus and second, that government programs should redistribute resources toward the "truly needy" and toward those discriminated against by society at large.

The two clauses in the Act do not necessarily contradict each other. The goal of efficiency need not necessarily conflict with the goal of serving the least advantaged, though it is sometimes perceived

that it does. The greatest economic and social benefit may be achieved by targeting the least advantaged among the disadvantaged. For example, transferring resources to the most disadvantaged may counteract market imperfections that prevent them from undertaking efficient investments in human capital. In this case, the apparent contrast in the stated goals of the legislation actually represents two different ways of stating the same objective, because pursuit of distributional objectives has no cost (and perhaps even yields benefits) in terms of social output foregone. This is an instance of socially efficient redistribution.

Pointing out that the two goals need not conflict in theory does not mean that they do not conflict in fact. Whether or not they do depends on the empirical relationship between the gains to participation in JTPA and the trainee's position in the distribution of skill or advantage within the JTPA-eligible population. In this paper, we document that for certain groups, primarily adults, and especially adult women, things are not so simple. There is a genuine trade-off between efficiency and equity for these groups. One can redistribute by providing services to the most disadvantaged or one can maximize total social output by focusing training resources on the least disadvantaged, but not both. However, even for these groups, the efficiency loss from redistribution toward the least skilled is modest.

Given the empirical trade-off between the goal of efficiency and the goal of serving the most disadvantaged, it is worthwhile to briefly consider the value of concentrating JTPA resources on those at the low end of the skill distribution. For several reasons, such a focus may not make good policy sense.

The argument that JTPA should focus on the most disadvantaged among its eligible population tacitly assumes that JTPA is the only program available to help the poor. In fact, JTPA is far from being the only training program providing services to this group. As documented by the National Commission for Employment Policy (1995), numerous other employment and training programs exist to serve the disadvantaged. In many cases, these programs provide services better suited to very low-skill individuals than the classroom training in occupational skills and on-the-job training at private firms that are the primary offerings of JTPA. These JTPA services presume levels of basic skills or of employability that the most disadvantaged among the eligible often lack. In addition to numerous other training programs, there are also many income and in-kind transfer program devoted to meeting the needs of the poor. Failure by JTPA to train members of certain groups does not prevent them from being trained elsewhere, nor does it condemn them to neglect.

It is sometimes argued that training fosters the work ethic, and that work is valuable even if the investment in training is not economically efficient when measured in conventional terms. Social benefits may be high even when economic benefits as measured by participant earnings are low. This argument

is an elusive one since social benefits are intrinsically subjective and thus difficult to establish any consensus about. This argument is also incomplete. Even if society values having the most disadvantaged work because of its beneficial effect on their families and on their communities, it does not necessarily follow that training is the most effective way to promote work. As argued in Heckman, Rebecca Roselius and Smith (1994) and Edmund Phelps (1994), employment subsidies may be a more effective tool to promote work than training programs.

3. Cream-skimming and the JTPA Performance Standards System

Even if maximization of the net benefits to training were universally accepted as the goal of the JTPA program and there were no discrimination by program officials, there might still be concerns about the effects of cream-skimming engineered by the JTPA performance standards system. An ideal JTPA program that maximized net benefits would raise the earnings of trainees at the lowest possible social cost. The total budget allocated to this ideal JTPA program would be set to make the marginal social opportunity costs of funds (including income foregone) equal to the marginal social returns. Implementing this ideal program in practice requires information about two important counterfactuals, both of which are difficult to measure. The first counterfactual is what the earnings of participants would have been had they not participated. This is required to compute the gain in earnings resulting from participation. The second required counterfactual is how the resources devoted to training would otherwise be used. This counterfactual is needed to calculate the marginal social cost of achieving that gain. As shown in Heckman and Smith (1995c) and elsewhere, both counterfactuals are required to evaluate the program at the margin, yet both are in practice difficult to obtain.

Given the difficulty in directly measuring the economically appropriate constructs, it is natural to turn to alternative measures that are related to the desired constructs but are less difficult to measure. This was the course taken in the design of the JTPA performance standards system. As discussed earlier, in that system, levels of outcome measures computed at the time trainees are terminated from the program or ninety days after they are terminated are used in place of changes in the outcomes brought about by the program. Average direct program costs are used in place of marginal social costs. Training centers are rewarded for attaining the average outcome levels specified in the performance standards and not (directly) for the impacts that they have on participant outcomes. When they were in place, the cost standards component of the performance standards system rewarded centers for achieving these average outcome levels with low direct average costs, rather than at low marginal social costs.

There need not necessarily be any conflict between efficiency and maximization of performance under the JTPA performance standards system if outcome levels predict real gains from participation. But because gains and levels need not be positively related within the JTPA-eligible population, performance standards based on levels may cause persons to be admitted into the program who do not contribute to maximizing the social surplus. However, if gains and levels are negatively related, and if potential participants can estimate their expected gains from the program, then training centers will find it difficult to cream-skim in practice because persons with high levels will not want to participate in the program.

4. Evidence on the Relationship Between Gains and Levels

In this section, we use data from the National JTPA Study described earlier to examine how program impacts are related to outcomes in the absence of training and to initial skill levels. As discussed in Heckman (1992), Nancy Clements, Heckman and Smith (1993) and Heckman and Smith (1995d), social experiments do not solve the fundamental evaluation problem that one cannot observe the same person simultaneously in the treated (participant) and untreated (non-participant) states. Some assumption must be made that relates a person's outcome in the untreated state with that person's outcome in the treated state. The conventional approach is to assume that treatment effects are the same for everyone irrespective of their outcome in either state.

Figures 1A-1D and 2A-2D, based on an analysis by Clements, Heckman and Smith (1993), relax the assumption that participating in JTPA has the same effect for everyone but preserve one feature consistent with that assumption, namely that a person's rank in the distribution of outcomes in the treated state is the same as their rank in the distribution of outcomes in the untreated state. Under this assumption, the person with the best outcome in the treated distribution also has the best outcome in the untreated distribution, the person with the worst outcome in the treated distribution also has the worst outcome in the untreated distribution, and so on.

Under this assumption, it is possible to use experimental data to derive the gross gain from participating in the program for each rank in the treated or untreated outcome distributions. Using the experimental data from the National JTPA Study, Figures 1A-1D plot the estimated gain (or loss) in earnings in the 18 months after random assignment for persons at each percentile of the outcome distribution in the two states. Figures 2A-2D plot the same thing, but using earnings for the 30 months

after random assignment.[4] These figures relate program benefits to outcomes in the absence of training. For example, for the adult women (age 22 and over) in Figure 1B, there is no impact up to about the 20th percentile. The treatment effect is flat and positive over the interval from the 20th to the 90th percentile, after which there is a discernible increase in the estimated impact in the final decile. This evidence suggests that among adult women, the greatest impacts are for those who would do well even without training, while the smallest impacts are for those who would do poorly without training. In between is a broad range over which the impact of the program is roughly constant. The 30 month evidence in Figure 2B is similar, with the exception of a dip in the gains around the 85th percentile.[5]

The increase in the impacts in the last decile does not necessarily mean that it is efficient to screen those persons into the program who would do the best without training. Their social opportunity costs may be higher than those for other groups, as they will earn more even if they don't participate. What Figure 1B reveals is that the gains are modest and roughly constant over a broad range of untreated outcomes, and that cream-skimming past the 20th percentile probably contributes little to efficiency. Past the 20th percentile, the pursuit of distributional goals is not likely to entail substantial efficiency costs for this group. However, a policy of targeting services at the bottom two deciles would likely entail considerable efficiency costs.

With the exception of the graphs for young men, the figures for the other groups tell a similar tale. For the adult men in Figure 1A, the curve is flat over the range from the 10th to the 50th percentile, after which it takes a dip and then begins to rise. This pattern is only little changed at 30 months. For young men, the curve is flat, but the estimated impacts are negative. For them the least harm comes from targeting those who would do worst in the absence of training. Only for this group do distributional objectives comport with efficiency considerations. A better policy, however, would be to deny services to male youth entirely. For young women, the curve is very flat in both Figures 1D and 2D, indicating no effect of the program and little scope for cream-skimming to affect the social surplus one way or the other.

Tables 1 and 2 present experimental estimates of the mean impact of JTPA participation on total earnings and employment in either the 18 or the 30 months after random assignment for the four

---

[4] The samples differ somewhat between the two sets of figures. The samples for the 18 month graphs are the same as those used in Bloom, et al., (1993). The sample for Figures 2A-2D consists of all persons with valid self-reported earnings values for the 30 months after random assignment.

[5] In work not reported here, we find that the difference between the 18 and 30 month patterns is not an artifact of the change in sample composition.

demographic groups just discussed.[6] The results are presented conditional on a variety of characteristics related to human capital and employability. For each characteristic, we include the p-value from a test of the null hypothesis of equal impacts across subgroups.

Results for adult males appear in Tables 1A and 2A. In general, the estimated impacts of JTPA are not statistically distinguishable for subgroups based on labor force status, education, receipt of Aid to Families with Dependent Children (AFDC) or time since most recent employment, each measured at the time of random assignment. The sole exception is the 30 month earnings impacts conditional on labor force status at random assignment. Here we find that men employed at random assignment have a higher earnings impact than those out of the labor force at that time.

Impact estimates for adult women appear in Tables 1B and 2B. At 18 months there are no statistically distinguishable subgroup differences in the estimated impact of JTPA on earnings or employment. At 30 months, there are statistically significant differences by AFDC receipt and time since most recent employment in both earnings and employment impacts. The AFDC estimates indicate a larger program impact for persons receiving AFDC at random assignment. The pattern of estimated impacts for time since most recent employment is not monotone, but the largest point estimates are for persons with some employment experience who had been out of work at least nine months as of random assignment.[7]

The subgroup impact estimates indicate only small differences across skill levels in returns to the program. For adult men, there is some evidence that impacts are positively related to initial employment, suggesting a positive relation between gains and levels. For adult women there is some evidence that impacts are positively related to AFDC receipt, which suggests a negative relation between gains and levels for this group. However, the overall impression is one of little difference in impacts across subgroups of the participant population as a function of their observable earnings-related characteristics.

V. The Relationship Between Short-Run Outcome Measures and Long-Run Gains

In this section, we address the question of whether short-run performance measures of the type used in the JTPA performance standards system predict long-run improvements in participant earnings.

---

[6] See the Appendix for a detailed discussion of the construction of the estimates.

[7] We do not present estimates for youth because the sample sizes for the youth groups are too small for an analysis based on subgroups.

Increasing the earnings of participants is the primary goal of the JTPA program. Because earnings impacts are difficult and expensive to measure directly, the performance standards system relies on alternative performance measures that are less difficult to obtain. As previously noted, these measures consist of levels of earnings and employment at termination from the program or shortly thereafter. In order for these alternative measures to guide training centers toward the true objective of earnings maximization, they must have a positive relationship with later earnings gains. If they do not, then the evidence in Courty and Marschke (1995a,b) and Heinrich (1995) that the performance standards alter bureaucratic behavior indicates only that the standards have led program bureaucrats to take actions that will not increase JTPA's effect on participant earnings.

Using the experimental data from the recent National JTPA Study, we now present evidence on the link between short-run outcome measures like those in the JTPA performance standards system and estimates of the longer term effect of the program on the earnings of participants. We find that the short-run measures used in the performance standards system are at best weakly related to the longer-run measures. In many cases, the short-run measures are negatively related to longer-run effects on earnings. The disconnection between the short-run goals embodied in the performance standards system and the program's long-run goal of increasing the earnings of its participants calls into question the usefulness of the current JTPA performance management system.

Tables 3 and 4 present estimates of the relationship between experimental earnings and employment impact estimates and various short-term outcomes measured at selected dates after random assignment.[8] These estimates are obtained from regressions using experimental impact estimates and estimates of the short-run measures for overlapping subsets of the experimental sample. These subsets are defined by characteristics such as training center, race, level of education and level of recent employment experience. The regression is optimally weighted using the estimated covariance matrix from the estimation of the subgroup impacts.

Given the evidence presented earlier that impacts are generally uniform across skill levels, we report only aggregate estimates that do not condition on initial skill levels. The columns in the tables present the various impact measures. The rows present various performance measures stated in terms of levels that can be measured by the training center and their monitoring organizations. The first panel of each table includes only measures that are actually used in the JTPA performance standards system. The second panel of each table includes other measures based on outcomes at the time of random

[8] See the Appendix for a detailed discussion of the construction of the estimates.

assignment, termination from the program, or 13 weeks after termination. These measures are for comparison purposes only; they are not used in the JTPA performance standards system. The A versions of the tables refer to adult males. The B versions of the tables refer to adult females. As above, we do not present results for youth as the samples sizes for the youth groups are too small for an analysis based on subgroups.

The two columns in Tables 3 and 4 correspond to cumulated earnings or employment gains over the eighteen and thirty month intervals. Each cell of these tables presents the regression coefficient associated with the column's dependent variable and the row's independent variable, the estimated (robust) standard error of the coefficient, the p-value from a test of the null hypothesis that the true coefficient is zero and the R^2 for the regression. The constant from the regression is omitted to reduce clutter.

Thus, for example, the first row of the first column of Table 3A reveals that a regression of cumulative earnings over the 18 months after random assignment on the hourly wage at termination from the JTPA program produces a regression coefficient on hourly wage of $465.41, with a standard error of $395, a p-value of 0.2452 and an overall R^2 of 0.0328.

Four striking findings emerge from Tables 3A and 4A for adult men. First, the R^2 values are quite low. The short-run performance standards measures are only weakly related to the long-run earnings and employment gains produced by the program. Second, moving from wage measures at termination to "longer-run" follow-up measures at three months after termination usually worsens the relationship between the performance standard measure and the longer-run earnings or employment impacts. Third, the performance measures often do worse at predicting the longer-run earnings gains cumulated over 30 months than at predicting earnings gains cumulated over only the first 18 months after random assignment. Fourth, the relationship between some of the performance measures and the experimental impacts is negative. The only evidence supporting the efficacy of short-run outcome measures is the link between employment at follow-up and earnings, which is positive at 18 months and positive and marginally statistically significant at 30 months.

Tables 3B and 4B reveal a similar pattern of results for adult women, except that the estimated relationships are almost all negative. Most of the performance measures are weakly negatively related to cumulated earnings and employment impacts measured over 18 and 30 months. The success of the employment at follow-up measure found for adult males does not recur for adult females; for women, the relationship is negative. In many cases, the short-term outcome measures held up as goals under the JTPA performance standards system are perversely related to the longer-run participant earnings and

employment gains which are the program's primary objective.

VI. Conclusions

Performance standards systems that attempt to motivate bureaucratic behavior by rewarding government workers, administrators or offices on the basis of short-run outcome measures are widely viewed as a solution to the problem of inefficiency in government, despite the absence of any strong evidence that such standards lead bureaucrats to increase their attainment of long-run program goals. In this paper, we examine the effects of the JTPA performance standards system -- widely cited as an example for other programs to follow -- on the efficiency and equity with which program services are distributed among the eligible, and on the long run earnings and employment gains attained by program participants.

The JTPA performance standards system must of necessity rely on short-run outcomes measured in terms of levels as proxies for the longer run changes in outcomes that the program is intended to produce. This focus on short-run outcomes gives bureaucrats an incentive to cream-skim, by admitting persons likely to find employment at good wages in the short run whether or not they benefit from JTPA services. The problem of cream-skimming by JTPA administrators has attracted a lot of attention in the literature and in policy circles. We reduce the confusion surrounding cream-skimming as it is presently discussed. By making clear the multiple definitions of cream-skimming used in the literature, we show the fundamentally empirical nature of the problem. Only with evidence about the relationship between outcome levels and program gains can the efficiency costs (or benefits) associated with cream-skimming be established, and only with such evidence can the efficiency costs of concentrating program services on the least disadvantaged among those eligible for JTPA be judged.

We present empirical evidence from the recent experimental evaluation of JTPA that suggests that for adults and for female youth, there is little efficiency loss associated with cream-skimming, because the relationship between outcome levels and the benefits of program participation is generally flat, except at the low end of the skill distribution. This also means that for these groups, satisfying distributional concerns by targeting services to those at the middle of the skill distribution of current participants would result in little, if any, loss in efficiency. For male youth, the effect of the program is negative overall, but least negative for the least advantaged. For this group, efficiency and distributional goals coincide.

Empirically more important than the efficiency and distributional effects of cream-skimming induced by the performance standards system is its effects on the longer-term earnings and employment

gains experienced by participants. A positive link between the short-run measures employed in a performance standards system and the longer-term effects that are the true goals of the program is a critical prerequisite to the effectiveness of a performance management system. The weak and often perverse relationship between the short-run outcome measures used in the JTPA performance standards system and longer-run measures of the impact of the program on participant earnings and employment reveals a fundamental weakness of the JTPA performance management system. Even if the system motivates training centers to attain its short-run goals, the evidence presented here indicates that the attainment of these short-run goals has no beneficial effects on participant earnings and employment.

The evidence presented in this paper shows that concerns about cream-skimming in JTPA are exaggerated. Instead, critical attention needs to be devoted to the relationship between the performance standards system and the earnings enhancement goals of the JTPA program. We find the link to be weak or even negative, which suggests that the JTPA performance system is at best ineffective and at worst is reducing the earnings gains achieved by the program's economically disadvantaged clientele.

References

Anderson, Kathryn, Richard Burkhauser, Jennie Raymond and Clifford Russell. "Mixed Signals in the Job Training Partnership Act," *Growth and Change*, 1991, *22(3)*, pp. 32-48.

Bloom, Howard. "Accounting for No-shows in Experimental Evaluation Designs," *Evaluation Review*, 1984, *8(2)*, pp. 225-246.

Bloom, Howard, Larry Orr, George Cave, Steve Bell and Fred Doolittle. *The National JTPA Study: Title IIA Impacts on Earnings and Employment at 18 Months*. Bethesda, MD: Abt Associates, 1993.

Clements, Nancy, James Heckman and Jeffrey Smith. "Making The Most Out of Social Experiments: Reducing the Intrinsic Uncertainty in Evidence from Randomized Trials with an Application to the National JTPA Experiment." Unpublished manuscript, University of Chicago, 1993.

Courty, Pascal and Gerald Marschke. "Do Incentives Motivate Organizations? An Empirical Test." Unpublished manuscript, University of Chicago, 1995a.

Courty, Pascal and Gerald Marschke. "Moral Hazard Under Incentive Systems: The Case of a Federal Bureaucracy." Unpublished manuscript, University of Chicago, 1995b.

Courty, Pascal and Gerald Marschke. "The JTPA Incentive System." unpublished manuscript, University of Chicago, 1995c.

Dickinson, Katherine, Richard West, Deborah Kagan, David Drury, Marlene Franks, Laura Schlichtmann and Mary Vencill. "Evaluation of the Effects of JTPA Performance Standards on Clients, Services and Costs." Research Report 88-16, National Commission for Employment Policy, September 1988.

Doolittle, Fred and Linda Traeger. *Implementing the National JTPA Study*. New York: Manpower Demonstration Research Corporation, 1990.

Gore, Albert. *Creating A Government that Works Better and Costs Less: Report of the National Performance Review*. Washington, DC: U.S. Government Printing Office, 1993.

Guskind, Robert. "Cheers and Bronx Cheers for the JOBS Law," *National Journal*, 1988, *20*, pp. 2407-2408.

Heckman, James. "Randomization and Social Program Evaluation", in Charles Manski and Irwin Garfinkel, eds., *Evaluating Welfare and Training Programs*, Cambridge, MA: Harvard University Press, 1992.

Heckman, James, Rebecca Roselius and Jeffrey Smith. "U.S. Education and Training Policy: A Re-Evaluation of the Underlying Assumptions Behind the 'New Consensus'," in Lewis Solomon and Alec Levenson, eds., *Labor Markets, Employment Policy, and Job Creation*, Boulder, CO: Westview Press, 1994.

Heckman, James and Jeffrey Smith. "Ashenfelter's Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies." Unpublished manuscript, University of Chicago, 1995a.

Heckman, James and Jeffrey Smith. "The Determinants of Participation in a Social Program: Evidence from JTPA." Unpublished manuscript, University of Chicago, 1995b.

Heckman, James and Jeffrey Smith. "Evaluating the Welfare State." Unpublished manuscript, University of Chicago, 1995c.

Heckman, James and Jeffrey Smith. "Assessing The Case For Randomized Evaluation of Social Programs", *Journal of Economic Perspectives*, 1995d, *9(2)*, pp. 85-110.

Heinrich, Carolyn. "A Case Study of the JTPA Participant Selection and Service Assignment Processes and the Intervening Influences of Performance Standards and Other Administrative Practices." Unpublished manuscript, University of Chicago, 1995.

Johnston, Janet. *The Job Training Partnership Act: A Report by the National Commission for Employment Policy*. Washington, DC: U.S. Government Printing Office, 1987.

Koretz, Daniel. *The Evolution of a Portfolio Program: The Impact and Quality of the Vermont Program in its Second Year*. Los Angeles, CA: National Center for Research on Evaluation, Standards and Student Testing (CRESST), 1994.

Levitan, Sar and Raymond Gallo. *A Second Chance: Training for Jobs*, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1988.

National Commission for Employment Policy. *Understanding Federal Training and Employment Programs*. Washington, DC: National Commission for Employment Policy, 1995.

Orr, Larry, Howard Bloom, Stephen Bell, Winston Lin, George Cave and Fred Doolittle. *The National JTPA Study: Impacts, Benefits, and Costs of Title IIA*, Bethesda, MD: Abt Associates, 1994.

Phelps, Edmund. *A Program of Low-Wage Employment Subsidies*, New York: Russell Sage Foundation, 1994.

U.S. Congress. *Government Performance and Results Act*, 1993.

TABLE 1A
EXPERIMENTAL EARNINGS IMPACT ESTIMATES BY SUBGROUP
Adult Males

Group	Earnings Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
Labor Force Status		
P-value for test of equal 18-month impacts across subgroups: 0.0407		
P-value for test of equal 30-month impacts across subgroups: 0.3469		
Employed	2839.24 (1145.51) p = 0.0132	6328.20 (4143.22) p = 0.1271
Unemployed	718.84 (710.16) p = 0.3115	3021.68 (2339.51) p = 0.1969
Out of the Labor Force	-2193.85 (1658.81) p = 0.1861	-2725.72 (4693.28) p = 0.5616
Education		
P-value for test of equal 18-month impacts across subgroups: 0.6077		
P-value for test of equal 30-month impacts across subgroups: 0.7939		
Highest grade completed < 10	680.26 (1193.62) p = 0.5688	1713.46 (3935.62) p = 0.6634
Highest grade completed 10-11	-64.77 (1020.79) p = 0.9494	-270.18 (3516.67) p = 0.9388
Highest grade completed 12	1438.13 (793.68) p = 0.0701	552.70 (2729.26) p = 0.8396
Highest grade completed > 12	-92.00 (1238.21) p = 0.9408	4886.81 (4155.34) p = 0.2399

1. National JTPA Study data used in estimation.
2. Monthly earnings is based on self-reports with top 1% trimming.
3. Least squares specification with Bloom estimator of earnings impact.
4. Earnings impacts are calculated using all sample members with valid observations for self-reported monthly earnings during each period.
5. The sample includes 4886 valid observations for the 18-month period after random assignment and 1147 valid observations for the 30-month period after random assignment.
6. Robust standard errors are used.
7. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.
8. Run on July 19, 1995.

TABLE 1A - Continued
EXPERIMENTAL EARNINGS IMPACTS ESTIMATES BY SUBGROUP
Adult Males

Group	Earnings Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
AFDC Receipt at Random Assignment		
P-value for test of equal 18-month impacts across subgroups: 0.5948		
P-value for test of equal 30-month impacts across subgroups: 0.5794		
Not Receiving AFDC	722.73 (556.43) p = 0.1941	2933.22 (1810.58) p = 0.1055
Receiving AFDC	-232.18 (1706.56) p = 0.8918	-274.82 (5495.50) p = 0.9601
Recent Employment		
P-value for test of equal 18-month impacts across subgroups: 0.5995		
P-value for test of equal 30-month impacts across subgroups: 0.6193		
Currently employed	2668.20 (1230.61) p = 0.0302	3053.96 (4174.11) p = 0.4646
Last employed 0-2 months ago	816.36 (1091.14) p = 0.4544	6126.54 (3637.23) p = 0.0924
Last employed 3-5 months ago	-425.61 (1162.99) p = 0.7144	1248.64 (3794.83) p = 0.7422
Last employed 6-8 months ago	-5.65 (1824.51) p = 0.9975	-790.27 (5453.91) p = 0.8848
Last employed 9-11 months ago	1191.58 (2328.58) p = 0.6089	-4914.81 (7657.02) p = 0.5211
Last employed ≥ 12 months ago	525.44 (1333.79) p = 0.6936	3885.63 (4722.38) p = 0.4108
Never employed	-799.52 (1606.04) p = 0.6186	-6377.68 (6242.27) p = 0.3072

1. National JTPA Study data used in estimation.
2. Monthly earnings is based on self-reports with top 1% trimming.
3. Least squares specification with Bloom estimator of earnings impact.
4. Earnings impacts are calculated using all sample members with valid observations for self-reported monthly earnings during each period.
5. The sample includes 4886 valid observations for the 18-month period after random assignment and 1147 valid observations for the 30-month period after random assignment.
6. Robust standard errors are used.
7. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.
8. Run on July 19, 1995.

TABLE 1B
EXPERIMENTAL EARNINGS IMPACT ESTIMATES BY SUBGROUP
Adult Females

Group	Earnings Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
Labor Force Status		
P-value for test of equal 18-month impacts across subgroups: 0.3919		
P-value for test of equal 30-month impacts across subgroups: 0.5745		
Employed	1223.78 (651.64) p = 0.0605	1487.38 (2461.08) p = 0.5458
Unemployed	507.42 (507.92) p = 0.3178	428.84 (1715.10) p = 0.8026
Out of the Labor Force	1543.72 (601.48) p = 0.0103	3274.29 (2089.21) p = 0.1174
Education		
P-value for test of equal 18-month impacts across subgroups: 0.6890		
P-value for test of equal 30-month impacts across subgroups: 0.4641		
Highest grade completed < 10	1029.22 (643.40) p = 0.1097	-2227.56 (2577.38) p = 0.3876
Highest grade completed 10-11	1341.37 (592.06) p = 0.0235	3088.46 (2179.51) p = 0.1567
Highest grade completed 12	460.29 (469.73) p = 0.3272	1503.23 (1711.16) p = 0.3799
Highest grade completed > 12	971.20 (816.54) p = 0.2343	795.14 (2997.34) p = 0.7908

1. National JTPA Study data used in estimation.
2. Monthly earnings is based on self-reports with top 1% trimming.
3. Least squares specification with Bloom estimator of earnings impact.
4. Earnings impacts are calculated using all sample members with valid observations for self-reported monthly earnings during each period.
5. The sample includes 6272 valid observations for the 18-month period after random assignment and 1273 valid observations for the 30-month period after random assignment.
6. Robust standard errors are used.
7. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.
8. Run on July 19, 1995.

TABLE 1B - Continued
EXPERIMENTAL EARNINGS IMPACT ESTIMATES BY SUBGROUP
Adult Females

Group	Earnings Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
AFDC Receipt at Random Assignment		
P-value for test of equal 18-month impacts across subgroups: 0.7224		
P-value for test of equal 30-month impacts across subgroups: 0.0371		
Not Receiving AFDC	712.26 (392.05) p = 0.0693	-947.01 (1462.17) p = 0.5173
Receiving AFDC	924.57 (451.07) p = 0.0404	3624.35 (1631.02) p = 0.0265
Recent Employment		
P-value for test of equal 18-month impacts across subgroups: 0.8614		
P-value for test of equal 30-month impacts across subgroups: 0.0492		
Currently employed	1104.08 (721.42) p = 0.1260	396.24 (2851.27) p = 0.8895
Last employed 0-2 months ago	594.01 (713.69) p = 0.4053	979.22 (2485.38) p = 0.6937
Last employed 3-5 months ago	171.44 (953.91) p = 0.8574	-7677.17 (3435.31) p = 0.0256
Last employed 6-8 months ago	1874.38 (1175.53) p = 0.1109	975.22 (3721.12) p = 0.7933
Last employed 9-11 months ago	1679.73 (1311.91) p = 0.2005	5244.59 (4437.63) p = 0.2375
Last employed ≥ 12 months ago	1304.36 (587.15) p = 0.0264	4919.73 (2020.46) p = 0.0151
Never employed	610.59 (609.42) p = 0.3164	-2490.44 (2763.46) p = 0.3677

1. National JTPA Study data used in estimation.
2. Monthly earnings is based on self-reports with top 1% trimming.
3. Least squares specification with Bloom estimator of earnings impact.
4. Earnings impacts are calculated using all sample members with valid observations for self-reported monthly earnings during each period.
5. The sample includes 6272 valid observations for the 18-month period after random assignment and 1273 valid observations for the 30-month period after random assignment.
6. Robust standard errors are used.
7. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.
8. Run on July 19, 1995.

TABLE 2A
EXPERIMENTAL EMPLOYMENT IMPACT ESTIMATES BY SUBGROUP
Adult Males

Group	Employment Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
Labor Force Status		
P-value for test of equal 18-month impacts across subgroups: 0.2679		
P-value for test of equal 30-month impacts across subgroups: 0.6517		
Employed	0.0300 (0.0166) p = 0.0703	0.0005 (0.0194) p = 0.9780
Unemployed	0.0056 (0.0105) p = 0.5920	0.0180 (0.0125) p = 0.1495
Out of the Labor Force	-0.0163 (0.0262) p = 0.5341	0.0289 (0.0281) p = 0.3042
Education		
P-value for test of equal 18-month impacts across subgroups: 0.9587		
P-value for test of equal 30-month impacts across subgroups: 0.7206		
Highest grade completed < 10	0.0114 (0.0203) p = 0.5736	0.0403 (0.0225) p = 0.0730
Highest grade completed 10-11	0.0120 (0.0163) p = 0.4639	0.0134 (0.0188) p = 0.4771
Highest grade completed 12	0.0030 (0.0119) p = 0.7995	0.0105 (0.0141) p = 0.4584
Highest grade completed > 12	0.0116 (0.0172) p = 0.5006	0.0201 (0.0221) p = 0.3626

1. National JTPA Study data used in estimation.

2. Least squares specification with Bloom estimator of employment impact.

3. Employment impacts are calculated based upon self-reports of non-zero earnings as a percentage of all sample members with valid observations for self-reported monthly earnings during each period.

4. The sample includes 4886 valid observations for the 18-month period after random assignment and 1147 valid observations for the 30-month period after random assignment.

5. Robust standard errors are used.

6. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.

7. Run on July 19, 1995.

TABLE 2A - Continued
EXPERIMENTAL EMPLOYMENT IMPACT ESTIMATES BY SUBGROUP
Adult Males

Group	Employment Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
AFDC Receipt at Random Assignment		
P-value for test of equal 18-month impacts across subgroups: 0.3813		
P-value for test of equal 30-month impacts across subgroups: 0.6678		
Not Receiving AFDC	0.0122 (0.0085) p = 0.1493	0.0161 (0.0099) p = 0.1027
Receiving AFDC	-0.0132 (0.0278) p = 0.6340	0.0306 (0.0322) p = 0.3423
Recent Employment		
P-value for test of equal 18-month impacts across subgroups: 0.9112		
P-value for test of equal 30-month impacts across subgroups: 0.7010		
Currently employed	0.0176 (0.0178) p = 0.3245	-0.0134 (0.0212) p = 0.5269
Last employed 0-2 months ago	0.0168 (0.0152) p = 0.2674	0.0205 (0.0180) p = 0.2531
Last employed 3-5 months ago	0.0037 (0.0176) p = 0.8346	0.0119 (0.0209) p = 0.5702
Last employed 6-8 months ago	-0.0135 (0.0256) p = 0.5969	0.0312 (0.0296) p = 0.2912
Last employed 9-11 months ago	0.0163 (0.0384) p = 0.6717	0.0098 (0.0478) p = 0.8376
Last employed ≥ 12 months ago	0.0284 (0.0224) p = 0.2041	0.0475 (0.0257) p = 0.0644
Never employed	0.0017 (0.0295) p = 0.9539	0.0145 (0.0319) p = 0.6484

1. National JTPA Study data used in estimation.

2. Least squares specification with Bloom estimator of employment impact.

3. Employment impacts are calculated based upon self-reports of non-zero earnings as a percentage of all sample members with valid observations for self-reported monthly earnings during each period.

4. The sample includes 4886 valid observations for the 18-month period after random assignment and 1147 valid observations for the 30-month period after random assignment.

5. Robust standard errors are used.

6. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.

7. Run on July 19, 1995.

TABLE 2B
EXPERIMENTAL EMPLOYMENT IMPACT ESTIMATES BY SUBGROUP
Adult Females

Group	Employment Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
Labor Force Status		
P-value for test of equal 18-month impacts across subgroups: 0.4715		
P-value for test of equal 30-month impacts across subgroups: 0.2286		
Employed	0.0017 (0.0135) p = 0.8967	-0.0158 (0.0168) p = 0.3493
Unemployed	0.0112 (0.0112) p = 0.3149	0.0184 (0.0128) p = 0.1504
Out of the Labor Force	0.0274 (0.0160) p = 0.0878	0.0184 (0.0188) p = 0.3280
Education		
P-value for test of equal 18-month impacts across subgroups: 0.8149		
P-value for test of equal 30-month impacts across subgroups: 0.4646		
Highest grade completed < 10	0.0135 (0.0164) p = 0.4087	0.0175 (0.0182) p = 0.3373
Highest grade completed 10-11	0.0289 (0.0147) p = 0.0493	0.0246 (0.0171) p = 0.1512
Highest grade completed 12	0.0129 (0.0109) p = 0.2390	-0.0053 (0.0129) p = 0.6803
Highest grade completed > 12	0.0115 (0.0172) p = 0.5062	0.0209 (0.0211) p = 0.3234

1. National JTPA Study data used in estimation.

2. Least squares specification with Bloom estimator of employment impact.

3. Employment impacts are calculated based upon self-reports of non-zero earnings as a percentage of all sample members with valid observations for self-reported monthly earnings during each period.

4. The sample includes 6272 valid observations for the 18-month period after random assignment and 1273 valid observations for the 30-month period after random assignment.

5. Robust standard errors are used.

6. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.

7. Run on July 19, 1995.

TABLE 2B - Continued
EXPERIMENTAL EMPLOYMENT IMPACT ESTIMATES BY SUBGROUP
Adult Females

Group	Employment Impacts Measured Over: 18-Month Period After Random Assignment	30-Month Period After Random Assignment
AFDC Receipt at Random Assignment		
P-value for test of equal 18-month impacts across subgroups: 0.0277		
P-value for test of equal 30-month impacts across subgroups: 0.2607		
Not Receiving AFDC	0.0028 (0.0087) p = 0.7507	0.0026 (0.0105) p = 0.8084
Receiving AFDC	0.0343 (0.0113) p = 0.0025	0.0211 (0.0127) p = 0.0973
Recent Employment		
P-value for test of equal 18-month impacts across subgroups: 0.5708		
P-value for test of equal 30-month impacts across subgroups: 0.0139		
Currently employed	0.0138 (0.0151) p = 0.3589	0.0056 (0.0197) p = 0.7751
Last employed 0-2 months ago	0.0099 (0.0161) p = 0.5365	0.0060 (0.0181) p = 0.7387
Last employed 3-5 months ago	-0.0063 (0.0199) p = 0.7514	-0.0589 (0.0220) p = 0.0074
Last employed 6-8 months ago	0.0451 (0.0263) p = 0.0860	0.0502 (0.0305) p = 0.0997
Last employed 9-11 months ago	0.0310 (0.0305) p = 0.3092	0.0636 (0.0382) p = 0.0958
Last employed ≥ 12 months ago	0.0341 (0.0155) p = 0.0285	0.0347 (0.0180) p = 0.0539
Never employed	0.0335 (0.0168) p = 0.0467	-0.0059 (0.0191) p = 0.7558

1. National JTPA Study data used in estimation.
2. Least squares specification with Bloom estimator of employment impact.
3. Employment impacts are calculated based upon self-reports of non-zero earnings as a percentage of all sample members with valid observations for self-reported monthly earnings during each period.
4. The sample includes 6272 valid observations for the 18-month period after random assignment and 1273 valid observations for the 30-month period after random assignment.
5. Robust standard errors are used.
6. P-values within cells are from tests of the null hypothesis that the impact is zero for each subgroup.
7. Run on July 19, 1995.

TABLE 3A
DETERMINANTS OF EARNINGS IMPACTS OF PARTICIPATION IN JTPA
Adult Males

	Earnings Impact Measured Over:	
Actual JTPA Performance Standards	18-Month Period After Random Assignment	30-Month Period After Random Assignment
Hourly Wage at Time of Termination	465.41 (394.76) p = 0.2452 R^2 = 0.0328	-1405.68 (1653.30) p = 0.4001 R^2 = 0.0173
Weekly Earnings at Time of Follow-up	6.74 (7.42) p = 0.3690 R^2 = 0.0197	-20.76 (31.79) p = 0.5174 R^2 = 0.0103
Employment at Time of Termination	2542.99 (1384.72) p = 0.0737 R^2 = 0.0778	3673.71 (5869.08) p = 0.5349 R^2 = 0.0097
Employment at Time of Follow-up	2579.24 (2486.91) p = 0.3058 R^2 = 0.0256	18716.00 (9842.28) p = 0.0643 R^2 = 0.0810

TABLE 3A
DETERMINANTS OF EARNINGS IMPACTS OF PARTICIPATION IN JTPA
Adult Males

	Earnings Impact Measured Over:	
Other Short-Run Measures	18-Month Period After Random Assignment	30-Month Period After Random Assignment
Hourly Wage at Time of Random Assignment	756.15 (367.69) p = 0.0461 R^2 = 0.0935	-1813.07 (1574.18) p = 0.2561 R^2 = 0.0313
Weekly Earnings at Time of Random Assignment	29.79 (10.62) p = 0.0076 R^2 = 0.1611	-2.69 (47.03) p = 0.9547 R^2 = 0.0001
Employment at Time of Random Assignment	3757.22 (2037.89) p = 0.0725 R^2 = 0.0766	18678.00 (8136.92) p = 0.0269 R^2 = 0.1139
Weekly Earnings at Time of Termination	3.72 (8.98) p = 0.6807 R^2 = 0.0042	5.12 (38.29) p = 0.8943 R^2 = 0.0004
Hourly wage at Time of Follow-up	239.67 (399.11) p = 0.5515 R^2 = 0.0087	-823.90 (1657.69) p = 0.6218 R^2 = 0.0060

1. National JTPA Study data used in estimation.

2. Actual JTPA Performance Measures are defined as follows:
Hourly Wage at Placement: the average wage at program termination for employed adults.
Weekly Earnings at Follow-up: the average weekly wage of adults who were employed 13 weeks after program termination.
Employment Rate at Placement: the fraction of adults employed at program termination.
Employment Rate at Follow-up: the fraction of adults who were employed 13 weeks after program termination.

3. In our analysis, employment rates were calculated based on the presence or absence of a job spell within 30 days of each reference date (random assignment, termination, or follow-up. Hourly wages were calculated based on the highest reported hourly wage for all job spells reported within 30 days of each reference date. Weekly earnings were calculated by averaging the product of hourly wages and hours worked per week across all reported job spells within 30 days of each reference date weighted by the fraction of the 30-day window spanned by each job spell.

4. Run on July 20, 1995.

TABLE 3B
DETERMINANTS OF EARNINGS IMPACTS OF PARTICIPATION IN JTPA
Adult Females

	Earnings Impact Measured Over:	
Actual JTPA Performance Standards	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Termination	-577.61 (304.00) p = 0.0645 R^2 = 0.0809	-1729.66 (1280.64) p = 0.1842 R^2 = 0.0426
Weekly Earnings at Time of Follow-up	-3.74 (8.78) p = 0.6726 R^2 = 0.0044	-12.05 (36.54) p = 0.7432 R^2 = 0.0026
Employment at Time of Termination	-117.72 (941.92) p = 0.9012 R^2 = 0.0004	-2065.61 (3928.63) p = 0.6019 R^2 = 0.0069
Employment at Time of Follow-up	1513.28 (1482.04) p = 0.3132 R^2 = 0.0248	-1873.03 (6236.83) p = 0.7655 R^2 = 0.0022

TABLE 3B
DETERMINANTS OF EARNINGS IMPACTS OF PARTICIPATION IN JTPA
Adult Females

	Earnings Impact Measured Over:	
Other Short-Run Measures	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Random Assignment	-688.90 (308.62) p = 0.0311 R^2 = 0.1084	-2845.35 (1321.78) p = 0.0373 R^2 = 0.1015
Weekly Earnings at Time of Random Assignment	-2.24 (12.42) p = 0.8578 R^2 = 0.0008	4.32 (51.45) p = 0.9336 R^2 = 0.0002
Employment at Time of Random Assignment	511.33 (1314.50) p = 0.6993 R^2 = 0.0037	-6250.53 (5417.38) p = 0.2553 R^2 = 0.0314
Weekly Earnings at Time of Termination	-9.20 (10.20) p = 0.3725 R^2 = 0.0194	-20.91 (42.35) p = 0.6240 R^2 = 0.0059
Hourly wage at Time of Follow-up	-312.24 (319.96) p = 0.3348 R^2 = 0.0227	-1283.42 (1316.65) p = 0.3354 R^2 = 0.0226

1. National JTPA Study data used in estimation.

2. Actual JTPA Performance Measures are defined as follows:
Hourly Wage at Placement: the average wage at program termination for employed adults.
Weekly Earnings at Follow-up: the average weekly wage of adults who were employed 13 weeks after program termination.
Employment Rate at Placement: the fraction of adults employed at program termination.
Employment Rate at Follow-up: the fraction of adults who were employed 13 weeks after program termination.

3. In our analysis, employment rates were calculated based on the presence or absence of a job spell within 30 days of each reference date (random assignment, termination, or follow-up. Hourly wages were calculated based on the highest reported hourly wage for all job spells reported within 30 days of each reference date. Weekly earnings were calculated by averaging the product of hourly wages and hours worked per week across all reported job spells within 30 days of each reference date weighted by the fraction of the 30-day window spanned by each job spell.

4. Run on July 20, 1995.

TABLE 4A
DETERMINANTS OF EMPLOYMENT IMPACTS OF PARTICIPATION IN JTPA
Adult Males

	Employment Impact Measured Over:	
Actual JTPA Performance Standards	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Termination	0.003 (0.005) p = 0.4914 R^2 = 0.0116	-0.005 (0.010) p = 0.6230 R^2 = 0.0059
Weekly Earnings at Time of Follow-up	0.000 (0.000) p = 0.9921 R^2 = 0.0000	-0.000 (0.000) p = 0.3274 R^2 = 0.0234
Employment at Time of Termination	0.005 (0.017) p = 0.7559 R^2 = 0.0024	-0.059 (0.034) p = 0.0850 R^2 = 0.0723
Employment at Time of Follow-up	0.050 (0.028) p = 0.0848 R^2 = 0.0707	0.021 (0.061) p = 0.7338 R^2 = 0.0029

TABLE 4A
DETERMINANTS OF EMPLOYMENT IMPACTS OF PARTICIPATION IN JTPA
Adult Males

	Employment Impact Measured Over:	
Other Short-Run Measures	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Random Assignment	0.003 (0.004) p = 0.4747 R^2 = 0.0125	0.004 (0.010) p = 0.7019 R^2 = 0.0036
Weekly Earnings at Time of Random Assignment	0.000 (0.000) p = 0.4983 R^2 = 0.0113	0.000 (0.000) p = 0.3073 R^2 = 0.0254
Employment at Time of Random Assignment	0.027 (0.024) p = 0.2668 R^2 = 0.0300	-0.032 (0.050) p = 0.5335 R^2 = 0.0095
Weekly Earnings at Time of Termination	-0.000 (0.000) p = 0.7201 R^2 = 0.0032	-0.000 (0.000) p = 0.9339 R^2 = 0.0002
Hourly wage at Time of Follow-up	0.001 (0.005) p = 0.7728 R^2 = 0.0021	-0.006 (0.010) p = 0.5626 R^2 = 0.0082

1. National JTPA Study data used in estimation.

2. Actual JTPA Performance Measures are defined as follows:
Hourly Wage at Placement: the average wage at program termination for employed adults.
Weekly Employment at Follow-up: the average weekly wage of adults who were employed 13 weeks after program termination.
Employment Rate at Placement: the fraction of adults employed at program termination.
Employment Rate at Follow-up: the fraction of adults who were employed 13 weeks after program termination.

3. In our analysis, employment rates were calculated based on the presence or absence of a job spell within 30 days of each reference date (random assignment, termination, or follow-up. Hourly wages were calculated based on the highest reported hourly wage for all job spells reported within 30 days of each reference date. Weekly earnings were calculated by averaging the product of hourly wages and hours worked per week across all reported job spells within 30 days of each reference date weighted by the fraction of the 30-day window spanned by each job spell.

4. Run on July 19, 1995.

TABLE 4B
DETERMINANTS OF EMPLOYMENT IMPACTS OF PARTICIPATION IN JTPA
Adult Females

	Employment Impact Measured Over:	
Actual JTPA Performance Standards	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Termination	-0.018 (0.008) p = 0.0202 R^2 = 0.1246	-0.010 (0.011) p = 0.3559 R^2 = 0.0208
Weekly Earnings at Time of Follow-up	-0.000 (0.000) p = 0.2728 R^2 = 0.0293	-0.000 (0.000) p = 0.3277 R^2 = 0.0234
Employment at Time of Termination	-0.023 (0.023) p = 0.3213 R^2 = 0.0246	-0.029 (0.033) p = 0.3767 R^2 = 0.0196
Employment at Time of Follow-up	-0.067 (0.037) p = 0.0767 R^2 = 0.0745	-0.024 (0.053) p = 0.6521 R^2 = 0.0050

TABLE 4B
DETERMINANTS OF EMPLOYMENT IMPACTS OF PARTICIPATION IN JTPA
Adult Females

	Employment Impact Measured Over:	
Other Short-Run Measures	**18-Month Period After Random Assignment**	**30-Month Period After Random Assignment**
Hourly Wage at Time of Random Assignment	-0.023 (0.008) p = 0.0055 R^2 = 0.1732	-0.012 (0.012) p = 0.3372 R^2 = 0.0225
Weekly Earnings at Time of Random Assignment	-0.000 (0.000) p = 0.2653 R^2 = 0.0302	-0.000 (0.000) p = 0.6742 R^2 = 0.0044
Employment at Time of Random Assignment	-0.075 (0.031) p = 0.0214 R^2 = 0.1225	-0.050 (0.046) p = 0.2898 R^2 = 0.0273
Weekly Earnings at Time of Termination	-0.001 (0.000) p = 0.0177 R^2 = 0.1296	-0.001 (0.000) p = 0.1511 R^2 = 0.0496
Hourly wage at Time of Follow-up	-0.008 (0.008) p = 0.3382 R^2 = 0.0224	-0.008 (0.011) p = 0.4914 R^2 = 0.0116

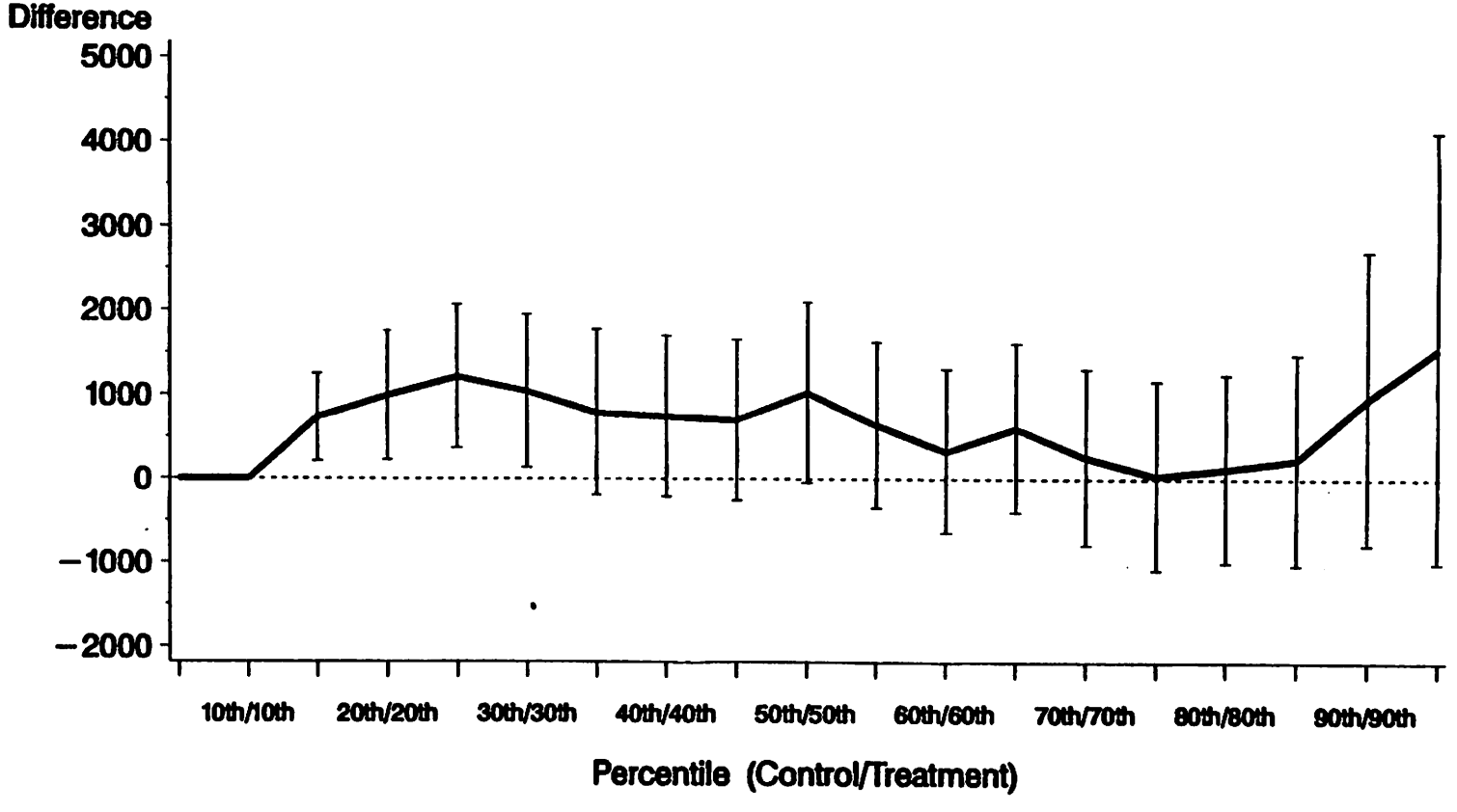
1. National JTPA Study data used in estimation.

2. Actual JTPA Performance Measures are defined as follows:
Hourly Wage at Placement: the average wage at program termination for employed adults.
Weekly Employment at Follow-up: the average weekly wage of adults who were employed 13 weeks after program termination.
Employment Rate at Placement: the fraction of adults employed at program termination.
Employment Rate at Follow-up: the fraction of adults who were employed 13 weeks after program termination.

3. In our analysis, employment rates were calculated based on the presence or absence of a job spell within 30 days of each reference date (random assignment, termination, or follow-up. Hourly wages were calculated based on the highest reported hourly wage for all job spells reported within 30 days of each reference date. Weekly earnings were calculated by averaging the product of hourly wages and hours worked per week across all reported job spells within 30 days of each reference date weighted by the fraction of the 30-day window spanned by each job spell.

4. Run on July 19, 1995.

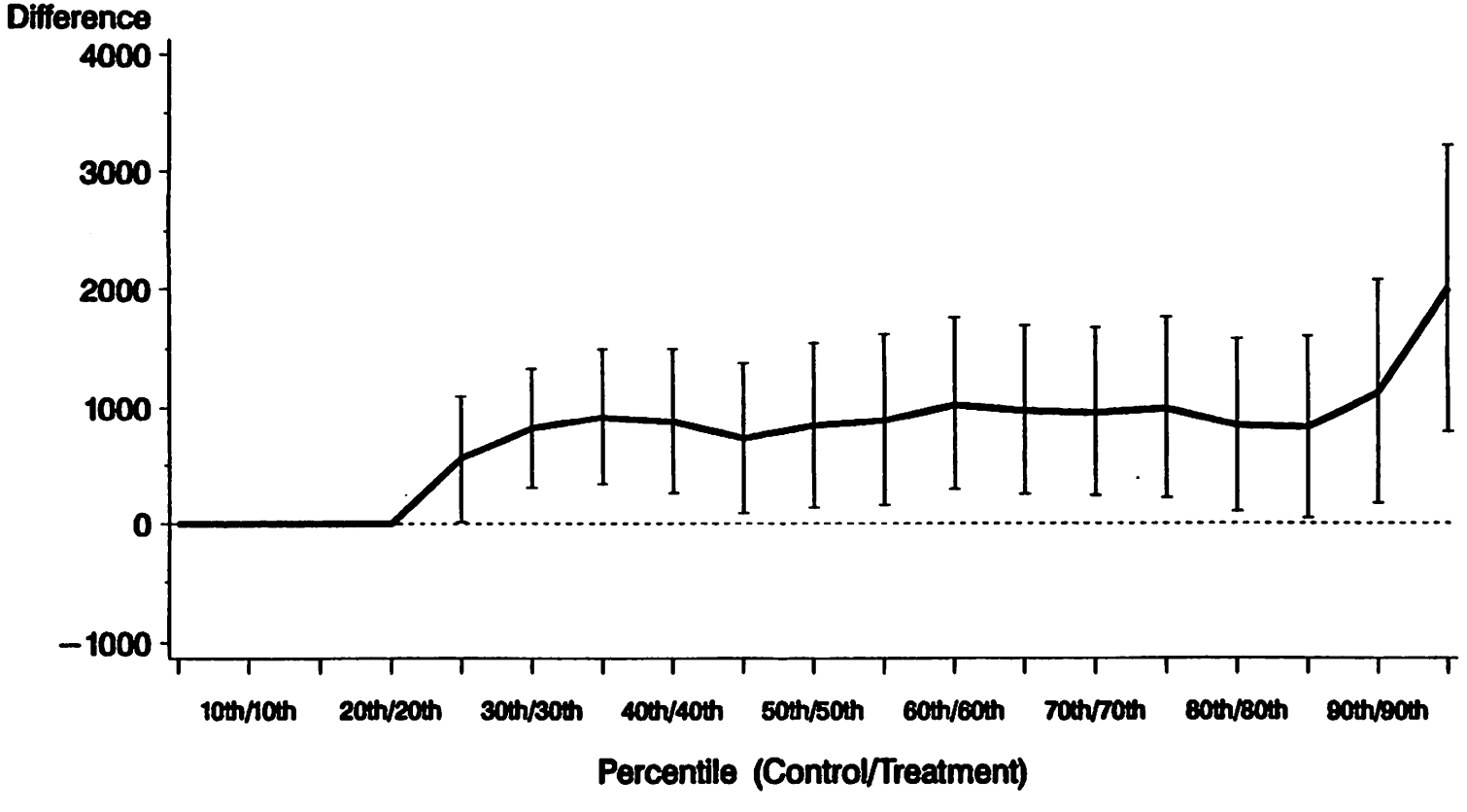
Figure 1A
Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution
Perfect Positive Dependence Case
Adult Males – Full Sample



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

Figure 1B
Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution

Perfect Positive Dependence Case
Adult Females – Full Sample

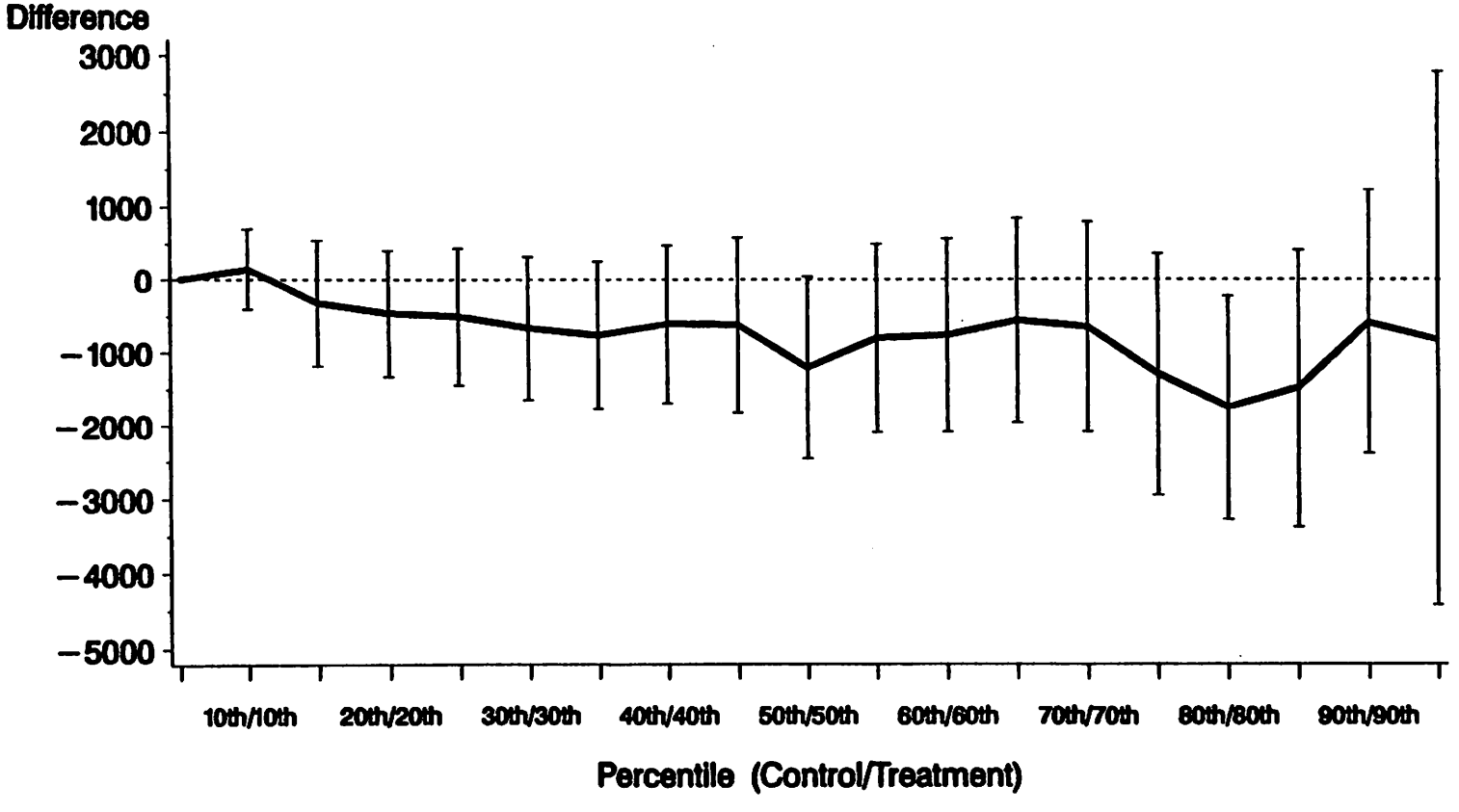


1. Sample consists of ABT's experimental 18-month study sample
2. ABT imputed values were used in place of outlying values
3. Standard errors for the quantiles are obtained using methods described in Csorgo (1993).

Figure 1C
Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution

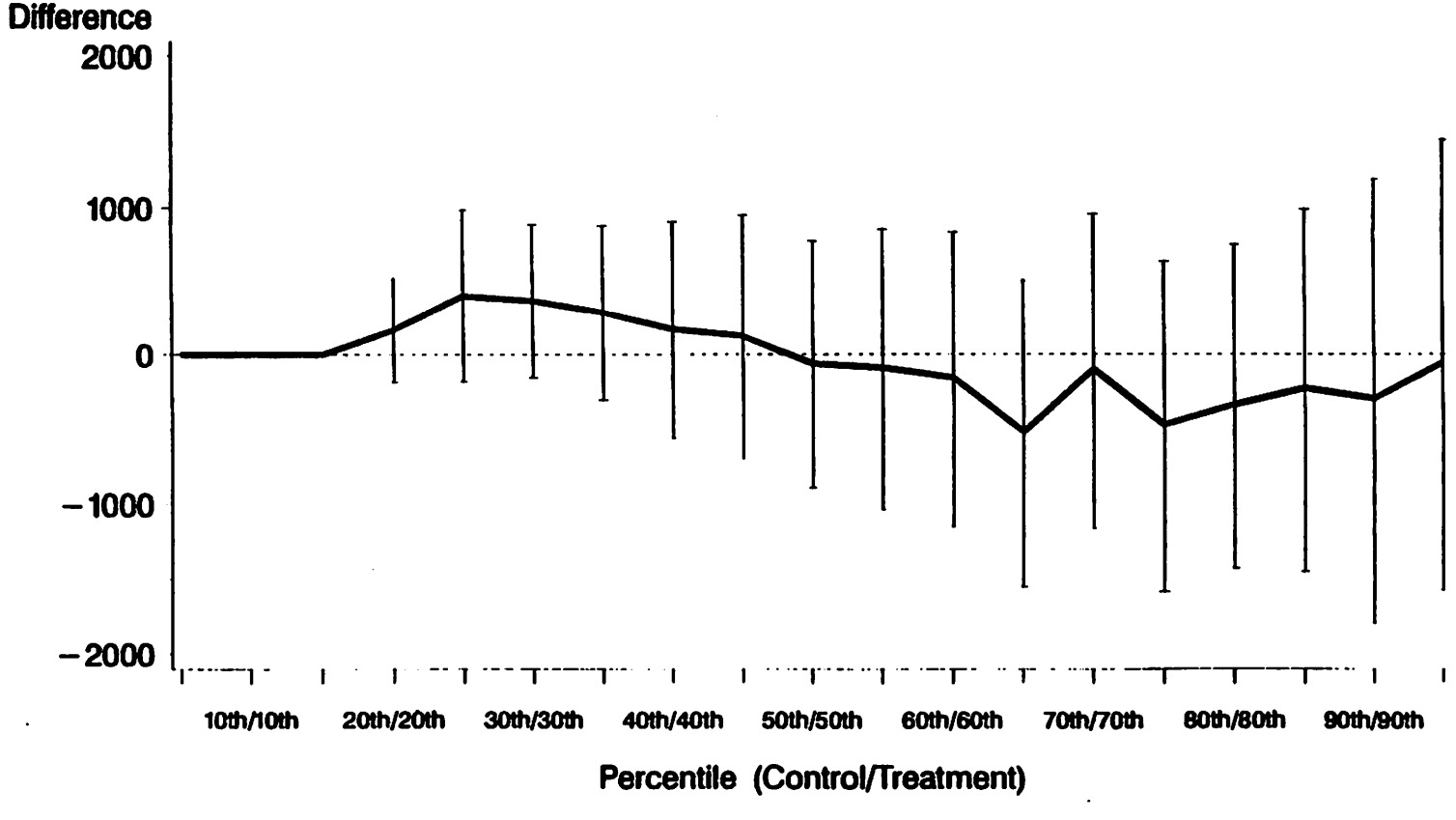
Perfect Positive Dependence Case

Male Youth – Full Sample



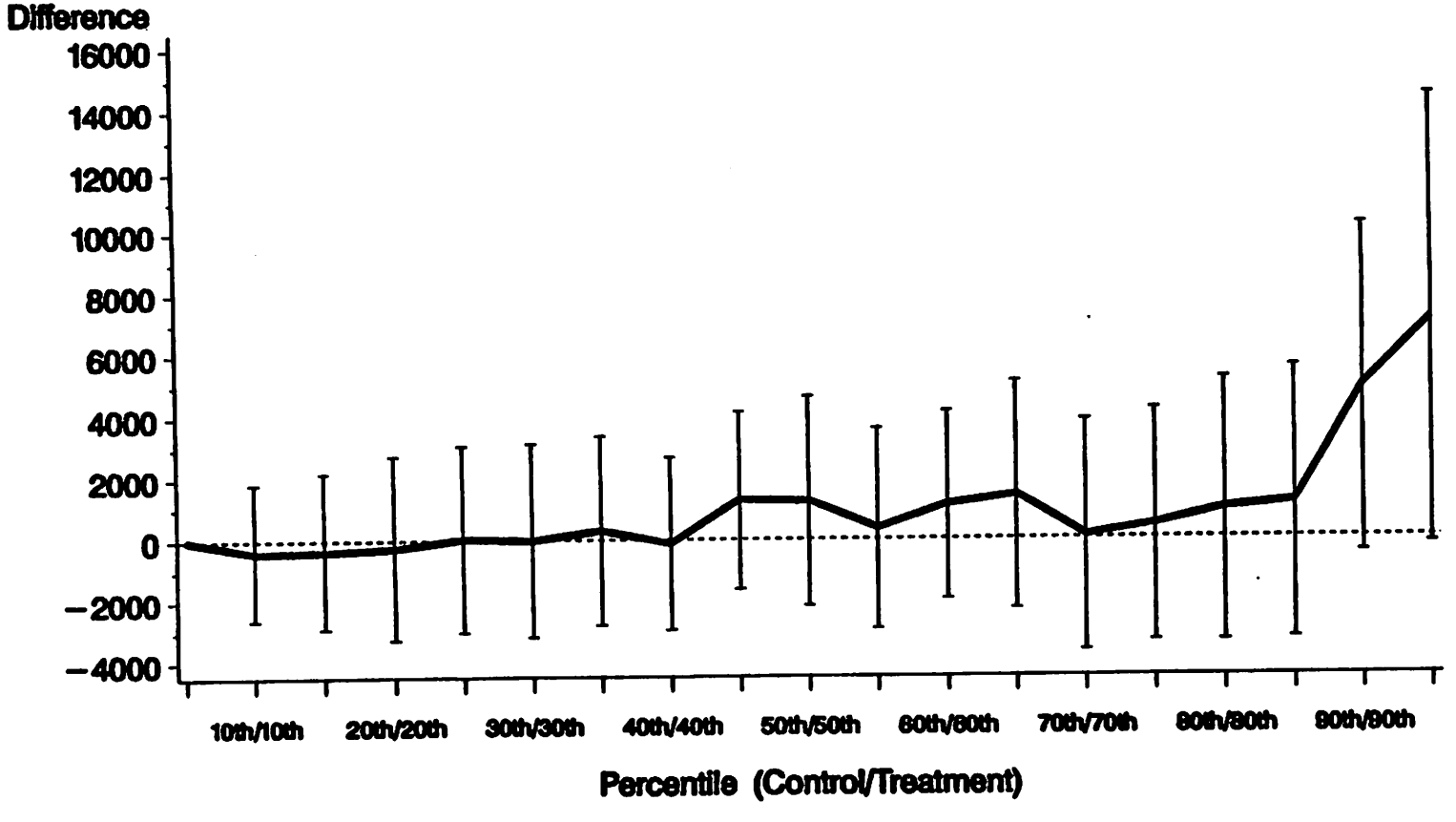
1. Sample consists of ABT's experimental 18–month study sample
2. ABT imputed values were used in place of outlying value
3. Standard errors for the quantiles are obtained using methods described in Csorgo (1993).

Figure 1D
Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution
Perfect Positive Dependence Case
Female Youth – Full Sample



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

Figure 2A
Treatment – Control Differences at Percentiles of the 30 Month Earnings Distribution
Perfect Positive Dependence Case
Adult Males – Full Sample

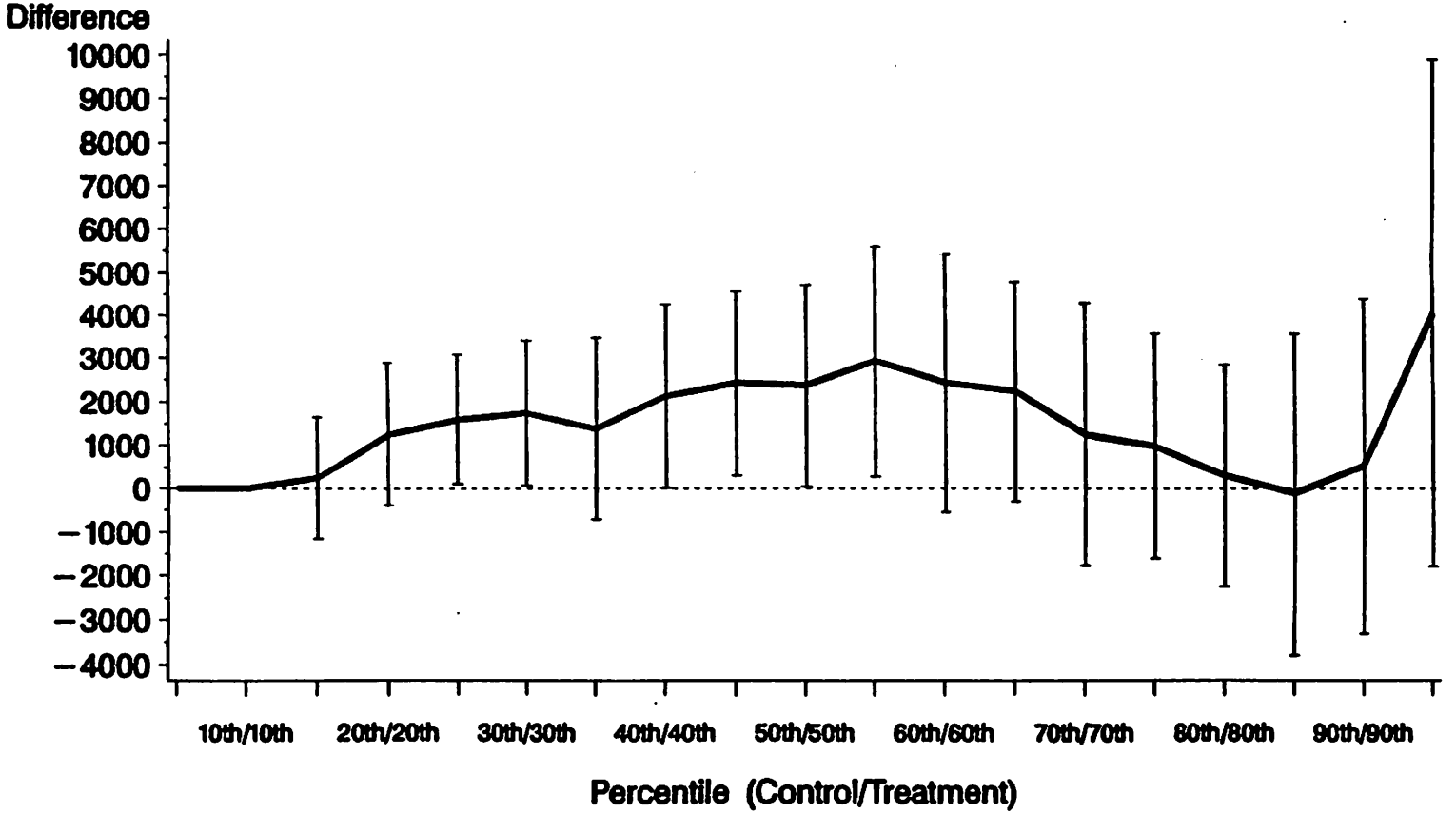


1. NJS experimental data. Sample consists of all persons with 30 months of valid self–reported earnings data.
2. Standard errors for the quantiles are obtained using methods described in Csorgo (1983).

Figure 2B
Treatment – Control Differences at Percentiles of the 30 Month Earnings Distribution

Perfect Positive Dependence Case

Adult Females – Full Sample

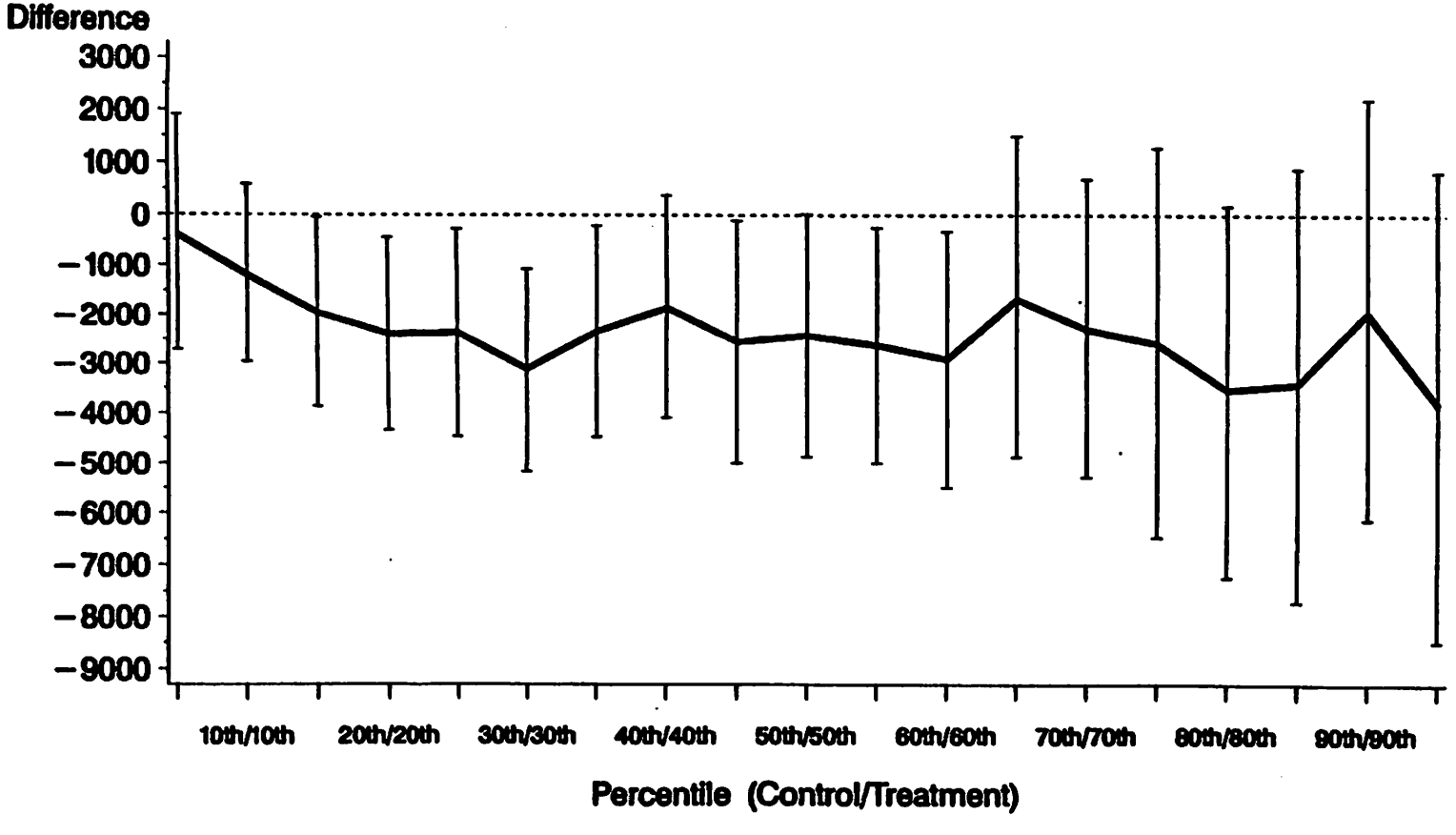


1. NJS experimental data. Sample consists of all persons with 30 months of valid self-reported earnings data.
2. Standard errors for the quantiles are obtained using methods described in Csorgo (1983).

Figure 2C
Treatment – Control Differences at Percentiles of the 30 Month Earnings Distribution

Perfect Positive Dependence Case

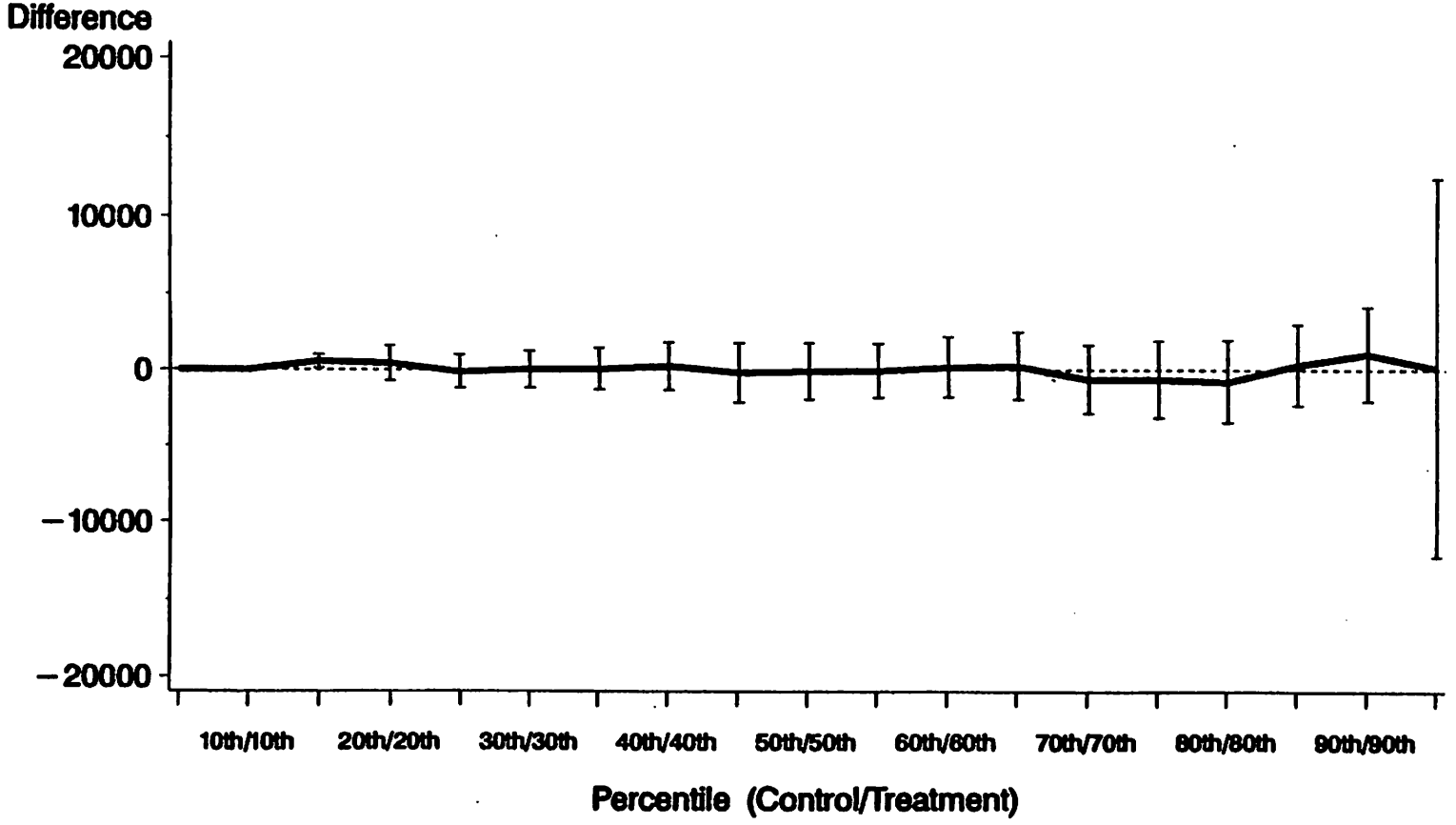
Male Youths – Full Sample



1. NJS experimental data. Sample consists of all persons with 30 months of valid self-reported earnings data.
2. Standard errors for the quantiles are obtained using methods described in Ceorgo (1993).

Figure 2D
Treatment – Control Differences at Percentiles of the 30 Month Earnings Distribution

Perfect Positive Dependence Case
Female Youths – Full Sample



1. NJS experimental data. Sample consists of all persons with 30 months of valid self–reported earnings data.
2. Standard errors for the quantiles are obtained using methods described in Csorgo (1993).