

Leaving Welfare and Joining the Labor Force:

Does Job Training Help?

Evidence from an Innovative Intervention in New York City

by

John Ifcher^{1,2}

JEL classification codes: I38, H52, H72

Keywords: welfare reform, job training, natural experiment

October 2006

¹ John Ifcher, Santa Clara University, Department of Economics, Adjunct Assistant Professor, 500 El Camino Real, Kenna 207, Santa Clara, CA, 95053, 408-554-5579 (phone), 408-554-2331 (fax), jifcher@scu.edu

² I wish to thank Alan Auerbach, David Card, Ken Chay, Swati Desai, Nada Eissa, Guido Imbens, John Quigley, Steve Raphael, and Emmanuel Saez. I also wish to thank the Burch Center at the University of California, Berkeley for its generous support and the New York City Human Resources Administration for making the data available and funding a portion of this work. All findings and conclusions expressed in this paper are those of the author.

Abstract

Starting in 1999, welfare recipients were required to participate in a job training and outplacement assistance program. Initially, recipients were enrolled in biweekly waves. I identify the effect of the program using an innovative natural experiment in which enrollees are compared to concomitantly eligible, non-enrolled recipients. Adjusting for control group contamination, I find that enrollees were over thirteen percentage points more likely to start a job. The majority of the newly employed remained off welfare for at least two years. Observed differences were not due to macroeconomic shocks or measurement error. The program passes a cost benefit test.

I. Introduction

In the early 1990s, a consensus was developing that welfare programs were failing. Critics charged that the programs were expensive, not preparing recipients for gainful employment, fostering welfare dependency, and ironically, not helping the individuals for whom the programs were designed.

In 1995, in response to these concerns, New York City (NYC) initiated a series of welfare reforms³. NYC's welfare caseload precipitously started contracting, and continued contracting, after having grown for the preceding six years. In late 1999, NYC created the Employment Services and Placement (ESP) program, a job training and outplacement assistance program for welfare recipients. Recipients were generally required to participate in the ESP program 14 hours per week and in the workfare program 21 hours per week⁴. Private contractors were hired to provide these services and paid strictly on a performance basis. The number of recipients who started a job annually doubled after the program's implementation.

There are many potential explanations for NYC's success in reforming its welfare programs. These explanations can be neatly divided into two categories, those related to the underlying economic conditions, and those related to changes in institutional factors. Prior research indicates that economic conditions alone generally cannot explain post-reform caseload reductions (Blank, 2002). This appears to be true in NYC as well.

³ NYC's welfare reforms were initiated prior to the passage of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996.

⁴ Prior to the implementation of the ESP program, recipients were required to participate in the workfare program 21 hours per week.

Comparing the welfare caseload and the unemployment rate in NYC, one observes an inconsistent relationship between the two prior to, and after, the reforms (see Figure 1). For example, they move in opposite directions from 1992 to 1997, and in the same direction from 1998 to 2001. Furthermore, assuming that there is a lag in the effect of the unemployment rate on the caseload does not explain the observed, inconsistent relationship. Thus, institutional factors appear to have had an impact.

NYC's welfare reforms were multi-faceted. In addition to the ESP program, other components included enhanced substance abuse treatment, improved fraud detection, and mandatory workfare participation. The success of the reforms could have been the result of any single component, or some combination thereof.

In this paper, I demonstrate the ESP program contributed to the success of NYC's welfare reforms. To do so, I take advantage of a quirk in the program's administration. Specifically, when the program was initiated, the entire cohort of eligible recipients could not be enrolled concurrently due to capacity constraints. Rather, recipients were enrolled in biweekly "waves." Computer programmers selected recipients for each wave centrally with the intention of making the process random. Intake interviews and objective assessments were not performed.

The wave enrollment process creates the opportunity to identify the effect of the ESP program using an innovative, natural experiment in which recipients who were enrolled on a given date, the "treatment group," are compared to those who were eligible, but not enrolled, on that date, the "control group." After adjusting for control group contamination, the observed post intervention employment difference peaks at over thirteen percentage points. Treated recipients are more than twice as likely to have

started a job than are control group members (23.7% versus 10.2%) eight months post treatment. Additionally, I demonstrate that the observed differences are not the result of macroeconomic shocks or measurement error.

The next section of this paper provides a brief overview of the previous literature. The third section presents a short description of NYC's welfare reforms between 1995 and 2000. The fourth section describes the identification strategy. The fifth section discusses the empirical implementation. The sixth section presents the results. The seventh section describes a robustness check that was performed. The final section discusses the implications of the results and presents a brief cost benefit analysis.

II. Previous literature

Many researchers have investigated the impact of welfare reform⁵ and there appears to be a consensus developing that it has had the intended effect, reducing welfare caseloads and increasing the labor force participation of former recipients (for example, Blank, 2002, and Moffitt, 2003). Unfortunately, that is where the consensus seems to end. Attempts to identify the effect of individual components of welfare reform, for example, time limits or job training, or the impact of welfare reform on recipients' well being, have not been as persuasive.

One branch of the literature has attempted to use aggregate data, or large nationally representative datasets, to identify the effects of welfare reform. For example, Blank (2001) uses the month in which recipients within a state were exposed to a bundle

⁵ Here welfare reform refers to both the waiver reforms authorized under the Family Support Act of 1988 and the Personal Responsibility and Work Opportunity Reconciliation Act of 1996.

of reforms to identify the effect of welfare reform. Such a strategy necessarily limits one's ability to identify the effect of an individual component of welfare reform. Attempts to solve this problem have been hampered by the large variety of welfare reform programs and the timing of the reforms (Blank, 2002).

This research is methodologically more similar to another branch of the literature that generally uses administrative data, in combination with random assignment experiments, to attempt to identify the impact of welfare reform. Generally, such research has found that mandatory welfare-to-work employment programs have decreased welfare use and increased annual earnings between \$200 and \$600 (Blank, 2002).

This study differs from the previous studies in several important respects. First, these studies generally investigate the impact of waiver reforms authorized under the Family Support Act of 1988. Recipients exposed to such reforms typically had not previously participated in mandatory welfare-to-work programs. This is a concern since Ifcher (2006) found that simply enrolling recipients (who had been previously unexposed to welfare reform) in a mandatory workfare program significantly increased the probability of exiting welfare; this effect was caused by enrollment in the program and was observable regardless of whether, or not, the enrollee participated in the program. Consequently, previous papers' estimates of the effect of welfare reform programs may be overstated and partly due to simply enrolling recipients (who had been previously unexposed to welfare reform) in the program. In contrast, the recipients in this study had been exposed to stringent "work first" welfare reform for up to five years. So the

estimated post intervention differences are likely due to the ESP program itself and not simply enrolling the recipients in a welfare reform program.

Second, the previous studies generally included randomized experiments. Although the results of randomized experiments are highly convincing, these studies are still hindered by the above concern as well as by the fact that multiple components of welfare reform were often implemented at approximately the same time. Consequently, the observed effects may be the result of a combination of components. Furthermore, randomized experiments can suffer from control group contamination, in which recipients in the treatment group talk to recipients in the control group during the study period (Blank, 2002). In contrast, the ESP program was implemented in isolation of other welfare reform components, and the control group contamination (described above) was limited due to the fact that most eligible recipients were enrolled within six months. This study also identifies the effect of a genuine, non-experimental welfare reform using a natural experiment.

Third, previous studies investigated the impact of welfare reform on Family Assistance (FA) recipients. This study investigates the impact of welfare reform on General Assistance (GA) recipients⁶, and includes approximately as many men as women. Previous studies found that while mandatory employment programs had a positive effect on the mean annual earnings of female recipients of FA, they had no effect on male recipients of FA (Friedlander et al, 1997). In contrast, this study found that the ESP program had a similar effect on men and women. Finally, there was one noteworthy

⁶ GA is welfare for childless adults. Concerns regarding generalizing the findings to FA recipients are explored in the discussion section of the paper.

and atypical feature of the ESP program, that is, that the contractors who provided the job training services for NYC were paid strictly on a performance basis.

This paper contributes to the growing literature on welfare reform in several ways. First, it identifies the effect of a single isolated component of welfare reform on recipients who had been previously exposed to welfare reform. Second, it develops an approach for identifying the effect of an intervention in which individuals are treated in waves. Third, it evaluates the effect of welfare reform in the NYC, where over five percent of all US welfare recipients live. Finally, it clearly demonstrates that welfare reforms can increase labor force participation and exits from welfare.

III. An overview of NYC's welfare reforms, 1995-2000

In 1994, Rudolph Giuliani, NYC's newly elected mayor, made reducing the city's welfare caseload a priority. At the time, NYC had over one million welfare recipients, including approximately 300,000 Home Relief (HR)⁷ recipients, and the city was spending approximately three billion dollars annually on welfare programs, including over one billion dollars on HR.

In early 1995, NYC initiated the Work, Accountability, and You (NYCWAY) program for HR recipients⁸. A central tenet of NYCWAY was that able-bodied welfare recipients were expected to work in exchange for their benefits. Almost all recipients fulfilled this requirement by participating in a workfare assignment 21 hours per week⁹.

⁷ HR was the name of NYC's GA program.

⁸ Starting in 1996, FA recipients were required to participate in NYCWAY as well.

⁹ Recipients typically worked for a city agency performing light manual labor or clerical tasks.

NYC's caseload declined from 1,150,509 to 533,284¹⁰, or by over 50%, over the next six years, after having grown by over 40% in the six prior years. Other components of NYCWAY included enhanced detection of welfare fraud, increased substance abuse treatment, and diversion to federally funded disability insurance where appropriate.

In late 1999, to enhance NYCWAY, and to mollify critics who complained that the program was too harsh and not helping recipients join the labor force, NYC created the ESP program, a job training and outplacement assistance program. In the subsequent two years, over 100,000 welfare recipients started a job annually, more than twice the pre-ESP program rate.

A. Description of the ESP program

Welfare recipients were now required to participate in a workfare assignment three days a week and in the ESP program two days a week, increasing the number of hours per week, from 21 to 35, that recipients were required to spend in structured activities. Eleven private contractors were hired to provide these services. Eight of the ESP contractors were not-for-profit organizations, and three were for-profit organizations. All of the contractors had a history of providing social services in NYC.

The ESP contractors were paid strictly on a performance basis, that is, they received a substantial payment¹¹ for each recipient they placed in a job. The payment was weakly increasing in the recipient's salary and tenure in the job. The contractors focused on developing participants' "soft skills," for example, résumé writing and

¹⁰ In the same period, the HR caseload declined by over 70%, from 297,102 to 87,293.

¹¹ The average payment was approximately \$3,000.

interview skills, and helped participants arrange job interviews¹². ESP contractors were required to attempt to place each recipient enrolled in the program. If, however, after six weeks a participant had not been placed, the contractor was permitted to stop attempting to place the recipient.

B. The population – job ready HR recipients with a workfare assignment

Prior to implementing the ESP program, there were approximately 15,000 HR recipients participating in workfare. Post implementation, each of these “pre-existing” participants was required to be in the ESP program if they were “job ready.” Private contractors were hired to conduct the job readiness evaluations, and 12,695 recipients were evaluated. 11,526, or 91%, were deemed to be job ready and form the population for this study.

C. The selection process – choosing enrollees for the biweekly waves

Since not all pre-existing recipients could be enrolled simultaneously, recipients were selected for the ESP program in biweekly waves. On November 3, 1999, the first wave of enrollees was selected for the ESP program. Enrollees were instructed by mail to report to their ESP contractor at a prescribed date and time¹³. The enrollment process continued until each member of the population was selected or had become ineligible¹⁴. Ultimately, 7,537 recipients were enrolled, or 65% of the population.

¹² The contractors were free to provide as much, or as little, training as they saw fit.

¹³ Typically, the reporting date was two weeks after the enrollment date.

¹⁴ To be eligible, a recipient had to remain on welfare and in a workfare assignment.

Computer programmers conducted the biweekly selection process centrally with the intention of generating a random sample of recipients, stratified by borough¹⁵, for each wave¹⁶. Enrollees were selected without an intake interview or objective assessment, in which human discretion could have played a role.

IV. Identification strategy

Generally, to identify the effect of a program, one compares the outcomes of treated and untreated individuals. In the case of the ESP program, each member of the population who remained eligible was ultimately treated. The only members of the population who were never treated must have become ineligible prior to being selected for the ESP program, and thus, do not make a valid comparison group, since becoming ineligible is endogenous¹⁷.

A. An alternative approach to identify the effect of the ESP program

It is possible, however, to estimate the effect of the ESP program taking advantage of the wave enrollment process. Specifically, the effect of the program is estimated using a natural experiment in which all enrollees from a wave are compared to all members of the population who were eligible to be enrolled in the same wave but

¹⁵ This reduced the difficulty of participating in the ESP program by limiting commute distances.

¹⁶ In section V, I demonstrate that the samples were probably not random.

¹⁷ For example, recipients who exited welfare, or were AWOL from workfare, became ineligible.

were not. By conducting this comparison for each of the first 17 waves¹⁸, and combining all enrollees into one group, the treatment group, and all eligible non-enrollees into another group, the control group, one can estimate the post intervention difference (PID) (see Figure 2). Specifically, the PID is defined as,

$$E[Y_i^T(t)] - E[Y_i^T(c)] \quad (1)$$

where $Y_i^T(g)$ is an indicator function which equals one if individual i started a job within T days of the enrollment date and zero otherwise, and is a function of which group g individual i was a member, where t is the treatment group and c is the control group.

The PID does not suffer from control group attrition, since each recipient who is placed in the control group remains in the control group for the entire study. For example, control group members who were subsequently treated or exited welfare remain in the control group. On the other hand, the PID does suffer from control group contamination, since 75% of control group members were treated subsequent to being included in the control group. Consequently, the PID will be negatively biased and a conservative estimate of the ESP program's true effect¹⁹.

B. Adjusting for control group contamination

To adjust for control group contamination, a set of additional, restricted control groups is created in which recipients are removed from the control group if they were treated within a given number of days of their inclusion in the control group. For example, a control group with a two-week (one-wave) restriction on being treated is

¹⁸ After which, 6,782 recipients were enrolled, 953 remained eligible, and 3,791 were ineligible.

¹⁹ Given that the ESP program increased the probability that recipients started a job.

created by removing all recipients from the original control group who were treated in the wave subsequent to the one in which they were placed in the control group. So, a member of the population who was eligible but not enrolled in the 1st wave, and enrolled in the 2nd wave, would be excluded from this restricted control group.

As described, the control group with a one-wave restriction on being treated introduces a selection problem. Specifically, it contains two distinct cohorts:

Cohort 1. Members who became ineligible prior to the subsequent enrollment date

Cohort 2. Members who were eligible but not enrolled in the subsequent wave

All recipients who are in the original unrestricted control group but are not in the new restricted control group must have been eligible on the subsequent enrollment date to be selected, and thus, would have been in cohort two if they had not been selected for the subsequent wave.

For the new restricted control group to be comparable to the treatment group the proportion of cohort two members in the new restricted control group must be the same as in the original unrestricted control group²⁰. Accordingly, the weight placed on each member of cohort two (in the new restricted control group) is increased such that the proportion of cohort two members in the new restricted control group is the same as in the original unrestricted control group. Specifically, the weight placed on each member of cohort two is the reciprocal of the probability that members of the unrestricted control

²⁰ This is true for the following two reasons. First, a recipient's membership in one of the two above cohorts is endogenous. Second, the design of this natural experiment ensures that the original control group is comparable to the treatment group.

group were not selected, conditional on being eligible to be, for the subsequent wave (see Figure 3).

One can extend this approach to create additional, more restrictive control groups. Control groups with a one-, three-, and six-month restriction on being treated, denoted respectively as control group one, control group three, and control group six, are used in this paper. This approach, unfortunately, cannot be extended beyond six months, since there were few control group members who remained eligible and were not treated within six months of their inclusion in the control group. Thus one can partially, but not fully, adjust for control group contamination, and the estimate of the PID using the restricted control groups will still be negatively biased.

V. Empirical implementation

With NYC's permission, the case history and demographic characteristics of each member of the population was extracted from an administrative database. From this data the treatment and control groups were formed as described in the previous section.

A. Comparing the treatment and control groups

Comparing the demographic characteristics of members of the treatment and control group, the most striking difference is that treated recipients are, on average, one year older, and have been continuously on welfare for five additional months (see Table 1). The distribution of the borough of residence is disparate as well²¹. To determine whether each wave approximated a stratified random sample as intended, one can test whether recipients' demographic characteristics significantly impacted the probability of being treated. Specifically, the probit equation below was estimated,

²¹ Stratifying the sample by borough and wave may have caused the observed differences.

$$P[D_i^{wb} = 1] = F(\alpha^{wb} + \sum_{c=1}^C \lambda_c^{wb} x_{ic}^{wb} + \varepsilon_i^{wb}) \quad (2)$$

where D_i^{wb} is a treatment dummy that equals one if individual i was treated in wave w in borough b and zero if individual i was not treated, but was eligible to be, in wave w in borough b ; and x_{ic}^{wb} is a series of C demographic characteristics for individual i in wave w in borough b . The number of enrollees per borough varied widely across waves²². To decrease the chance of generating spurious results, equation (2) is only estimated for the three largest waves in each of the four largest boroughs.

The coefficients on years continuously on welfare provide strong evidence that the selection process was probably not random²³. Of the twelve coefficients, eight are positive and significantly different from zero (see Table 2). A similar, but weaker, pattern was observed for the coefficients on age, seven are positive and significantly different from zero. It seems that the probability of being treated increased with age and

²² For example, while the 2nd wave had 40 enrollees from Brooklyn, the 3rd wave had 240.

²³ The computer programmers who conducted the selection process mistakenly believed that sorting the list of eligible recipients by borough would result in the recipients being randomly ordered in the sorted list. Thus they thought that by simply selecting recipients off the top of the sorted list for each borough they would be randomly selecting recipients for the ESP program.

the length of time continuously on welfare²⁴. For the other demographic characteristics there was no discernable selection pattern.

One thing is certain though; eligible recipients were selected solely using information that was stored in the database. The selection process was centralized and conducted by computer programmers. Individual caseworkers were not involved in any manner, and no intake interviews, or objective assessments, were conducted prior to selection. In other words, the selection process was conducted without human discretion. Such a selection process, even if it did not approximate a random one, should not disturb the necessary assumption that there was no systematic selection on unobserved characteristics. Consequently, by including covariates in the analysis, one should be able to adjust for the observed differences.

B. Adjusting for observed demographic differences

The PID, as defined, is probably negatively biased, since older recipients on welfare longer should have been harder to place in a job. To test whether this is true, a treatment dummy and a series of demographic characteristics are regressed on an outcome dummy,

$$y_i^T = \alpha^T + \beta^T D_i + \sum_{c=1}^C \lambda_c^T x_{ic} + \varepsilon_i^T \quad (3)$$

where y_i^T is an outcome dummy, which equals one if individual i started a job within T days of his or her inclusion in the treatment or control group and zero otherwise; D_i is a

²⁴ One credible explanation for the observed selection pattern is that the order in which the recipients were initially accepted for welfare benefits and entered into the database, was positively correlated with their position on the sorted borough lists.

treatment dummy, which equals one if individual i was treated and zero otherwise; and x_{ic} is a series of C demographic characteristics for individual i at the time of his or her inclusion in the control or treatment group. Since a new wave started every 14 days, regression coefficients are estimated for values of T that are multiples of 14²⁵ and between 14 and 728²⁶. Regression coefficients are calculated using OLS, and corrected standard errors are calculated by clustering the observations by individual²⁷.

C. Controlling for wave and borough of residence

One might be concerned that the stratification of the sample by wave and borough plays an important role in the observed PID. To eliminate this concern, wave, borough, and interaction (between wave and borough) dummies are added to equation (3),

$$y_i^T = \alpha^T + \beta^T D_i + \sum_{c=1}^C \lambda_c^T x_{ic} + \sum_{j=1}^4 \delta_j^T B_{ij} + \sum_{k=1}^{16} \gamma_k^T W_{ik} + \sum_{j=1}^4 \sum_{k=1}^{16} \eta_{jk}^T (B_{ij} * W_{ik}) + \varepsilon_i^T \quad (4)$$

where y_i^T , D_i , and x_{ic} are defined as above, B_{ij} is a borough dummy, which equals one if individual i resides in borough j and zero otherwise, and W_{ik} is a wave dummy, which equals one if individual i was added to the treatment or control group in wave k and zero otherwise²⁸.

²⁵ This approach enables one to estimate a very general, non-parametric hazard rate.

²⁶ The maximum days post enrollment for which data is available for all recipients in the study.

²⁷ This is necessary since many recipients appear in the dataset repeatedly.

²⁸ Again regression coefficients are estimated using OLS with corrected standard errors for values of T that are multiples of 14 and between 14 and 728.

VI. Results

The PID peaks at seven and half percentage points 98 days post enrollment, at which time members of the treatment group are more than twice as likely to have started a job than are members of the control group (13.2% versus 5.7%) (see Figure 4). The PID then declines to four percentage points 728 days post enrollment (36.3% versus 32.3%). As was discussed previously, this estimate of the PID is negatively biased since there is extensive control group contamination.

A. Adjusting for control group contamination

Calculating the PID using the restricted control groups (one, three, and six), one finds that the PID increases as the length of the restriction increases (see Figure 5). The PID peaks at thirteen and a half percentage points 252 days post enrollment using control group six. At the peak, members of the treatment group are more than twice as likely to have started a job than are members of the control group (23.7% versus 10.2%). The difference then declines steadily to seven and a half percentage points 728 days post enrollment (36.3% versus 28.7%). This PID is still potentially negatively biased since control group contamination cannot be adjusted for fully.

Additionally, the PID using control group six is superimposed over the PID using the control group three for the first 126 days post enrollment. The former does not diverge from the latter until after the restriction on membership in the latter control group expires. The same pattern is observed when comparing the PID using control group three and one, and using control group one and the unrestricted control group. Thus, the PID using control group six is probably the upper envelope of the true employment effect for

182 days post enrollment. After that, the true employment effect almost certainly diverges from, and lies above, this PID.

B. Adjusting for observed demographic differences

To determine if the observed differences in demographic characteristics between the treatment and control group have an impact, equation (3) is estimated with, and without, the demographic characteristics included. The coefficients on the treatment dummy increase slightly when the demographic characteristics are included in the regression for all values of T (see Figure 6). This is not surprising since the probability of being treated was higher for older recipients who were on welfare longer, and such recipients were presumably less likely to start a job. This presumption is supported by the fact that the coefficients on age are negative and significantly different than zero for all value of T (see columns (2) and (5) of Table 3). Finally, the coefficients on the treatment dummy are positive and highly significant with t-scores between eight and 22 for values of T greater than 14. The above findings hold regardless of which control group is used.

C. Controlling for wave and borough of residence

To control for wave and borough of residence, equation (4) is estimated. The regression coefficients on the treatment dummy and age are stable regardless of which additional explanatory variables are included in the regression (see Figure 6 and columns (3) and (6) of Table 3)²⁹. Moreover, including these variables should control for any macroeconomic shocks that may have occurred during the study. Thus, the findings are

²⁹ The above findings hold regardless of which control group is used.

not simply the result of economic conditions. Furthermore, the central findings hold even if the study period is prematurely concluded prior to September 11th, 2001.

D. The long-term effect of the ESP program

Starting a job is not the sole outcome by which the ESP program should be evaluated. For it to be considered a true success, it must not only have increased the probability that recipients started a job, but also increased the probability that they remained employed and off welfare. Unfortunately NYC did not collect job retention data for all recipients who started a job. In truth, however, one is not concerned with whether recipients retained their first job, but rather whether they remained off welfare and became long-term labor force participants.

The long-term effect of the ESP program is thus estimated by determining the program's effect on the probability that recipients started a job and "permanently" exited welfare^{30, 31}. Specifically, equation (4) is estimated with a new dependent variable y_i^T , which equals one if individual i has started a job and permanently exited welfare within T days of his or her inclusion in the treatment or control group, and zero otherwise. The observed PID peaks at almost eight percentage points 490 days post enrollment using

³⁰ "Permanently" means that the recipients exited welfare, and did not return, within two years of their inclusion in the treatment or control group. Only two years of post enrollment data was provided for each recipient.

³¹ Data regarding recipients' labor force participation is not available. To date, New York State has been unwilling to provide the quarterly earned income data for the recipients in the study.

control group six (see Figure 7). The coefficients on age continue to be highly significant and negative.

The same pattern of superimposed differences is observed when comparing control groups with longer restrictions to those with shorter restrictions. Thus, it seems likely that for 182 days post enrollment the PID using control group six is the upper envelope of the true long-term effect. After that, the true effect almost certainly diverges from, and lies above, this PID.

VII. Robustness check

One might be concerned that “starting a job” is better observed for members of the treatment group than for members of the control group³². This is unlikely for two reasons. First, 75% of the control group was ultimately treated. Second, in 2000, NYC set the ambitious goal of placing 100,000 recipients in a job annually. To meet this goal, NYC had a strong incentive to accurately record each job placement.

To rule out the possibility that the observed differences are the result of measurement error (in identifying which recipients started a job), one can estimate the impact of the ESP program on the probability of permanently, and unconditionally, exiting welfare. Specifically, equation (4) is estimated with a new dependent variable y_i^T , which equals one if individual i has (1) permanently exited welfare within T days of his or her inclusion in the treatment or control group and (2) has been off welfare for at

³² This concern arises since ESP contractors were paid for each job placement. Thus the contractors had a strong incentive to make sure that each placement was recorded.

least six months at the end of the study³³, and zero otherwise. The observed PID peaks at over seven percentage points 364 days post enrollment (see Figure 8). The coefficients on the treatment dummy are positive and significantly different than zero for values of T between 112 and 546 using control group six.

Comparing the observed effect of the ESP program on the probability that a recipient started a job and a recipient permanently exited welfare, one observes that the former effect peaks approximately six percentage points higher than does the latter. This is explained by two factors. First, there were many paths off welfare other than starting a job, for example, moving out of NYC or failing to comply with program requirements. This factor appears to explain over half the variation, since that the probability of exiting welfare, conditional on not starting a job, is about one-third for both the treatment and control group, and more members of the treatment group started a job. Second, recipients who were treated and started a job were slightly more likely to return to welfare than were recipients who were not treated and started a job. Thus, the observed differences appear to be robust.

VIII. Discussion

The ESP program appears to have had a persistent positive impact on enrollees. It increased the probability that a recipient will start a job by at least thirteen percentage points, and it increased the probability that a recipient will start a job and permanently exit welfare by at least eight percentage points. Again, these estimates may be

³³ This restriction prevents exits at the end of the study period from being counted as permanent exits. Thus, coefficients can only be estimated for values of T between 14 and 546.

conservative since control group contamination cannot be controlled for fully. Finally, in a recent paper that investigates whether temporary help jobs foster long-term employment, Autor and Houseman (2005) demonstrate that enrolling welfare applicants in Work First³⁴ increased their earnings, and quarters of employment, when applicants found a direct hire job³⁵. Since the ESP program placed participants in direct hire jobs, it is likely, based on Autor and Houseman's findings, that recipients who started a job increased their earning and quarters of employment.

A. Understanding the Results

A strength of this study is that all enrollees were considered treated regardless of whether, or not, they actually participated in the ESP program. Thus, the PID estimates the “Intent to Treat” effect (not the “Treatment on the Treated” effect), and does not suffer from self-selection bias³⁶. A limitation of this study is that only job-ready HR recipients, who were participating in workfare, were included in this study. This limitation is the direct result of the program's design and the identification strategy, and may affect the conclusions one can draw in the following three ways. First, the PID may be biased for welfare recipients who are not job ready. This is probably not a major concern since finding employment for such recipients is, presumably, not an appropriate objective. Furthermore, 90% of workfare participants were found to be job ready. Second, the observed effects may be biased for welfare recipients who have not

³⁴ Work First is a welfare-to-work program in Michigan.

³⁵ Applicants who found a temporary help job did no better than those who did not find a job.

³⁶ See Katz et al (2001) for a comparison of these effects.

participated in workfare³⁷. From a policy perspective, this also does not appear to be a major concern since concurrent workfare participation could be incorporated into any similar program. Finally, the observed effects may be biased for FA recipients. This, however, seems not the case. In a similar study, using a smaller sample of FA recipients, Ifcher (2004) found that the ESP program had virtually the same impact on FA recipients as on GA recipients.

Perhaps the most intriguing finding is that the ESP program had an impact at all. Recall that the ESP contractors were paid on a performance basis; were free to provide as much, or as little, training as they saw fit; and primarily focused on developing the participants' soft skills³⁸. Consequently, one could credibly claim that participating in the ESP program did not substantially increase recipients' human capital. Furthermore, the majority of recipients in this study had unfavorable demographic characteristics, barriers to employment, and been exposed to stringent welfare reforms for many years. In 1995, when NYCWAY was initiated, the average age of an able-bodied HR recipient was 39. In 2000, when the ESP program was initiated, the average age of an able-bodied HR recipient had increased to 47. Strikingly, in just five years, the average age had increased by 8 years. Apparently, the younger HR recipients, who presumably had better alternative opportunities, had already left the rolls by 2000.

³⁷ The observed PID might result from an interaction between the ESP and workfare programs.

³⁸ They also typically employed "job developers" who identified and arranged job interviews.

One possible explanation for the effectiveness of the ESP program is that it substantially increased the cost, or disutility, of being on welfare. Specifically, the introduction of the ESP program increased the number of hours per week, from 21 to 35, that recipients were required to spend in structured activities. This change may have reduced the net utility of being on welfare enough that some recipients started a job and or exited welfare. This explanation implies that NYC could have engaged the recipients in any structured activity for the additional 14 hours per week and the effect would have been the same. One reason to doubt the accuracy of this explanation is that members of the treatment group were less likely to have (permanently and unconditionally) exited welfare than were members of the control group for the first 56 days post enrollment (see Figure 8), possibly indicating that the net utility from receiving welfare actually increased after enrollment in the ESP program.

Another, more appealing, explanation is that the ESP program enabled participants to get a “foot in the door.” That is, once participants were able to find, and start a job, they were able to remain employed. In such a scenario, recipients had, or quickly developed, the requisite skills, and had been on welfare simply because they could not find a job. There are many avenues through which the ESP program could have helped participants get their foot in the door. For example, participating in the ESP program could have helped by sending a positive signal to prospective employers or developing participants’ soft skills and “networks.”

One additional explanation is that the ESP program was a beneficial complement to the workfare program. In other words, the ESP program in conjunction with the workfare program not only helped recipients get their foot in the door, but also helped

them develop basic work skills, for example, being timely and working well with others, that were necessary for them to find a job and remain employed.

B. Cost Benefit Analysis

Finally, one should consider whether the fiscal benefits of the ESP program outweighed the costs. The observed cost of a placement was approximately \$3,000. However, many of the participants who were placed in a job by the ESP contractors would have found a job on their own. Consequently, the relevant cost is that of a “new” placement, that is, a placement that would not have occurred without the ESP program. Given that approximately 35% of the treatment group started a job and that the PID was approximately 13 percentage points, NYC had to pay for 35 placements to generate 13 new placements. Consequently, the cost of a new placement was approximately \$8,100, or $(\$3,000 \times 35) \div 13$.

The fiscal benefit per placement is harder to calculate since it is time dependent, that is, it depends on how long a recipient remained off welfare and/or employed³⁹. In early 2000, a typical HR recipient received \$350 per month in GA and \$120 per month in food stamps. This totals almost \$6,000 per year⁴⁰. Additionally, one must add administrative costs, income taxes paid by former recipients, and Medicaid costs (each HR recipient received Medicaid), to the benefit calculation. This presumably adds at least a \$1,000. Thus, an estimate of the fiscal benefit is \$7,000 per year.

Unfortunately, only two years of post enrollment data is available, and it is not possible to determine whether recipients who started a job, remained off welfare past the

³⁹ It also varies by recipient. Different recipients may have received different benefits.

⁴⁰ Employed former recipients with very low earnings qualified for some food stamps.

break even point, 1.2 years, or $\$8,100 \div 7000$. At the end of the study, over two-thirds of the recipients who were treated and started a job were still off welfare. These recipients had, on average, already remained off welfare for approximately a year. Projecting the benefits into future, assume that in each subsequent year one-third of the treated recipients, who started a job and were still off welfare, returned to welfare⁴¹. With such an assumption, the fiscal benefits of the ESP program easily surpass the costs. Even with an extremely conservative assumption that 60% of those still off welfare return to it in each subsequent year, the fiscal benefits of the program outweigh the costs. Thus it seems highly likely that the ESP program passes a fiscal cost benefit test⁴².

⁴¹ Previous studies have suffered from the same limitation and have projected the benefits of short-term training programs into the future as well (Friedlander et al, 1997).

⁴² The fiscal benefits of placing a FA recipient were larger, totaling about \$13,000 per year. Thus an equally effective program for FA recipients would pay for itself in under a year.

References

- Autor, David, Houseman, Susan, 2005. Do Temporary Help Jobs Improve Labor Market Outcomes for Low-Skill Workers? Evidence from Random Assignments. Manuscript (http://econ-www.mit.edu/faculty/download_pdf.php?id=1185).
- Blank, Rebecca M., 2002. Evaluating Welfare Reform in the United States. *Journal of Economic Literature* 40, 1105 – 1166.
- _____, 2001. What Causes Public Assistance Caseloads to Grow? *Journal of Human Resources* 36, 85 – 118.
- Friedlander, David H., Greenberg, David H., Robins, Philip K., 1997. Evaluating Government Training Programs for the Economically Disadvantaged. *Journal of Economic Literature* 35, 1809 – 1855.
- Ifcher, John, 2006. The Abrupt Impact of Welfare Reform: Evidence from Enrolling Recipients in New York City's Mandatory Workfare Program. Manuscript.
- _____, 2004. The Impact of Job Training on Welfare Recipients with Children. Dissertation chapter 3.
- Katz, Lawrence F., Kling, Jeffrey R., Liebman, Jeffrey B., 2001. Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *The Quarterly Journal of Economics* 116, 607 – 654.
- Moffitt, Robert A., 2003. The Temporary Assistance for Needy Families Program. In: Moffitt, Robert A. (Eds.), *Means-Tested Transfer Programs in the United States*. The University of Chicago Press, Chicago, IL, 291 - 363.

Table 1

Demographic Characteristics of Members of the Treatment and Control Groups

Demographic characteristic	Treatment group	Control group
Observations	6,782	58,051
Gender		
Male	53.9%	54.1%
Female	46.1%	45.9%
Race		
Asian	0.9%	1.2%
Black	47.8%	49.3%
Hispanic	37.1%	34.9%
White	8.6%	9.6%
Other	0.6%	0.7%
Not reported	5.0%	4.2%
Borough of residence		
Bronx	30.8%	39.9%
Brooklyn	26.5%	31.4%
Manhattan	28.9%	16.8%
Queens	12.4%	11.0%
Staten Island	1.4%	0.6%
Not reported	0.0%	0.2%
Average age	48.2 (8.315)	47.2 (8.727)
Average number of years continuously on welfare	3.7 (3.084)	3.3 (3.098)
Is a U.S. citizen	90.8% (.289) {687}	90.0% (.300) {5,080}

Standard deviations are given in parenthesis and number of missing observations are given in brackets

Table 2

Coefficients on Demographic Characteristics from Estimating Equation (2) for Select Borough-Wave Pairs

Demographic characteristic	Bronx 3rd wave	Bronx 4th wave	Bronx 12th wave	Brooklyn 3rd wave	Brooklyn 11th wave	Brooklyn 16th wave	Manhattan 1st wave	Manhattan 2nd wave	Manhattan 3rd wave	Queens 4th wave	Queens 5th wave	Queens 6th wave
Male	-0.043	0.059	0.009	0.188 **	-0.132	0.398 ***	0.140 **	-0.020	0.018	0.150	-0.105	-0.016
Race												
Asian	-0.023	-	-	-	-0.161	-1.541 ***	-0.579	-0.661	0.527	0.140	-0.305	0.266
Black	-0.072	-0.048	0.116	-0.086	0.281	-0.417	-0.299 **	-0.040	-0.220	-0.056	0.051	0.293
Hispanic	0.087	-0.049	0.234	-0.050	0.371	-0.353	-0.249 *	-0.020	0.113	0.099	0.117	0.513
White	0.206	-0.122	-0.013	-0.318 *	0.187	-0.687 **	-0.142	-0.226	-0.429	-0.175	-0.031	0.042
Average age	0.008 **	0.002	0.018 ***	0.012 ***	0.003	0.008	0.007	0.009 **	0.011 *	0.007	0.021 ***	0.018 **
Years on welfare continuously	0.080 ***	0.067 ***	-0.020	0.044 ***	-0.057 ***	0.045 ***	0.109 ***	0.079 ***	-0.010	0.093 ***	0.012	-0.042

* signifies that $p < 0.10$, ** signifies that $p < 0.05$, and *** signifies that $p < 0.01$

- variable was dropped because there were not enough observations

Table 3

Coefficients of Explanatory Variables at 252 and 728 days Post Enrollment Using Control Group Six

Demographic characteristic	252 days post treatment (peak effect)			728 days post treatment		
	(1)	(2)	(3)	(4)	(5)	(6)
Enrollment date dummies	NO	NO	YES	NO	NO	YES
Interaction dummies[^]	NO	NO	YES	NO	NO	YES
Treatment dummy	0.134 ***	0.139 ***	0.134 ***	0.075 ***	0.083 ***	0.082 ***
Male		-0.007	-0.004		-0.020	-0.016
Race						
Asian		0.109	0.093		0.080	0.061
Black		0.012	0.003		0.022	0.013
Hispanic		-0.007	-0.006		-0.052	-0.043
White		0.004	-0.020		-0.014	-0.046
Average age		-0.002 ***	-0.002 ***		-0.004 ***	-0.004 ***
Average number of years continuously on welfare		0.002	0.003		0.001	0.000
Borough						
Bronx			-0.038			-0.076
Brooklyn			0.051			0.041
Manhattan			-0.004			0.026
Queens			0.045			0.022

[^] between borough and enrollment date* signifies $p < 0.10$, ** signifies $p < 0.05$, and *** signifies $p < 0.01$

Figure 1

The Unemployment Rate and Welfare Caseload in NYC, January 1978 – June 2004

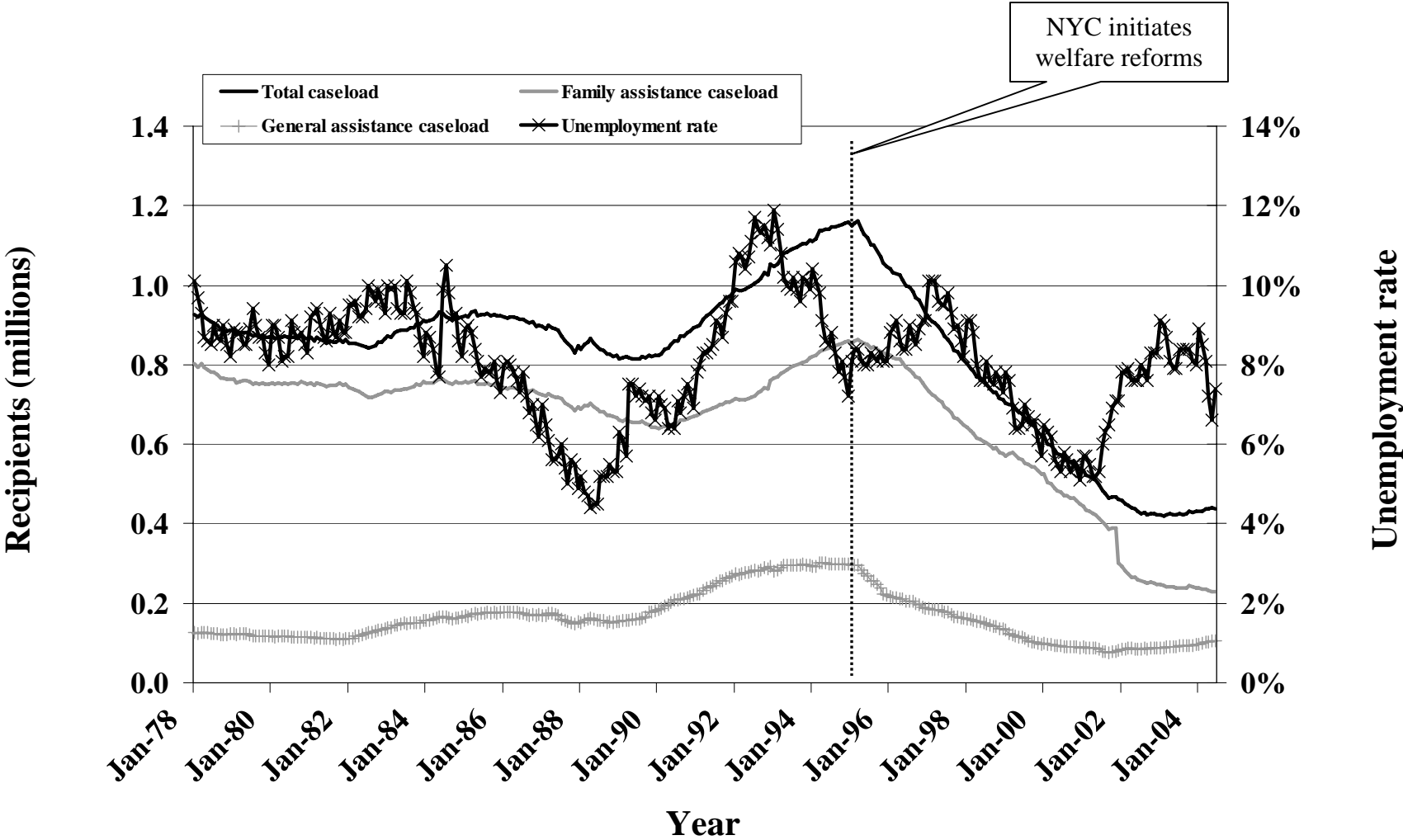


Figure 2

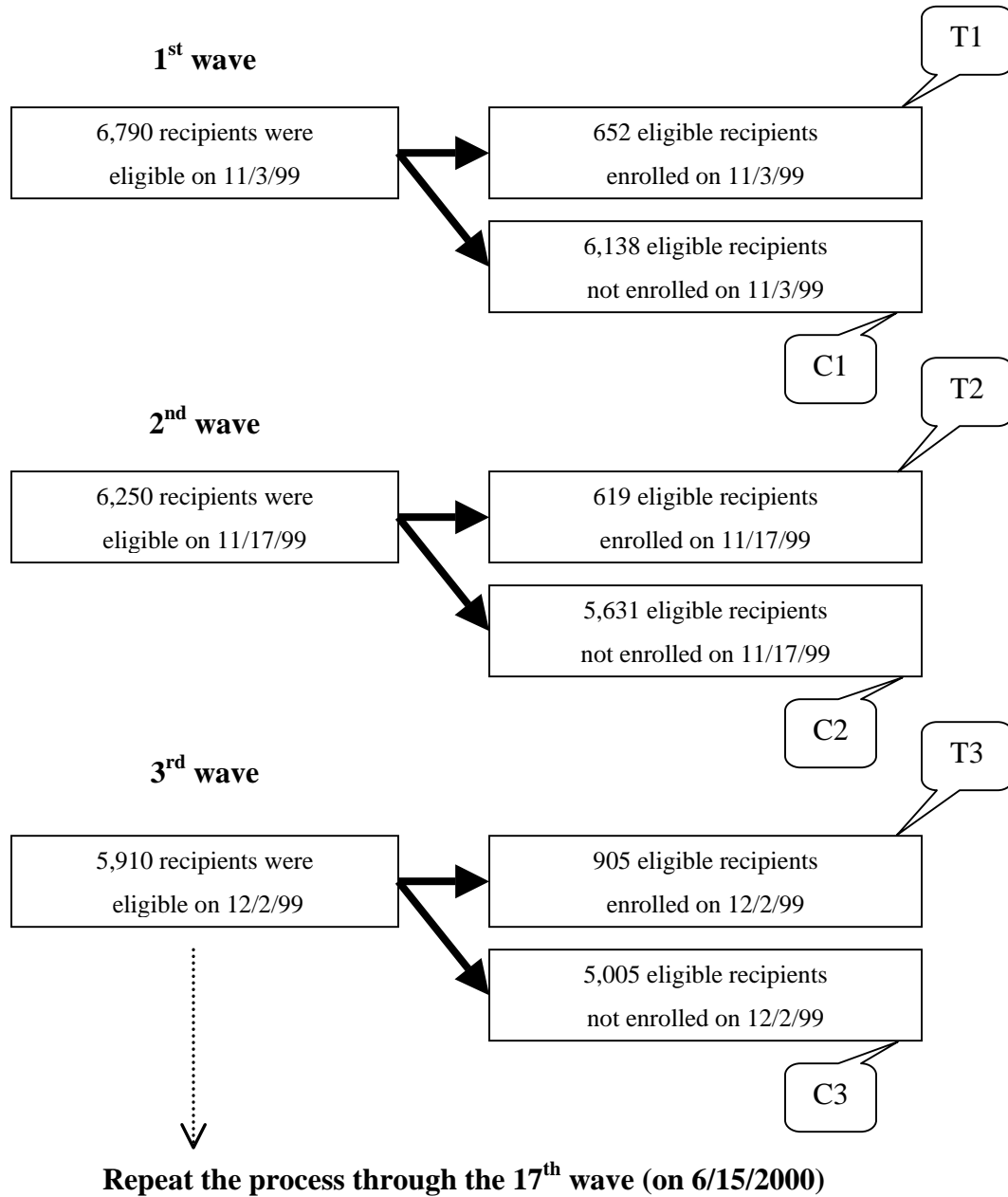
The Formation of the Treatment Group and the Control Group

The treatment group:

All recipients enrolled in the ESP program during the first 17 waves. Specifically it is the union of T1, T2, T3,, T17.

The control group:

All the members of the population, who were eligible to be enrolled, but were not, in each of the first 17 waves. Specifically it is the union of C1, C2, C3,, C17¹.



¹ Note that many members of the population were eligible to be enrolled, but were not, in multiple waves. Consequently, many recipients are members of the control group multiple times.

Figure 3

The Formation of the Treatment Group and the Control Group with a Two-Week Restriction on Being Treated

The treatment group:

All recipients enrolled in the ESP program during the first 17 waves. Specifically it is the union of T1, T2, T3,, T17.

The control Group with a two-week restriction on being treated

All recipients who were

1. eligible but not enrolled during any of the first seventeen waves and
2. not treated in the subsequent wave.

Specifically it is a weighted union of C1a, C1b, C2a, C2b, C3a, C3b,, C17a, C17b. The weight placed on each member of each a-series cohort is equal to one. The weight placed on each member of each b-series cohort is equal to the reciprocal of the probability of not being enrolled in the subsequent wave, conditional on be eligible to be enrolled in that wave, e.g., for members of C1b the weight is the reciprocal of .896 (5,248/5,860), or 1.116.

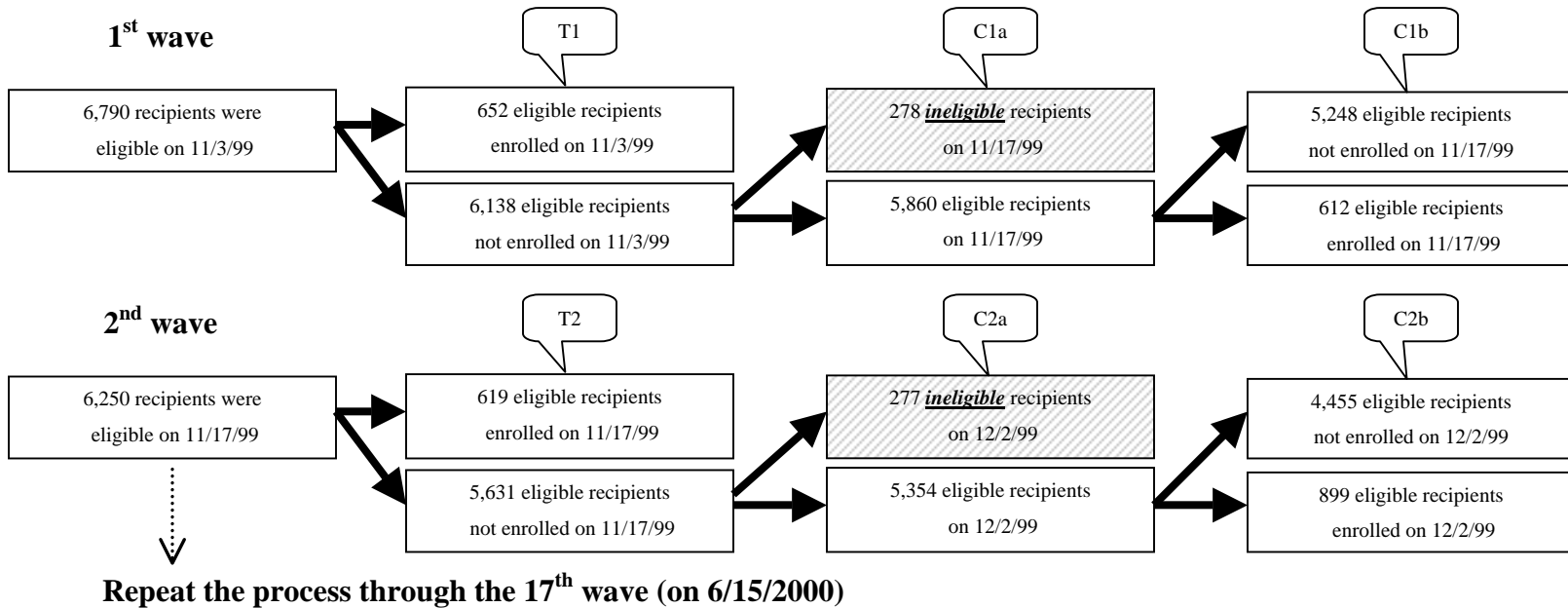


Figure 4

The PID

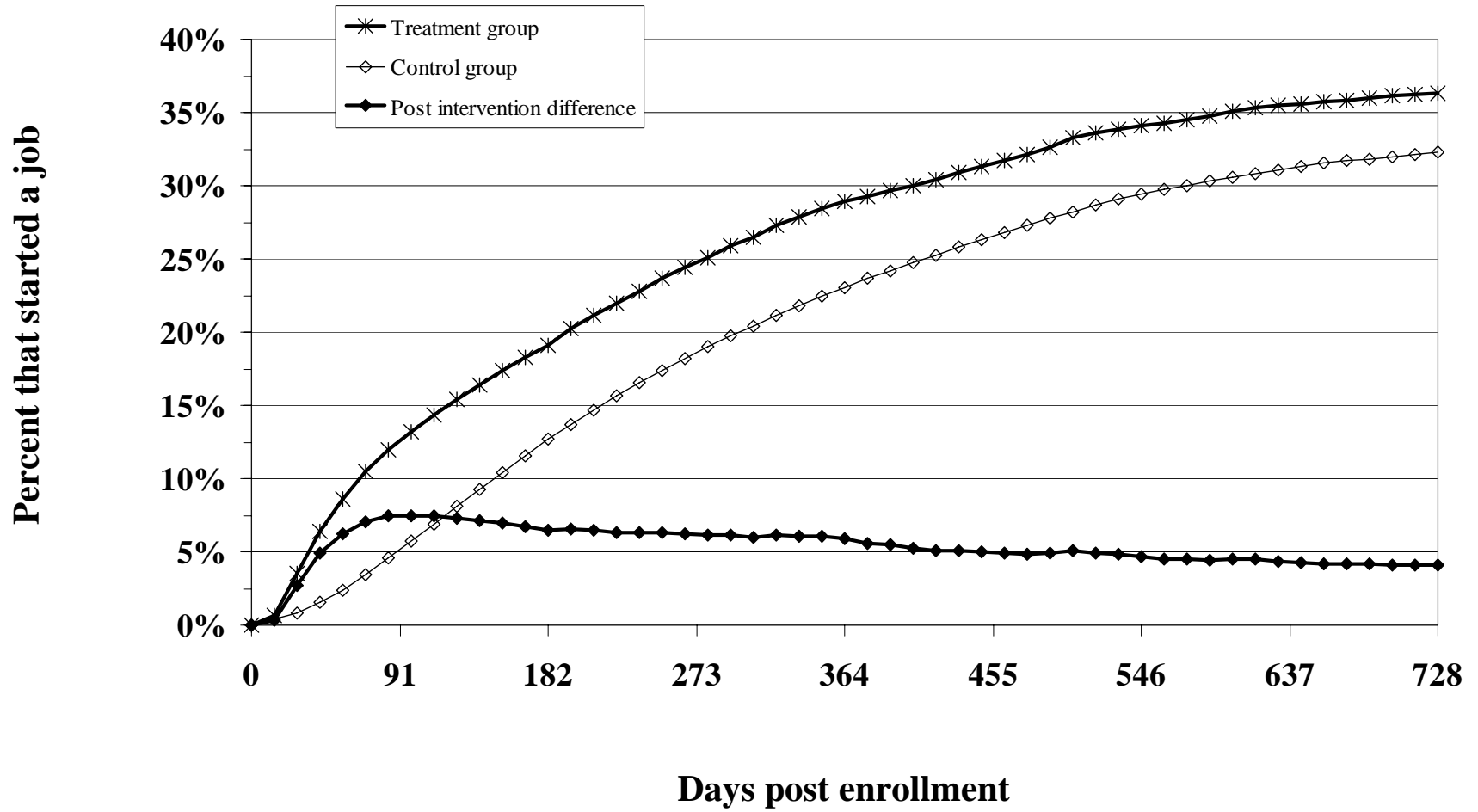


Figure 5

The PID Using the Original Control Group and Control Groups One, Three, and Six

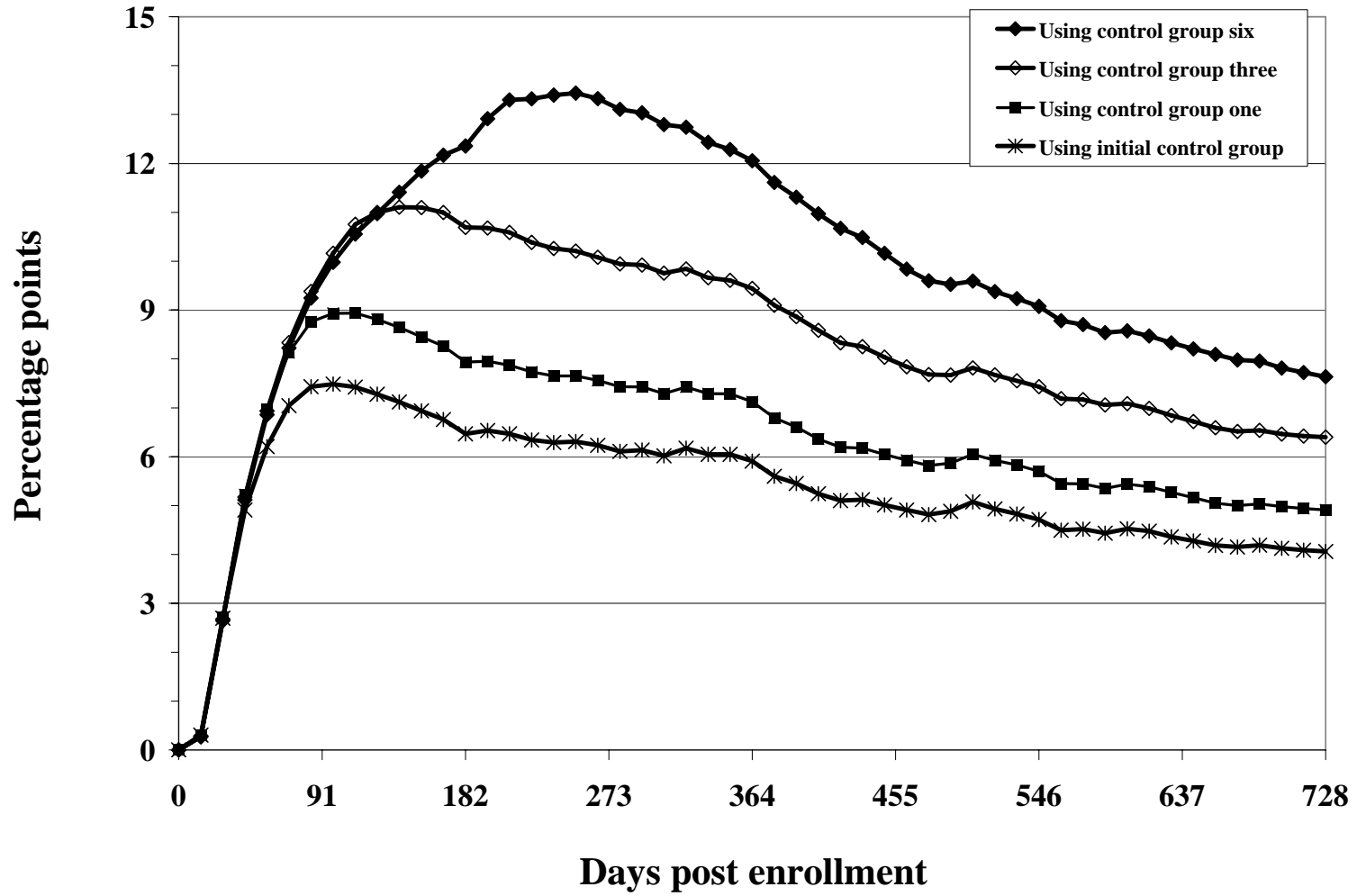


Figure 6

The PID with Various Covariates Included Using the Control Group Six

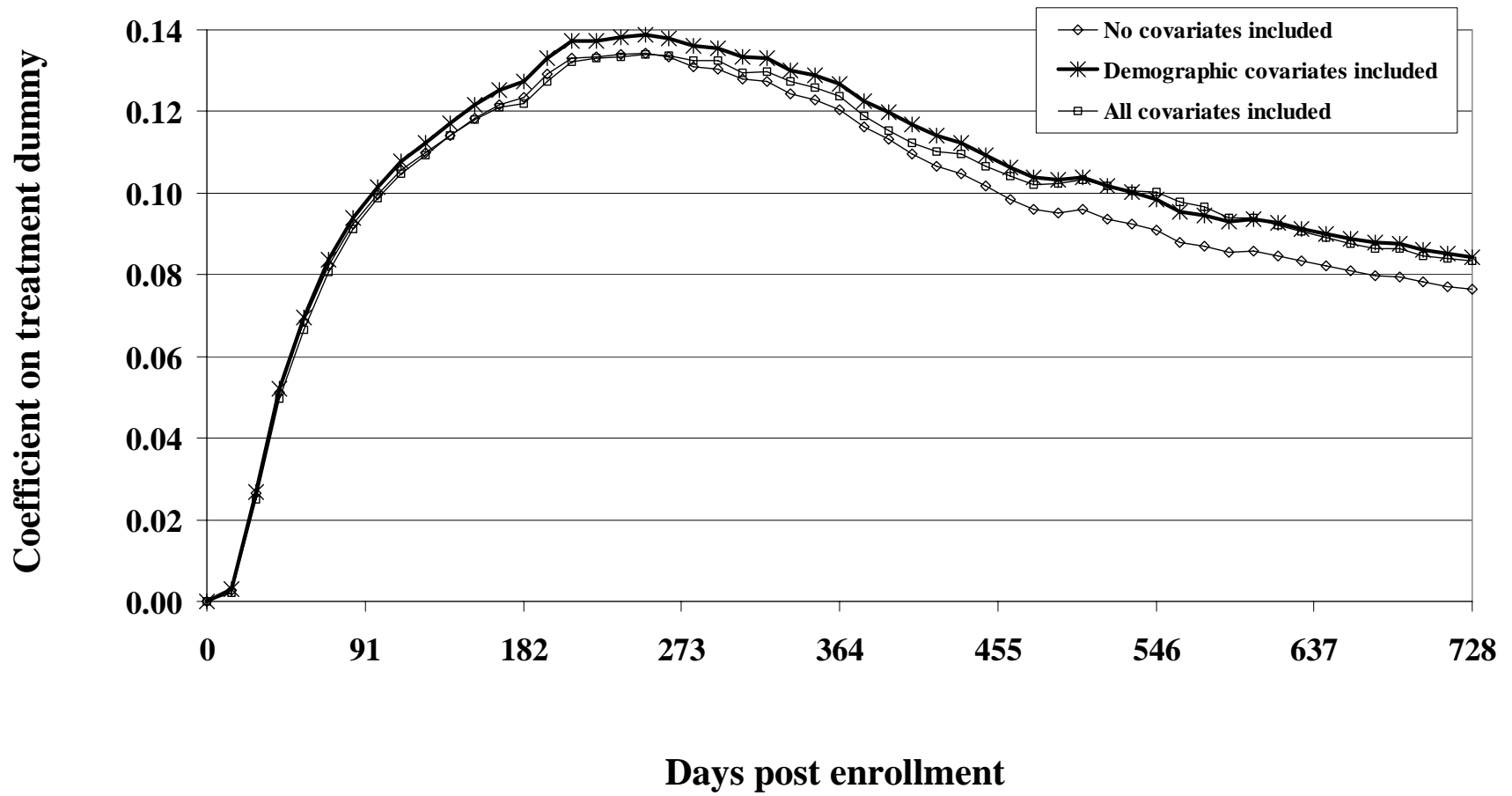


Figure 7

The PID with Started a Job and Permanently Exited Welfare as the Dependent Variable and All Covariates

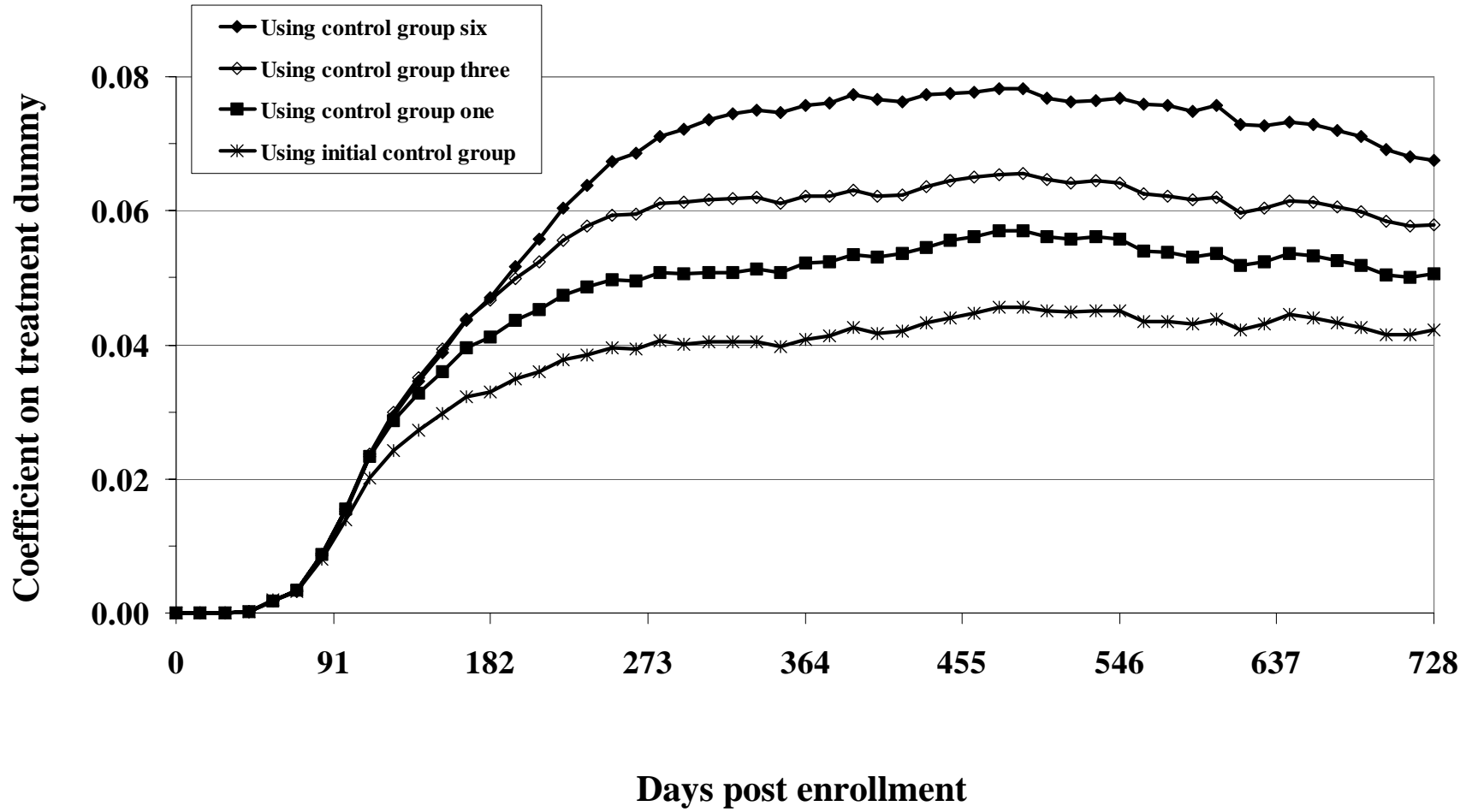


Figure 8

The PID with Permanently Exited Welfare as the Dependent Variable and All Covariates

