SRDC Working Paper Series 04-05

An Econometric Analysis of the Impact of the Self-Sufficiency Project on Unemployment and Employment Durations

The Self-Sufficiency Project

Jeffrey Zabel
Tufts University

Saul Schwartz
Carleton University

Stephen Donald
University of Texas

July 2004

SOCIAL RESEARCH AND DEMONSTRATION CORPORATION

This paper is part of the Social Research and Demonstration Corporation's program of analysis for the Self-Sufficiency Project (SSP) sponsored by the federal Department of Human Resources and Skills Development Canada (HRSDC). The Self-Sufficiency Project is sponsored by HRSDC. This paper was produced for the Social Research and Demonstration Corporation (SRDC). The opinions expressed herein are the author's and do not necessarily reflect those of SRDC or HRSDC.

Table of Contents

Tables and Figures

Acknowledgements

We would like to thank Kelly Foley and Shawn de Raaf for excellent research assistance. This paper was informed by the work done by the SSP research team at MDRC and we acknowledge their influence (though they should not, of course, be blamed for any shortcomings in this work).

Abstract

The Self-Sufficiency Project (SSP) was a Canadian research and demonstration project that investigated how to "make work pay," by supplementing the earnings of long-term income assistance (IA) recipients. The long-term goal of SSP was to get lone parents permanently off income assistance and into the paid labour force. The purpose of this study is to evaluate the short-term and long-term impact of SSP on the duration of employment and unemployment. This paper focuses on generating estimates of the "effect of the treatment on the treated" where the "treated" are those in the program group who actually received an SSP supplement (take-up program group). Following the work of Ham and LaLonde (1996) this paper estimates a joint model of unemployment and employment durations that controls for unobserved heterogeneity and non-random selection into work. This provides an unbiased estimate of the impact of SSP on unemployment and employment durations, and finds evidence of significant short-term impacts of SSP on unemployment and employment durations. It is also found that SSP appears to have a long-term positive impact on the employment rate of the take-up program group. This appears to be due to the long-term decrease in the probability of exit from employment for take-up program group members.

Introduction

The Self-Sufficiency Project (SSP) was a Canadian research and demonstration project that attempted to "make work pay," by supplementing their earnings of long-term income assistance (IA) recipients.[1] Lone parents on IA qualified for a generous earnings supplement if they took up full-time work and left the welfare rolls within 12 months of entering the project. Once qualified, they received a supplement that roughly doubled their pre-tax earnings during periods of full-time work in the next three years. Participation was entirely voluntary; individuals could choose not to participate without penalty. Those who qualified for the supplement could return to income assistance when they were not working; if they subsequently found full-time work within their three-year eligibility period, they could again receive the earnings supplement. Participants received only minor services beyond the financial incentive provided by the earnings supplement.

It is well known that the benefits paid to welfare recipients are reduced, often dollar-for-dollar, as their labour market earnings rise. Thus, work may not "pay" because increased earnings are offset by lower social assistance payments. The short-term goal of SSP was to encourage lone parents on IA — 95 per cent of whom are lone mothers — to take up paid work and increase their total income significantly by doing so.

To evaluate whether participation in SSP resulted in increased earnings and employment, SSP was designed as a social experiment with participants randomly divided into a program group and a control group. The rigorous evaluation methods associated with social experimentation were an integral part of the project from its inception. A series of surveys — a baseline survey at the point of random assignment and follow-up surveys 18, 36 and 54 months after random assignment — was undertaken by Statistics Canada.

The long-term goal of SSP was to get lone parents to permanently leave IA and enter into the paid labour force. Program designers theorized that the SSP earnings supplement would induce women who would otherwise have stayed on IA to enter the labour force. Once in the labour force, program group members would be more likely to stay off the welfare rolls, either because they would experience large enough increases in labour income so they would be better off working (even without the supplement) or because they would come to appreciate the non-monetary benefits of work. The purpose of this study is to evaluate the short-term and long-term impact of SSP on the duration of employment and unemployment. We focus on generating estimates of the "effect of the treatment on the treated" where the "treated" are those in the program group who actually received an SSP supplement.

The experimental nature of the SSP evaluation provides clear and unambiguous answers to a variety of policy-relevant questions. Randomization ensures that the experimental status

[1]The project took place in two Canadian provinces (British Columbia and New Brunswick) between November 1992 and December 1999, and was funded by Human Resources and Skills Development Canada (HRSDC). Operating outside the provincially run social assistance systems, SSP was managed by the Social Research and Demonstration Corporation (SRDC). Evaluations undertaken thus far have been carried out by Manpower Demonstration Research Corporation (MDRC) and SRDC.

of participants is statistically independent of any observed and unobserved characteristics. For that reason, simple comparisons of means and proportions provide methodologically powerful and statistically valid answers to certain questions. For example, SSP did not lead to greater employment or earnings for the program group, as compared with the control group:

> *... in the middle of the fifth year [after random assignment], about 27 per cent of the control group worked full time compared with 28 per cent of the program group, and average earnings for both groups were close to $500 per month. Moreover, a comparison of IA use in the sixth year found virtually no difference between the program and control groups.*
> (Michalopoulos et al., 2002, p. ES-12)

Even though the program and control groups had similar rates of employment and welfare use at the end of the experiment, SSP had important effects on program group members — it substantially raised average employment rates and earnings while the experiment was operating (and while some program group members were eligible for the earning supplement). During the experiment, program group members and their families had considerably higher incomes, on average, than control group members and their families.

Despite the power of experimental evaluation, many policy-relevant issues cannot be answered by simple program–control comparisons. Among them are the effects of the program on wages, unemployment durations, and employment durations prior to, during, and after the supplement period. These issues are all important to policy development. Many welfare policies are aimed at getting welfare recipients into the paid labour force (shortening the duration of an unemployment spell) while paying relatively little attention to policies that enable welfare recipients to stay in their new jobs (lengthening the duration of employment). Perhaps most importantly, if SSP program group members experienced neither wage growth nor increased labour force attachment, then the case for greater self-sufficiency is seriously undermined. Simple program–control comparison cannot fully evaluate these impacts.

Researchers must turn to non-experimental methods to address issues such as these because understanding them depends on comparisons of program group members who worked with control group members who worked. Such comparisons are *not* unbiased, even if program and control members have been randomly assigned. In their study of the experimental effects of training on unemployment and employment durations, Ham and LaLonde (1996) note:

> *... simple comparisons between trainees' and controls' employment rates yield unbiased estimates of training's effect on the probability of employment. However, a similar comparison between the durations of trainees' and controls' employment and unemployment spells, or their hazard rates out of those spells, yields potentially biased and economically misleading estimates of the effect of training. Although program administrators used random assignment to create the treatment and control groups, there is no reason to believe that the treatments and controls experiencing subsequent employment and unemployment spells are random subsets of the experimental sample.* (p. 178)

This paper develops an econometric framework that allows us to obtain unbiased estimates of the long-term impact of SSP. As with most econometric modelling, however, the properties of the estimates depend crucially on the assumption that the model is correctly specified. Following the work of Ham and LaLonde (1996) a joint model of unemployment and employment durations that controls for unobserved heterogeneity and non-random selection into work is estimated. This provides an unbiased estimate of the impact of SSP on unemployment and employment durations.

Evidence of significant short-term impacts of SSP on unemployment and employment durations is found, where "short-term" means, "while the take-up program group was eligible for the earnings supplements." We also found that SSP appears to have a long-term positive impact on the employment rate of the take-up program group, where "long-term" means "beyond the period of program eligibility." This appears to be due to the long-term decrease in the probability of exit from employment for take-up program group members.

In the second section, we discuss some of the basic information about SSP that has been released in previous analyses by Lin, Robins, Card, Harknett, and Lui-Gurr (1998), Michalopoulos, Card, Gennetian, Harknett, and Robins (2000), Morris and Michalopoulos (2000), and Michalopoulos et al. (2002). We derive other basic results from our own calculations. In the third section, we describe the data and present some descriptive analysis of the initial unemployment and employment behaviour of the program and control groups. We estimate models of unemployment and employment duration in the fourth section. The fifth section simulates the long- and short-run impacts of SSP. The sixth section summarizes our sense of how SSP affected the durations.

Basic Information About the Self-Sufficiency Project

One of the most important goals of the Self-Sufficiency Project (SSP) was to increase the rate at which lone parents on social assistance left their provincial welfare programs and took up full-time work. It was decided early in the project to incorporate a random assignment evaluation into the project.[2] In total, 6,028 lone parents completed the baseline survey and were randomly assigned between November 1992 and March 1995. Of the 6,028 participants 2,880 were assigned to the program group, 2,849 were assigned to the control group, and 299 were assigned to a special program called SSP Plus. In this paper, we use the results of the follow-up survey administered by Statistics Canada roughly 54 months after each participant was randomly assigned either to the program group or to the control group. Excluding those who were assigned to SSP Plus, the number of 54-month survey respondents was 4,852 — 2,460 respondents were in the program group and 2,392 were in the control group. We exclude another 356 program and control group members because they were working full time in the month prior to random assignment. A few other participants had missing values on one or more of the survey questions used in the multivariate models that we estimate and are excluded for that reason. The resulting sample size for our analysis is 4,420 individuals (2,255 in the program group and 2,165 in the control group).

The data are organized into "relative months" or months since random assignment. The "first" month for each SSP participant is the month in which they were randomly assigned, which could be in any calendar month from November 1992 to March 1995. Because all participants who were employed full-time in the month prior to random assignment were excluded, all sample members are considered to be unemployed in the first month after random assignment. In effect, employment status in month t is determined in month t-1.

After random assignment, SSP participants were followed until the end of a follow-up period marked by the date of the 54-month survey interview. For a number of participants, however, the survey occurred before 54 calendar months had elapsed, so that complete data on unemployment and employment durations is available for only 51 months after random assignment. Thus, in this paper, the follow-up period is 51 months in length. Since employment status in the 51st month determines employment status in the 52nd month, there are 52 months over which spells of employment and unemployment can be calculated.

[2]In each month of the enrolment period, Statistics Canada used administrative data provided by the provincial social assistance agencies in British Columbia and New Brunswick to identify all recipients who: (a) were lone parents, (b) were 19 years of age or older, and (c) had received benefits in the current month and 11 of the 12 preceding months. A random sample of those who met these criteria were then contacted, interviewed by a Statistics Canada interviewer, and invited to participate in the project. However, "to participate in the project," meant either being in the program group that would be offered the earnings supplement, or in the control group that would receive nothing. The interviewers carefully explained the project to each potential participant. Following the explanation, the person was asked to sign an "informed consent" form verifying that she understood the project and giving researchers access to administrative data. The baseline survey was completed at this time. Roughly 90 per cent of the fielding sample completed the baseline survey and signed the informed consent.

Each survey asked about every job that the respondent held between the survey date and the last time the respondent had been surveyed. For example, the 18-month survey collected detailed information on every job held since the date of the baseline survey. Respondents were asked if they were still working in jobs that they held at the time of the baseline survey and were asked about every new job held since then.[3]

Given this evaluation design, it is a simple matter to assess the impact of SSP on *some* dimensions of participants' labour force participation. From the survey information, any number of labour force "outcomes" can be defined. In reports published thus far, the emphasis has been on comparing the proportion working full time *at a single point in time*.

Using data from the follow-up surveys, Manpower Demonstration Research Corporation (MDRC) created a series of 0-1 variables, one for every month after random assignment, that took the value "1" if the respondent worked 30 or more hours in any week during the month.[4] In each month after random assignment, the proportion of the program group who were employed full time (as just defined) could be calculated and compared with the proportion of the control group who were employed full time. Lin et al. (1998) report that the largest impact of SSP on full-time employment occurred after random assignment, in Quarter 5. At that point, the average quarterly proportion working full time was 29.3 per cent in the program group and 14.0 per cent in the control group, a difference of 15.3 percentage points.[5]

KEY FEATURES OF THE SSP EARNINGS SUPPLEMENT

Full-time work requirement. Supplement payments were made only to eligible single parents who worked full time (an average of at least 30 hours per week over a four-week or monthly accounting period, whether in one or more jobs) and who were not receiving income assistance (IA).

Substantial financial incentive. The supplement is calculated as half the difference between a participant's earnings from employment and an "earnings benchmark" set by the program for each province. The benchmark was set at a level that would make full-time work pay better than IA for most recipients. During the first year of operations, the benchmark was $37,000 in British Columbia. The benchmark was adjusted over time to reflect changes in the cost of living and in the generosity of IA. The supplement was reduced by 50 cents for every

[3]We note that "seam effects" may exist in the SSP data. Seam effects arise because respondents tend to "move" starting and ending dates of events to the dates on which they were interviewed. SRDC has noted this seam problem in the SSP data, and in work done in 2002, taken steps to mitigate the problem. We have not, however, made any adjustment for this potential problem in this paper.

[4]Respondents were asked how many hours per week they "typically" worked on each job. If the person was working on a job whose typical weekly hours were 30 or more during a month, then she was assigned the value "1" on the full-time indicator for that month.

[5]The Lin et al. result was calculated using all those who responded to the 18-month survey, which included 2,645 program group members and 2,643 control group members. Michalopoulos et al. (2000) calculate the same impact using the sample of those who responded to the 36-month survey. In Quarter 5 for that sample, 29.6 per cent of the program group worked full time as compared with 14.8 per cent of the control group, a difference of 14.7 percentage points. Lin et al. chose to aggregate the monthly impacts into quarterly impacts. The percentage of the program and control group working in each quarter was calculated as the average of the percentages in each of the three months making up the quarter. For example, the quarterly average "employed full time" for the program group in the first quarter — defined as the three months following random assignment — was 13.0 per cent, the average of 11.5 per cent in Month 1, 12.9 per cent in Month 2, and 14.7 per cent in Month 3.

dollar of increased earnings. Unearned income (such as child support), earnings of other family members, and number of children did not affect the amount of the supplement.

Targeted at long-term recipients. Eligibility for the supplement was limited to long-term welfare recipients (with at least one year of IA receipt).

One year to take advantage of the offer. Eligible IA recipients were informed that they could sign up for the supplement if they found full-time work within the 12 months following random assignment. If they did not sign up within 12 months, they would be permanently ineligible to receive the supplement.

Three-year time limit on supplement receipt. A person could collect the supplement for up to three calendar years from the time it was received, as long as that person was working full time and not receiving IA.

Voluntary alternative to welfare. Participants could not receive IA payments while receiving the supplement. Participation in the supplement program was voluntary, and even after beginning supplement receipt, people could decide at any time to return to IA as long as they gave up supplement receipt and met the eligibility requirements for IA. They could also renew their supplement receipt by going back to work full time at any point during the three-year period in which they were eligible to receive the supplement.

Simple Descriptive Measures of the Duration of Unemployment and Employment

In this section, we describe the duration of unemployment and employment spells with only minimal modeling of the processes that determine the durations. Where possible, we present experimental contrasts, involving all members of the Self-Sufficiency Project (SSP) program and control groups. When this is not possible — when, for example, we describe the non-random subsets of the program and control groups who worked full time during the follow-up period — we describe the durations without accounting for any self-selection or unobserved heterogeneity. Our description includes estimates of hazard rates and mean durations. The estimation of econometric models that account more formally for self-selection, unobserved heterogeneity and duration dependence appears in the next section.

In this paper "Employment" means working in a job that normally involves 30 or more hours per week; that is, "employment" is really "full-time employment." Furthermore, respondents are "unemployed" in this paper when they are not working full time; part-time workers are counted as "unemployed." There are two reasons for this choice. The first is that the explicit goal of SSP was to increase full-time employment. The second is that, empirically, "SSP increased full-time work primarily by persuading people who would not have worked otherwise to work full time" (Michalopoulos et al., 2002, p. ES-9). In describing employment and unemployment spells, we therefore adopt the terminology sometimes used in the Manpower Demonstration Research Corporation (MDRC) and Social Research and Demonstration Corporation (SRDC) experimental evaluations.[6]

THE DURATION OF UNEMPLOYMENT SPELL FOLLOWING RANDOM ASSIGNMENT

We provide the definitions of a number of observable characteristics of SSP participants used in our analyses in Table 1. We give the means and standard deviations of these variables for all program and control group members in Columns 1 to 4 of Table 2. One can see that the mean values are very similar across the two groups. For each variable, we conduct t-tests for the equality of the population means. Only for MONTHIA — the number of months on income assistance in the three years prior to the baseline survey — do we reject the null hypothesis at the one per cent significance level (p-value = 0.0058) though note that even here the difference is only 0.65 months. These results suggest that random assignment "worked" and that the two groups are quite similar in terms of their observable characteristics.[7]

[6]It would perhaps be clearer if we substituted "non-full-time employment" for "unemployment" but the improvement in clarity would be counterbalanced by such awkward language.

[7]See Lin et al. (1998), p. 125, for similar analyses that also confirm that random assignment "worked."

Table 1: Variable Descriptions

Variables that do not change over time	
PROGRAM	1 if program group member; 0 if control group member.
TAKEUP	1 if program group member who found work in qualifying period (13 months following survey); 0 otherwise.
NON-TAKEUP	1 if program group member who did not find work within qualifying period; 0 otherwise.
IDBC	1 if British Columbia respondent; 0 if New Brunswick respondent.
NKIDS	Number of children in respondent's household at baseline.
YGCHLE4	1 if a child in the household is less than or equal to four years; 0 otherwise.
BLTHS	1 if respondent has less than a high school education at baseline; 0 otherwise.
NVRMAR	1 if respondent is single, never married at baseline; 0 otherwise.
MONTHIA	Number of MONTHS on IA in the three years prior to the baseline interview.
EMOPROB	1 if respondent had limited activity due to a long term emotional, psychological, nervous or mental health condition or problem at baseline; 0 otherwise.
MBENEFIT	Monthly IA benefit in the month immediately prior to random assignment (in hundreds).
EEMPFTBR	Employed and working full time (total hours/week at least 30 hours in all jobs) in at least one of the 12 months preceding the baseline survey (0-1 variable).
BAGE1922	1 if respondent's age is 19 to 22 years of age at baseline; 0 otherwise (omitted).
BAGE2325	1 if respondent's age is 23 to 25 years of age at baseline; 0 otherwise.
BAGE2629	1 if respondent's age is 26 to 29 years of age at baseline; 0 otherwise.
BAGE3034	1 if respondent's age is 30 to 34 years of age at baseline; 0 otherwise.
BAGE3539	1 if respondent's age is 35 to 39 years of age at baseline; 0 otherwise.
BAGEGE40	1 if respondent's age is 40 or more years of age at baseline; 0 otherwise.
Variables that change over time	
Pre-incentive	1 in months when program group members could qualify for the SSP supplement by finding full-time work; 0 otherwise.
Post-incentive	1 for program group members who had qualified for the SSP supplement in months after their eligibility for the supplement had expired; 0 otherwise.
Monthly unemployment rate	The unemployment rate for women, 15+, in the respondent's province, for each calendar month corresponding to each month after random assignment for the respondent.

Table 2: Variable Means and Standard Deviations

	Control Group (n=2,165)		Program Group (n=2,255)	
Variable	Mean (1)	St. Dev. (2)	Mean (3)	St. Dev. (4)
IDBC	0.52	0.50	0.51	0.50
NKIDS	1.71	0.89	1.71	0.84
YGCHLE4	0.50	0.50	0.47	0.50
BLTHS	0.55	0.50	0.54	0.50
NVRMAR	0.50	0.50	0.49	0.50
MONTHIA	29.53	8.03	30.18	7.71
EMOPROB	0.078	0.27	0.085	0.28
BENEFIT ($00)	8.47	2.56	8.47	2.52
EEMPFTBR	0.099	0.30	0.10	0.30
BAGE2325	0.14	0.35	0.12	0.33
BAGE2629	0.17	0.37	0.17	0.38
BAGE3034	0.22	0.42	0.23	0.42
BAGE3539	0.16	0.37	0.17	0.38
BAGEGE40	0.17	0.38	0.17	0.38

Note: Participants who were employed full time in the month before random assignment are excluded from this table. The vast majority of such participants also worked in the first month after random assignment and thus are likely not to have had a period of unemployment after random assignment. Also excluded from this table are participants who have missing values on covariates included in later analyses.

Table 3 shows the percentages of the program and control groups in each of a series of unemployment duration categories. The relevant unemployment spell here is the *first* spell, one that is measured (for all sample members) from the date of random assignment.[8]

Table 3: Length of Time From Random Assignment to First Full-Time Job

Category	Program Group	Control Group	Impact
0 < Duration < 6 Months	16.98	9.42	7.56
6 ≤ Duration < 12 months	16.78	7.53	9.25
12 ≤ Duration < 18 months	5.90	6.28	-0.38
18 ≤ Duration < 24 months	2.88	4.20	-1.32
24 ≤ Duration < 30 months	3.41	5.03	-1.62
30 ≤ Duration < 36 months	2.97	3.46	-0.19
36 ≤ Duration < 42 months	3.59	4.48	-0.49
42 ≤ Duration < 48 months	2.00	2.96	-0.96
Duration of 48 months or longer	0.75	1.43	-0.68
No full-time work after RA	44.75	55.20	-10.45
Total	100.00	100.00	0.00
Sample size (n=4,420)	2,255	2,165	—

Note: Participants who were employed full time in the month before random assignment are excluded from this table.

A significantly higher percentage of the control group, compared with the program group, was unemployed for the entire 52-month follow-up period. In the program group, 44.7 per cent never found full-time work while 55.2 per cent of the control group never

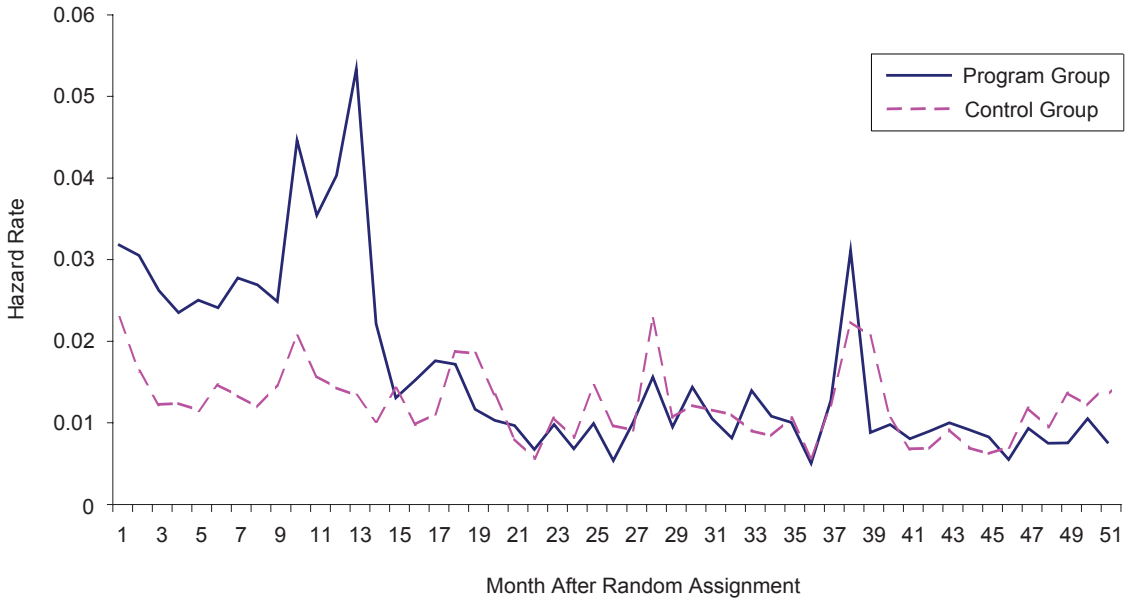
[8]We also chose to begin the spell of unemployment on the date of random assignment, thus "left censoring" the data.

found full-time work. The third column of the Table 3 shows the impact of the program, measured as the difference between the program and control group percentages.

The impact estimates in Table 3 are *experimental* impacts because all 4,420 program and control group members are distributed among the categories. That SSP had the effect of lowering the duration of unemployment for some program group members can be seen by the greater percentage of program group members in the two shortest unemployment duration categories. There is also evidence of "control group catch-up" here, as a greater percentage of the control group had unemployment durations of more than 13 months.

One way of viewing the difference in the unemployment behaviour of the program and control groups is to calculate the empirical hazard functions using the Kaplan–Meier product-limit estimator. Figure 1 shows the empirical hazards for the program and control groups in each of the 51 months for which hazard rates can be calculated; Figure 2 shows the difference between the hazard rate functions of the program and control groups.[9] The greater speed with which the program group found full-time employment can be seen by noting that the height of the empirical hazard function for the program group — representing the probability of finding full-time employment in any month for those who had not yet found full-time employment — is greater than that of the control group in each of the 12 months after random assignment. Figure 3 presents the same information as a survivor function, representing the probability of remaining in unemployment until the time period in question. The survivor function for the control group lies entirely above the survivor function of the program group, indicating that the control group was more likely to remain unemployed.

Figure 1: Empirical Hazard Rates for First Spell of Unemployment — Program and Control Groups



[9]Hazard rates can be calculated only for 51 months of the 52-month follow-up period. While we know employment status in the 52nd month (because it is determined by employment status in the 51st month) we do not know how many "failed" in that month.

Figure 2: Program–Control Difference in Hazard Rates

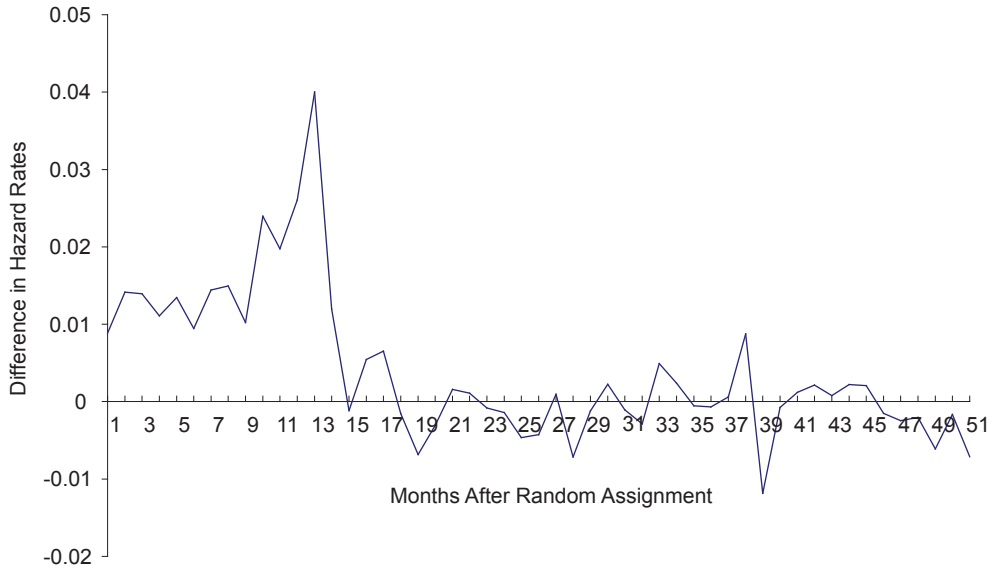
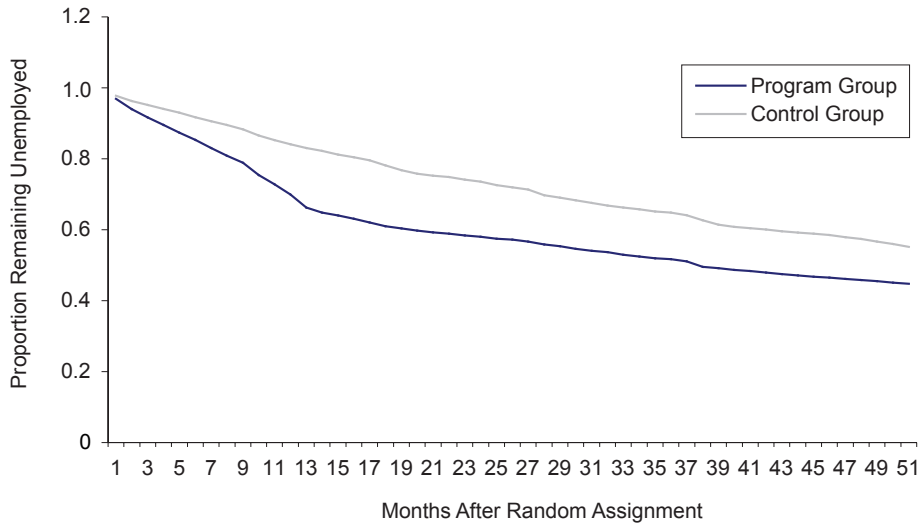


Figure 3: Kaplan–Meier Survivor Functions — First Unemployment Spell



Since program group members had a much greater incentive to find full-time work in the first 12 months, it is not surprising that their empirical hazards are higher than those of the control group. After the end of the twelve-month qualification period, however, the program group members who have not yet found a full-time job no longer have a greater financial incentive to find one than control group members. Figure 2 shows that the empirical hazards are no longer consistently higher than those of the control group.

The expected duration of unemployment is of particular interest to policy-makers. Several options are available for calculating average duration with censored data (Klein & Moeschberger, 1997). We use two of these techniques. The first assumes that all those who

were unemployed at the end of 51 months became employed immediately thereafter.[10] More realistically, however, we would expect that participants would gradually become employed as the months proceed. The second option captures this idea by assuming that the number of remaining unemployed individuals declines in the months after the follow-up period ends, according to a particular functional form. Table 4 shows these estimated durations of unemployment using both methods.[11]

Table 4: Estimated Average Duration of Unemployment (Months)

	Program Group	Control Group	Experimental–Control Diff.
Assuming all censored observations become employed at the censoring point	32.05	38.35	6.30
Assuming the survivor function declines exponentially	60.99	86.65	25.66

The first method almost certainly underestimates the eventual duration of unemployment for those who did not find a full-time job in the follow-up period. Using that method, the expected duration for the program group is about 32 months while the control group duration is estimated to be about 38 months. The second method allows the predicted durations to be very long for those who did not find a full-time job but does not allow for control group "catch-up." As a result, the predicted average durations are quite long and program–control differences become quite large. The expected duration for the program group is 61 months as compared with 87 months for the control group.

THE DURATION OF FIRST FULL-TIME EMPLOYMENT SPELL

We now describe the first spell of full-time employment for SSP participants. There are several important differences between the analysis of employment spells carried out here and the above analysis of the first unemployment spell. The first unemployment spell began at the point of random assignment for all participants and the program and control group differed only in that the program group had a stronger incentive to find full-time employment. By contrast, the first spell of full-time employment (if any) can begin at any point between random assignment and the end of the 51st month after random assignment. And, while there are still two groups that face different incentives, the two groups are different. We refer to program group members who qualified for the SSP supplement as the "take-up" group. They have a stronger incentive to stay employed than do control members or program group members who found their first full-time job after the deadline for qualifying for the SSP supplement. That difference in incentives ends, however, after the take-up program group members have received the supplement for 36 months. Program group members who did not qualify for the SSP supplement and control group members may

[10]Given that expected duration is not observed for roughly half of the sample (because they had not found a full-time job within 51 months of random assignment) we must either make some sort of assumption about their eventual employment or abandon the effort to calculate average duration.

[11]Many software programs used to estimate survivor functions provide estimates of average duration based on these methods. Those in Tables 3, 5, 7 and 8 were calculated using *Stata*.

face the same incentives to stay employed but the groups may no longer be similar to each other. It is possible that the most "job-ready" of the program group members qualified for the supplement, leaving only the less "job-ready" in search of jobs after the qualifying period. If so, this group should be less likely, on average, than the control group to maintain any employment that they found.

Table 5 contains an experimental analysis of the duration of first full-time employment, showing the proportion of the program and control groups in each of several duration categories. If SSP encouraged only short-term jobs taken for the purpose of qualifying for the supplement, most of the 10 percentage point difference in the proportion that worked full time would be concentrated in the shortest duration categories, "less than 6 months" and "6-12 months." By contrast, Table 5 shows that very little of the impact is in the shorter duration categories, with most spread across the longer categories. These results support the interpretation that the effect of SSP was to encourage full-time work that lasted for a relatively long period of time.

Table 5: Duration of First Spell of Full-Time Employment (All Participants)

Category	Program Group	Control Group	Impact
1 < Duration < 6 Months	23.24	24.02	-0.78
6 ≤ Duration < 12 months	9.76	7.44	2.32
12 ≤ Duration < 18 months	6.16	5.08	1.08
18 ≤ Duration < 24 months	4.26	2.82	1.44
24 ≤ Duration < 30 months	2.57	1.66	0.91
30 ≤ Duration < 36 months	2.57	1.39	1.18
36 ≤ Duration < 42 months	3.37	1.29	2.08
42 ≤ Duration < 48 months	2.13	0.69	2.44
Duration of 48 months or longer	1.20	0.42	0.78
No full-time work after RA	44.75	55.20	-10.45
Total	100.00	100.00	0.00
Sample size (n=4,420)	2,255	2,165	—

The SSP program group therefore acquired considerably more work experience than the control group. And, according to Michalopoulos et al. (2000):

> *If that additional work experience translated into higher wages, different attitudes about work, or both, the impacts of the program might persist after the constant encouragement of the SSP supplement is withdrawn.* (p. 47)

Within the framework of social experimentation, researchers cannot go much further in analyzing the duration of full-time employment. This is because the sub-sample of individuals who ever found full-time work is not random. In Table 6, we compare the mean values of observable characteristics for this sub-sample to those of the full sample. One can see that these means differ. When we conduct t-tests for the equality of the population means for the ever and never full-time employed groups, we reject the null hypothesis at the one per cent significance level for almost all of the variables. These results validate that the two groups are significantly different in terms of their observable characteristics.

Table 6: Variable Means and Standard Deviations for the Full Sample and for Those Who Worked

	Full Sample (n=4,420)		All Those Who Worked Full Time After Random Assignment (n=2,216)	
Variable	Mean (1)	St. Dev. (2)	Mean (3)	St. Dev. (4)
IDBC	0.53	0.50	0.49	0.50
NKIDS	1.71	0.87	1.62	0.80
YGCHLE4	0.49	0.50	0.50	0.50
BLTHS	0.55	0.50	0.45	0.50
NVRMAR	0.49	0.50	0.52	0.50
MONTHIA	29.86	7.87	29.01	8.08
EMOPROB	0.10	0.30	0.055	0.23
BENEFIT ($00)	8.47	2.54	8.01	2.51
EEMPFTBR	0.10	0.30	0.14	0.35
BAGE2325	0.13	0.34	0.14	0.35
BAGE2629	0.17	0.38	0.18	0.38
BAGE3034	0.23	0.42	0.21	0.41
BAGE3539	0.17	0.37	0.16	0.37
BAGEGE40	0.17	0.38	0.13	0.34

Note: Participants who were employed full time in the month before random assignment are excluded from this table. The vast majority of such participants also worked in the first month after random assignment and thus are likely not to have had a period of unemployment after random assignment. Also excluded from this table are participants who have missing values on covariates included in later analyses.

Program group members who took up full-time employment may also differ systematically, in unobservable ways, from control group members who found full-time employment. Any comparison of means or proportions that is limited to program and control group members who worked full time will not be "comparing comparables." The analysis of the impact of SSP on those who qualified for the supplement — the "effect of the treatment on the treated" — must therefore account non-experimentally for the differences between the two groups.

We begin, however, with several analyses that do not account for any potential systematic differences between the two groups. As before, all of these analyses are non-experimental and thus it is possible that the results are as much a function of the techniques employed as they are of the underlying economic behaviour.

A simple non-experimental comparison is shown in Table 7, where we show the percentage of respondents in each duration category but restrict the sample to those who found full-time work (2,216 of the 4,420 sample members). We see that the control group was considerably more likely to have a first spell of full-time employment that was relatively short.

Table 7: Duration of First Spell of Full-Time Employment for the Program Group and Those Control Group Members Who Worked Full Time in at Least One Month After Random Assignment

Category	Program Group	Control Group	Impact
1 < Duration < 6 Months	42.05	53.61	-11.56
6 ≤ Duration < 12 months	17.66	16.60	1.06
12 ≤ Duration < 18 months	11.16	11.34	-0.18
18 ≤ Duration < 24 months	7.70	6.29	1.41
24 ≤ Duration < 30 months	4.65	3.71	0.94
30 ≤ Duration < 36 months	4.65	3.09	1.56
36 ≤ Duration < 42 months	6.10	2.89	3.21
42 ≤ Duration < 48 months	3.85	1.55	2.30
Duration of 48 months or longer	2.17	0.93	1.24
Total	100.00	100.00	0.00
Sample size (n=2,216)	1,246	970	—

Table 8 divides the sample of full-time workers into three groups — (1) the take-up program group; (2) program group members who did not qualify for the supplement (the non-take-up program group); and (3) all control group members who worked full time — to reflect the different incentives mentioned above. Table 8 shows that take-up program group members were considerably less likely than the other groups to have short first spells of employment. The employment durations of the control group seem to be slightly longer than the durations of the non-take-up program group. A complication in interpreting this table is that, as noted above, take-up program group members lose their eligibility for the SSP supplement 36 months after they started working full time.

Table 8: Duration of First Spell of Full-Time Employment for the Take-Up Program Group, the Non-Take-Up Program Group, and Those Control Group Members Who Worked Full Time in at Least One Month After Random Assignment

Category	Take-Up Program Group Members	Non-Take-Up Program Group Members	Control Group Members
1 < Duration < 6 Months	32.24	55.90	53.61
6 ≤ Duration < 12 months	17.28	18.18	16.60
12 ≤ Duration < 18 months	10.84	11.61	11.34
18 ≤ Duration < 24 months	7.96	7.35	6.29
24 ≤ Duration < 30 months	6.17	2.51	3.71
30 ≤ Duration < 36 months	5.90	2.90	3.09
36 ≤ Duration < 42 months	9.88	0.77	2.89
42 ≤ Duration < 48 months	6.58	0.00	1.55
Duration of 48 months or longer	3.16	0.77	0.93
Total	100.00	100.00	100.00
Sample size (n=2,216)	729	517	970

We can also compare the average lengths of the first full-time employment spells in the program and control groups. As above, those who are still in their first employment spell at the end of the follow-up period complicate this non-experimental comparison. For example, in the program group, 28 individuals had first employment spells that lasted 12 months. However, 6 of these 28 were still working during those first employment spells at the end of the 51st month after random assignment. That means that what seems, according to the data, to be a 12-month spell is likely to continue for more than 12 months. Like the computation of the means for the duration of post-random-assignment unemployment above, there are two alternative ways to compute means in this case. When we simply restrict the maximum employment spell to be number of months at which the observation is censored, the mean for all program group members who worked was 18.3 months and the corresponding mean for the control group was 16.3 months. This calculation clearly underestimates the mean employment duration for both program and control groups. When we assume an exponential pattern of job loss after the censoring point, the mean durations of the first spell of full-time employment in the program group is 21.9 months while the mean for the control group is 20.3 months. These means are shown in Table 9.

Table 9: Estimated Average Duration of Employment (Months) — Program and Control Group Members

	Program Group	Control Group	Experimental–Control Diff.
Assuming all censored observations become employed at the censoring point	18.3	16.3	2.0
Assuming the survivor function declines exponentially	21.9	20.3	1.6

For reasons just discussed, we expect the mean duration for take-up program group members to be greater than for non-take-up program group members. The longer duration is partly the result of the earnings supplement being available only to the former group and partly the result of the presumably greater employability of the supplement recipients. Because supplement eligibility lasted for only 36 months after the initiation of the supplement, the financial incentive to stay employed for the supplement receivers would vanish when supplement eligibility ended.

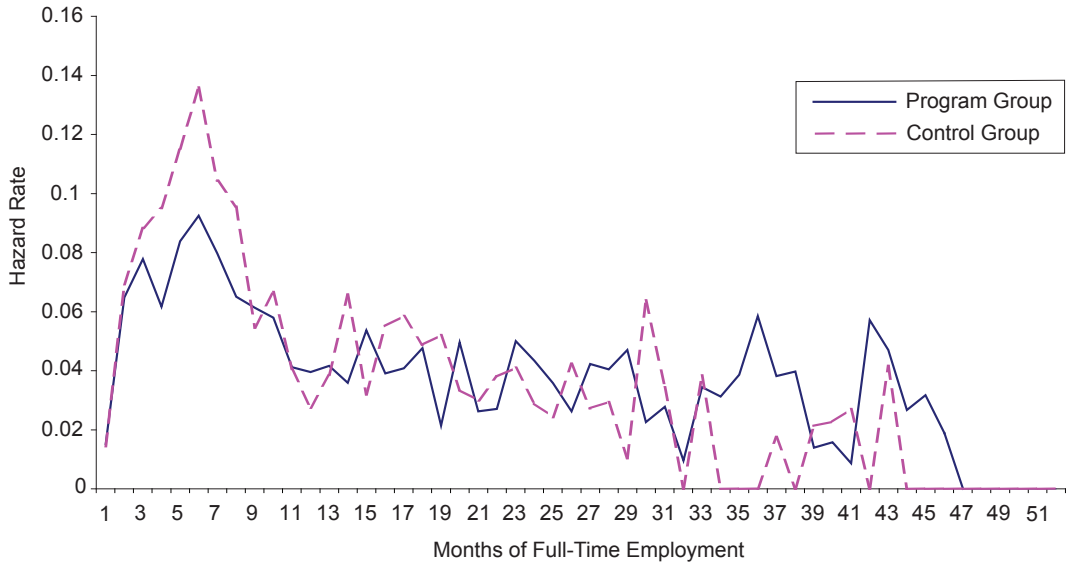
Assuming the duration ends at the censoring point, Table 10 shows that the take-up program group had an average first employment spell of 20.6 months, non-take-up program group members had a mean duration of 14.5 months, and as shown in Table 9, control group members had a mean duration of 16.3 months. Assuming exponential job loss, the mean duration for the program members who took up the supplement was 25.2 months, as compared with 16.4 months for program members who found work after Month 13, and 20.3 months for control group members.

Table 10: Estimated Average Duration of Employment (Months) — Take-Up Program Group, Non-Take-Up Program Group, and Control Group

	Take-Up Program Group	Non-Take-Up Program Group	Control Group
Assuming all censored observations become employed at the censoring point	20.6	14.5	16.3
Assuming the survivor function declines exponentially	25.2	16.4	20.3

Figure 4 shows the empirical hazard functions for the program and control groups, where the hazard is the end of the first full-time employment spell. Figure 5 shows the difference between the hazard rates of the program and control groups. The control group was considerably more likely to end their first employment spell fairly quickly, as indicated by their higher hazard of leaving full-time employment in the first months. Interestingly, program group members have a higher hazard of leaving full-time employment after 36 months (when their supplement eligibility ends).[12] This is some evidence that the take-up program group reverted back to behaviour similar to the control group after the incentive period ended. Figure 6 shows the Kaplan–Meier survivor functions associated with the program and control groups. The quicker job loss by control group members is illustrated by the lower values of the function that first appear at about Month 6; the quicker job loss by program group members as supplement eligibility runs out for some of them is illustrated by the narrowing of the gap between the two functions after about 36 months.

Figure 4: Empirical Hazard Rates for the First Spell of Full-Time Employment



[12]Because we are analyzing the *first* employment spell, we know that all take-up members whose duration was greater than 36 months were no longer eligible for the supplement and all those whose duration is less than or equal to 36 months remain eligible. Out of the 729 take-up program group members, 149 (20 per cent) had first spells that lasted more than 36 months.

Figure 5: Program–Control Differences in Empirical Hazard Rates for First Spell of Full-Time Employment

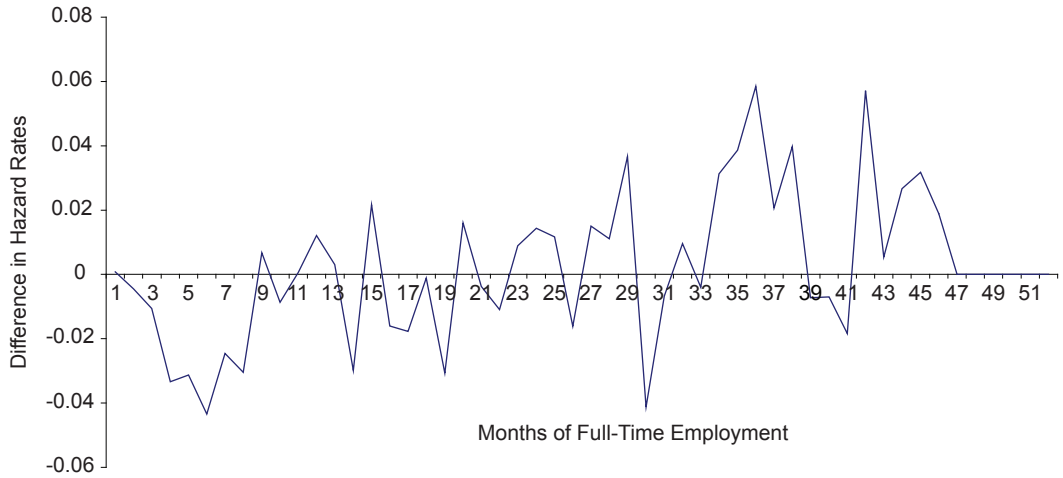


Figure 6: Kaplan–Meier Survivor Functions — Program and Control Groups

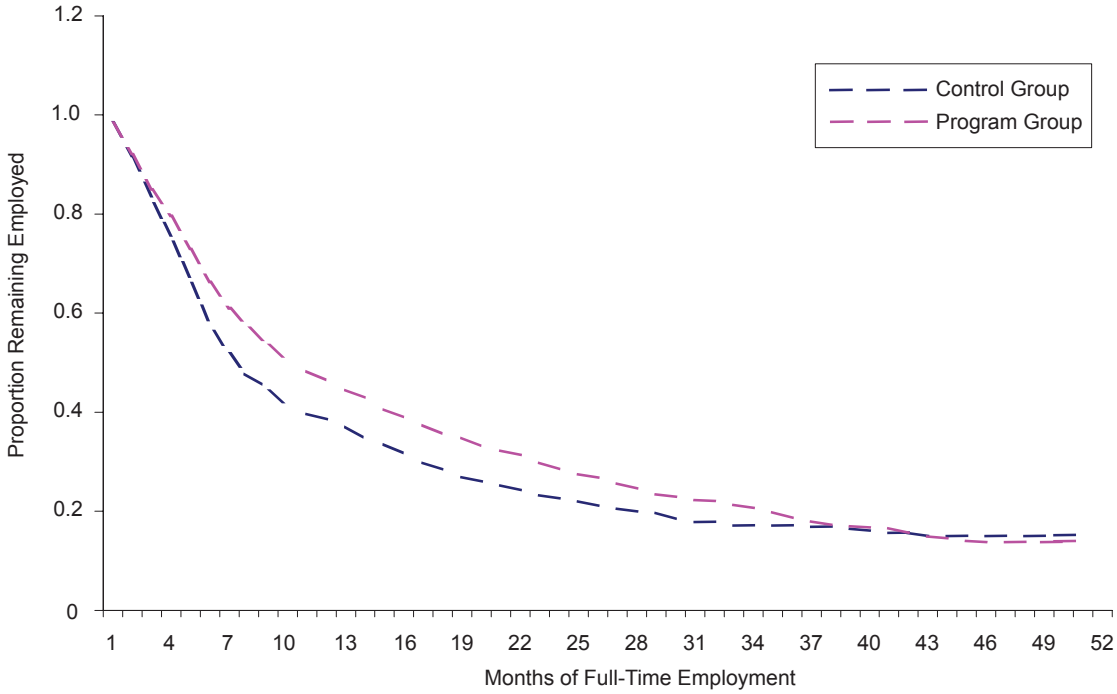
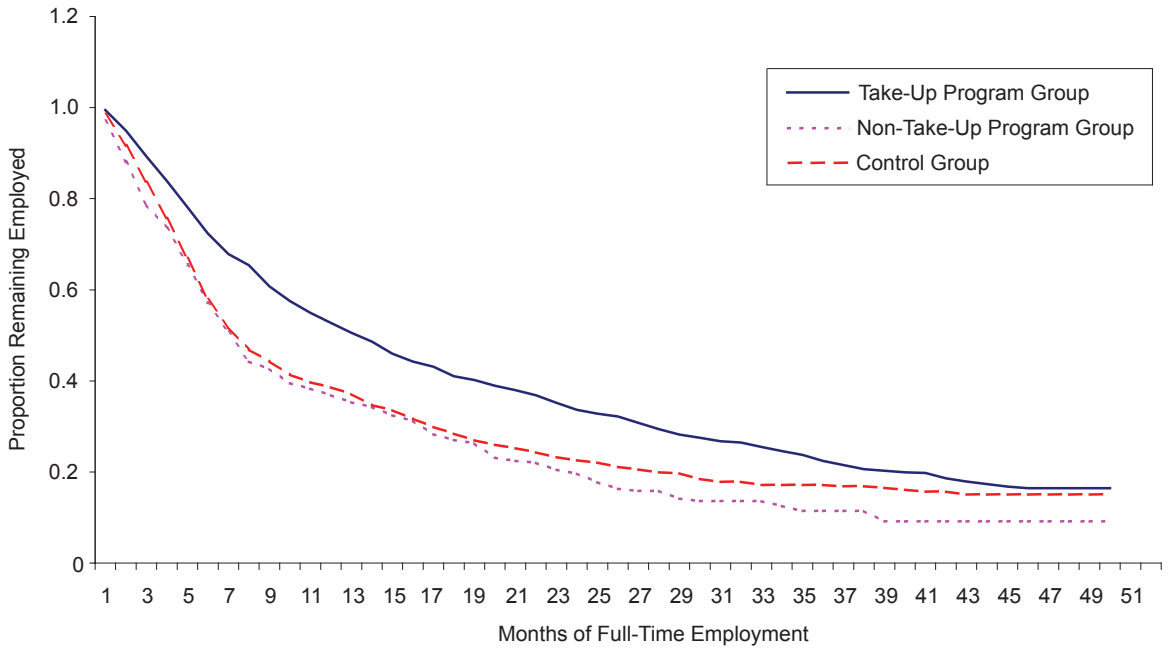


Figure 7 shows three survivor functions — for the take-up program group, the non-take-up program group, and the control group. Because of the key difference within the program group — some qualified for the supplement and others did not — it is important to control for that observed heterogeneity in the econometric modeling undertaken below.

Figure 7: Kaplan–Meier Survivor Functions — Take-Up Program Group, Non-Take-Up Program Group, and Control Group



MULTIPLE SPELLS OF UNEMPLOYMENT AND EMPLOYMENT

Many SSP participants had more than one unemployment or employment spell. It is important to include this information when evaluating the overall effects of the SSP program because one of the long-term goals of SSP was to increase the likelihood that the "take-up" group would work full time. In order to assess whether this goal was met we need to analyze the labour market behaviour of the take-up program group after their 36-month supplement eligibility period has ended. The first employment spell lasted less than 36 months for almost 80 per cent (580 of 729) of the take-up program group so spells beyond the first must be considered.

Table 11 shows the number of unemployment spells for each person while Table 12 shows the number of employment spells. Because we have excluded all participants who were employed full time in the month prior to random assignment, all sample members have at least one spell of unemployment. About one third of the control group and over 40 per cent of the program group had *more* than one spell of unemployment, meaning that they found a full-time job after random assignment and then left that job within the follow-

up period. Smaller percentages (24 per cent of the program group and 17.9 per cent of the control group) had more than one spell of full-time employment.

Table 11: Number of Unemployment Spells

Number of spells	Program Group	Control Group	Impact
1	57.43	67.76	-10.33
2	27.41	21.57	5.84
3	11.13	7.99	3.14
4+	4.03	2.68	1.35
Total	100.00	100.00	0.00
Sample size (n=4,420)	2,255	2,165	—

Note: Participants who were employed full time in the month before random assignment are excluded from this table; all of the remaining participants therefore have at least one spell of unemployment.

Table 12: Number of Full-Time Employment Spells

Number of Spells	Program Group	Control Group	Impact
0	44.75	55.20	-10.45
1	31.04	26.93	4.11
2	16.90	12.52	4.38
3	5.54	4.06	1.48
4+	1.77	1.30	0.47
Total	100.00	100.00	0.00
Sample size (n=4,420)	2,255	2,165	—

Note: Participants who were employed full time in the month before random assignment are excluded from this table.

The eligibility of each SSP participant for the SSP earnings supplement was determined by the date of their first eligible full-time employment spell. Either a person had started their first full-time employment spell within the eligibility period or they had not. Those who are eligible for the supplement should have shorter durations for the second and higher unemployment spells than would either control group members or non-take-up program group members. A complicating factor here is that even short spells of employment or unemployment may "cross" the point in time when the participant loses their eligibility for the supplement. For example, a supplement recipient might hold their first full-time job for 30 months, be unemployed for three months, and then find a new second full-time job. After three months on the second job, their eligibility for the supplement would end. Accounting for this phenomenon requires the more detailed econometric analysis that follows.

Figure 8 shows the Kaplan–Meier survival functions for the second and higher unemployment spells of these three groups. Here, and in the econometric modeling in the fourth section, we have combined all spells of unemployment and employment beyond the first. For example, if an individual had three unemployment spells, we would treat the two spells that follow the first as if they were the second spells of two different individuals. That is, we do not run separate models of the duration of the second, third, and higher unemployment spells. We do this because there is no reason, theoretically, to expect differences in the determinants of the durations of any higher order spells. Subject to the above caveats, Figure 8 makes clear that the supplement recipients were more likely to end

their spells of unemployment than were the control group. Both the control group and the take-up program group were more likely to end their unemployment spells than were program group members that were not eligible for the supplement (the non-take-up program group).

Figure 8: Kaplan–Meier Survival Functions — Second and Higher Spells of Unemployment

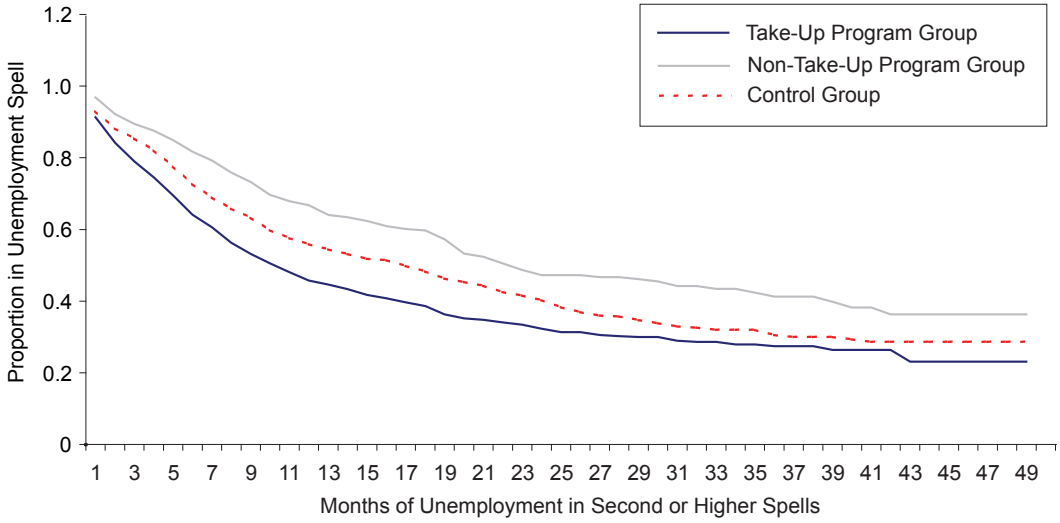
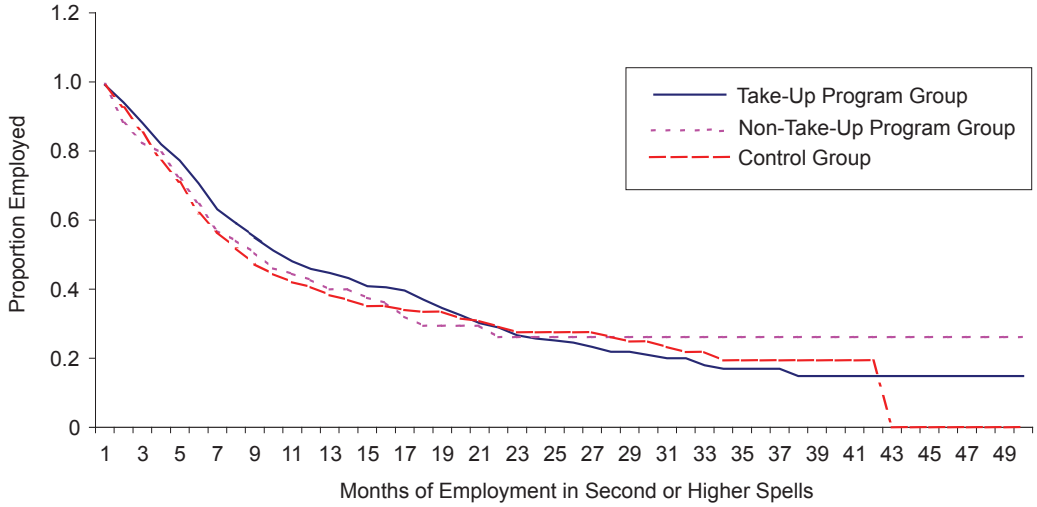


Figure 9 shows the Kaplan–Meier survival functions for the second and higher employment spells of these three groups. The survival functions for the three groups are roughly the same, indicating that the three groups had the same probability of ending their second and higher employment spells.

Figure 9: Kaplan–Meier Survival Functions — Second and Higher Spells of Employment



A Multivariate Model of Unemployment and Employment Durations

In this section we develop an econometric model of unemployment and employment durations that will allow us to obtain unbiased estimates of the Self-Sufficiency Project (SSP) impacts, assuming the specification of the model is correct. We can estimate the impact prior to uptake, during the incentive period, and after the incentive period is over. The latter estimate is particularly important since it is a measure of the long-term impact of SSP. This allows us to evaluate whether SSP achieved its goal of encouraging lone parents on welfare to leave income assistance permanently and move into the labour force.

The econometric analysis of SSP is an important advance over the data analysis of the previous section for at least two reasons. First, an important assumption in estimating the Kaplan–Meier hazard rate functions in Figures 1 to 9 is that the population at risk is homogeneous. One kind of participant is no more likely to "fail" than any other participant. This is clearly not true for the SSP program and control groups. For example, social assistance recipients who have worked in the recent past are more likely to be able to find work in the future than those who have no recent work history. The econometric analysis allows us to estimate unemployment and employment behaviour conditional on observable characteristics that affect this behaviour. Second, any analysis based on non-random sub-samples of the program and control groups is non-experimental in nature and hence requires a framework that controls for the non-random selection into employment and unemployment, unobserved heterogeneity and duration dependence. Our econometric analysis provides such a framework.

ECONOMETRIC MODEL

We specify a version of the model presented in Ham and LaLonde (1996). Where y^*_{uidt} and y^*_{eidt} are latent measures of unemployment and employment exit propensities for person *i*, spell duration *d*, and in (relative) time period *t*, we specify the unemployment and employment equations as follows:

$$\begin{aligned} y^*_{uidt} &= \beta_{0u} + \mathrm{X}_{it}\beta_{1u} + \beta_{21u}\mathrm{PRE\text{-}INCENTIVE}_{it} + \beta_{22u}\mathrm{TAKEUP}_{it} \\ &+ \beta_{23u}\mathrm{NON\text{-}TAKEUP}_{it} + \beta_{24u}\mathrm{POST\text{-}INCENTIVE}_{it} + h_u(d,\beta_{3u}) + \theta_{ui} + \varepsilon_{uit} \end{aligned} \tag{1}$$

$$\begin{aligned} y^*_{eidt} &= \beta_{0e} + \mathrm{X}_{it}\beta_{1e} + \beta_{22e}\mathrm{TAKEUP}_{it} + \beta_{23e}\mathrm{NON\text{-}TAKEUP}_{it} \\ &+ \beta_{24e}\mathrm{POST\text{-}INCENTIVE}_{it} + h_e(d,\beta_{3e}) + \theta_{ei} + \varepsilon_{eit} \end{aligned} \tag{2}$$

We observe:

$$y_{uidt} = \begin{cases} 0; \text{remain unemployed if } y^*_{uidt} \leq 0 \\ 1; \text{exit unemployment if } y^*_{uidt} > 0 \end{cases}$$

and

$$y_{eidt} = \begin{cases} 0; \text{remain employed if } y^*_{eidt} \leq 0 \\ 1; \text{exit employment if } y^*_{eidt} > 0 \end{cases}$$

and where X_{it} is a vector of observable covariates, $h(d,\beta_{3j}), j=u,e,$ are general functions that capture duration dependence, θ_{ji} represent unobserved heterogeneity and ε_{jit} are unobserved error terms in the unemployment and employment equations that are assumed to be *i.i.d.* logistic. Since the data are monthly, these specifications can be viewed as discrete hazard models or, equivalently, as panel logit models. We adopt the latter form because time-varying covariates are easier to handle in that model. For example, we include the monthly female unemployment rate to capture the effect of the business cycle. Note that the β's measure the change in the probability of exiting unemployment and employment. Positive values for the β's associated with X_{it} and the program variables will increase the probability of exit and hence decrease the duration of unemployment or employment.

In equations (1) and (2), θ_u and θ_e represent unobserved heterogeneity in individual propensities to exit unemployment and employment, respectively. Conditional on the observables, individuals with relatively large values of θ_u will be more likely to exit unemployment leaving those who remain unemployed with relatively low values of θ_u. Again conditional on the observables, individuals in the latter group will have a lower average probability of exiting unemployment. If we do not account for θ_u, this decline in the exit probability over the duration of the spell will be erroneously identified as negative duration dependence. We control for sample-selection bias by allowing θ_u and θ_e to be correlated; we discuss this in more detail when we specify the likelihood function below.

For our purposes, the key variables in equations (1) and (2) are the binary indicators PRE-INCENTIVE, TAKEUP, NON-TAKEUP, and POST-INCENTIVE. These measure the impact of SSP on unemployment and employment duration, prior to take-up, during the incentive period, and after the incentive period is over. These variables are always zero for control group members.

In the SSP context, the earnings supplement is only available to program group members who found eligible full-time employment in the 12 months following random assignment.[13] We capture this initial program effect with the time-varying variable PRE-INCENTIVE that takes the value 1 for program members in the first 13 months after random assignment and the

[13]We actually define the incentive period to be the first 13 months. Each program group member had exactly 12 months to qualify for the earnings supplement by finding a full time job. The employment variable measured in the surveys, however, was based on calendar months. A participant whose 12-month period started on January 21st, for example, would have until January 21st of the following year to qualify. However, "Month 1" for that person, would be the January in which random assignment occurred. If he or she found a full-time job in the first three weeks of the second January, full-time employment would have been coded as starting in "Month 13."

value 0 afterwards. Depending on whether they had ever received an SSP supplement, the program group members are divided into the take-up and non-take-up program groups. These groups are identified by the variables TAKEUP and NON-TAKEUP. The former is equal to "1" starting in the first month of full-time employment for program group members who found full-time employment in the first 13 months and is zero otherwise. The latter is equal to "1" starting in Month 14 for program group members who did not find full-time employment in the first 13 months and is zero otherwise.

To capture the long-term impact of the program on the take-up program group, we generate a variable, POST-INCENTIVE, that is one for take-up program group members after the three year eligibility period ends and is zero otherwise. POST-INCENTIVE measures the change in the likelihood of leaving unemployment or employment after the benefits period ends. If the coefficient for POST-INCENTIVE is zero, this means that there is no change in behaviour by take-up program group members once they stop receiving benefits. Note that because take-up program group members find their first full-time job in different months (since baseline), POST-INCENTIVE "turns on" at different months as well.

For obvious reasons, we include PRE-INCENTIVE in the initial unemployment hazard. We expect that its coefficient will be positive since program group members have an incentive to find a full-time job in the first 13 months. We also include NON-TAKEUP in this equation. NON-TAKEUP will capture any differences between members of the non-take-up program group and members of the control group in the period after the incentive to take up the program is no longer in effect. The availability of the SSP offer in the first 13 months after random assignment may have induced some individuals who were relatively more "job ready" to find jobs. If so, the members of the program group who were still unemployed after the qualifying period may have been less employable, on average, than the members of the control group who were still unemployed. The coefficient NON-TAKEUP will then be negative, indicating longer spells of unemployment. On the other hand, all members of the program group might have had a greater incentive than control group members to look for work during the qualifying period. This experience may make program group members generally more able than those in the control group to find work after the end of the qualifying period. Under this scenario, the coefficient on NON-TAKEUP will be positive. The overall impact of NON-TAKEUP on unemployment duration is thus an empirical issue.

In the employment equation, the SSP incentive is in place for the take-up program group (at least for the first 36 months after their first month of full-time employment) and irrelevant for the non-take-up program group. We include both TAKEUP and NON-TAKEUP to measure the effect of the incentive on the employment durations (relative to the control group) of these two groups. We expect that the coefficient for TAKEUP will be negative since take-up program group members will have longer employment spells due to the program incentive. The employment behaviour of the non-take-up program group might not be different from the control group workers, particularly those control group members who also did not find work in the 12 months after random assignment. On the other hand, because of the "digging deeper" phenomenon discussed above, workers in the non-take-up program group may be different than the members of the control group who found employment. Theoretically, the overall impact of NON-TAKEUP on employment duration is unknown. We also include the variable POST-INCENTIVE in the employment equation. This captures the long-term impact of SSP since it

measures the differential behaviour of the take-up program group after they stop receiving the SSP benefits. If this coefficient is less than or equal to zero, then (given a significantly negative coefficient for TAKEUP) then this is evidence that SSP has a long-term, positive, impact on the employment durations of the take-up program group. Even if the coefficient for POST-INCENTIVE is positive, there will still be a long-term impact of SSP as long as the sum of the coefficients for POST-INCENTIVE and TAKEUP is significantly less than zero.

We model the full labour market histories of all sample members. These will include strings of both unemployment and employment spells. We believe that it is important to use the full labour-market histories because the goal of SSP is to affect the long-term employment behaviour of the program members. In particular, the incentives are different in the second and higher unemployment spells. Also, since the goal of SSP is to affect the long-term labour-market behaviour of program members, the post-program-eligibility outcome must be identified. Specifically, after the three years of eligibility have expired, it must be known if the labour market behaviour of take-up program group members is different than what it would have been if they had not taken up the program. For the vast majority of these individuals, this can only be measured using multiple spells for unemployment and employment. Note that because of the selection problem, an unbiased estimate of this impact cannot be obtained from a simple and direct comparison of the take-up and control groups.

In the case of unemployment behaviour, the incentives are different for the program group during the first spell versus later spells of unemployment. The program effect for the second and higher unemployment spells is similar to that for employment spells. The program incentive is in effect for the take-up program group and is not in effect for the non-take-up program group. We thus include the variables TAKEUP and NON-TAKEUP in the second and higher unemployment hazards and we expect that the coefficient for TAKEUP will be positive. We also include POST-INCENTIVE to capture the long-term impact of the program incentive on the unemployment behaviour of the take-up program group. If this coefficient is greater than or equal to zero then (given a significantly positive coefficient for TAKEUP) this is evidence that SSP has a long-term negative impact on the unemployment durations of the take-up program group. Even if the coefficient for POST-INCENTIVE is negative, there will still be a long-term impact of SSP as long as the sum of the coefficients for POST-INCENTIVE and TAKEUP is significantly greater than zero.

Initially, we assume that the coefficients in the unemployment and employment hazards for second and higher spells are different from those in the initial coefficients for the non-program variables. For example, given the different incentive structure for the program group, the unemployment hazard for the first spell might well be different by more than just the direct program effect. When we estimate the unemployment and employment equations for all spells, we test for the equality of the coefficients for the non-program variables.

To enable the construction of the likelihood function, we divide the labour market histories into four scenarios:

1. Unemployed for the entire T=52 months.
2. Unemployed for t_{u1} periods and employed for the remaining t_{e1}= T – t_{u1} periods.
3. Unemployed for n>1 spells with a last, censored, unemployment spell.
4. Unemployed for n>1 spells with a last, censored, employment spell. Recall that the ε_{jit}, $j=u,e$ are specified as *i.i.d.* logistic. Then 1-λ (the exit rate) is specified as:

$$1-\lambda_{jidt}\left(t,d \mid \theta_{ji}\right)=\left(1+\exp\left(-y_{jidt}\right)\right)^{-1},\ \ j=u,e \tag{3}$$

The contributions to the likelihood function are:

1. $$\mathrm{L(T)}=\int\prod_{k=1}^{T}\lambda_u\left(t,k \mid \theta_u\right)\cdot d\mathrm{G}_u(\theta_u) \tag{4}$$

where $G_u(\theta_u)$ is the marginal distribution for θ_u (discussed below).

2. $$\mathrm{L}\left(d_{u1},d_{e1}\right)=\int\left(\left(1-\lambda_u(d_{u1} \mid \theta_{\mathrm{u}})\right)\cdot\prod_{k=1}^{d_{u1}-1}\lambda_u\left(k \mid \theta_u\right)\right)\cdot\left(\prod_{m=1}^{d_{e1}}\lambda_{\mathrm{e}}\left(m \mid \theta_e\right)\right)\cdot d\mathrm{G}(\theta_u,\theta_e) \tag{5}$$

where $G(\theta_u,\theta_e)$ is the joint distribution of θ_u and θ_e (discussed below).

3.
$$\begin{aligned}\mathrm{L}\left(d_{u1},\ldots,d_{un},d_{e1},\ldots d_{e,n\text{-}1}\right)=\int&\left(\prod_{i=1}^{n-1}\left(\left(1-\lambda_u(d_{ui} \mid \theta_u)\right)\cdot\prod_{k=1}^{d_{ui}-1}\lambda_u\left(k \mid \theta_u\right)\right)\cdot\prod_{k=1}^{t_{un}}\lambda_u\left(k \mid \theta_u\right)\right)\\&\cdot\left(\prod_{j=1}^{n\text{-}1}\left(\left(1-\lambda_e(d_{ej} \mid \theta_e)\right)\cdot\prod_{m=1}^{d_{ej}-1}\lambda_e\left(m \mid \theta_e\right)\right)\right)\cdot d\mathrm{G}(\theta_u,\theta_e)\end{aligned} \tag{6}$$

4.
$$\begin{aligned}\mathrm{L}\left(d_{u1},\ldots,d_{un},d_{e1},\ldots d_{en}\right)=&\int\left(\prod_{i=1}^{n}\left(\left(1-\lambda_u(d_{ui} \mid \theta_u)\right)\cdot\prod_{k=1}^{d_{ui}-1}\lambda_u\left(k \mid \theta_u\right)\right)\right)\\&\cdot\left(\prod_{j=1}^{n\text{-}1}\left(\left(1-\lambda_e(d_{ej} \mid \theta_e)\right)\cdot\prod_{m=1}^{d_{ej}-1}\lambda_e\left(m \mid \theta_e\right)\right)\cdot\prod_{k=1}^{d_{en}}\lambda_e\left(k \mid \theta_e\right)\right)\cdot d\mathrm{G}(\theta_u,\theta_e)\end{aligned} \tag{7}$$

where d_{ji}, $j=u,e$ are the duration lengths for spell i. Note that we have suppressed the i and t subscripts (as well as X_{it}) to ease notation.

Following Ham and LaLonde (1996) we model the unobserved heterogeneity terms as:

$$\theta_u = \alpha_{1u} + \alpha_{2u}\theta^* \quad \text{and} \quad \theta_e = \alpha_{1e} + \alpha_{2e}\theta^* \tag{8}$$

where θ^* is drawn from a two point distribution. Normalizing α_{1u}=0 and α_{2u}=1 means that we only need to estimate α_{1e} and α_{2e}. The main reason that we allow for only two points of support is due to the computational complexity of the model.[14] While the fact that we only allow for two types of unobserved heterogeneity might seem limiting, increasing the number of points of support usually has little impact on the results.[15]

The specification in equation (8) allows for the unobserved heterogeneity terms to be correlated across the unemployment and employment durations. This controls for non-random selection into and out of work and hence eliminates what Eberwein, Ham, and LaLonde (1997) call "dynamic selection bias." In their analysis, Eberwein, Ham, and LaLonde (1997) use a slightly modified version of the model in Ham and LaLonde (1996) to estimate the impact of classroom training on the duration of unemployment and employment using data from the Connecticut *National Job Training Partnership Act* study (JTPA-CT). They find significant effects on unemployment duration but not on employment duration. They also find no evidence of dynamic selection bias. This is in contrast to Ham and LaLonde (1996), where there is significant evidence of dynamic selection bias using data from the National Supported Work – Work Experience Study (WE). Eberwein, Ham, and LaLonde (1997) conclude that the nature of the program will dictate whether dynamic selection bias will exist. In particular, they note that programs like WE, where the experimental group is removed from the labour force to receive training, are more likely to exhibit dynamic selection bias than programs like JTPA-CT, where this is not the case. This is because when comparing new unemployment spells for the experimental and control groups, the former group includes all its members who have just finished training while the latter group is a sub-sample of members who have found a job and then returned to the unemployment state. Thus we are comparing the stock of experimental group members with the flow of control group members. These samples are unlikely to be similar and hence dynamic selection bias will be an issue. On the other hand, when the experimental group is not removed from the labour force, the comparison of new unemployment spells will be based on flows of both experimental and control group members that are likely to be relatively similar in their composition and hence dynamic selection bias will not be so important. SSP is like JTPA-CT in that the experimental group is not removed from the labour force. Hence we do not expect that dynamic selection bias will be much of a factor when we estimate our model. Still, it is important to follow through with joint estimation the unemployment and employments equations since the Eberwein, Ham, and LaLonde (1997) conclusion is only based on two programs.

[14]Ondrich and Rhody (1999) provide a method for organizing the data that simplifies the maximum likelihood estimation of the multiple-spell model with unobserved heterogeneity.

[15]For example, we estimate the unemployment and employment equations individually using the *Stata* command "xtlogit," with the minimum number of points of support (4). The results do not change when we increase this number.

EMPIRICAL RESULTS – EQUATIONS ESTIMATED INDIVIDUALLY

In this section, we present the empirical results for the econometric model presented above. We first estimate the unemployment and employment equations individually. We report on the joint estimates in the next sub-section. We do this in two stages since we can use the panel logit estimator for the individual equations and it is then much easier to include the unemployment rate in the model (since it varies each month). The data are monthly observations on the unemployment and employment status (working full time or not) of the 2,255 program group members and 2,165 control group members. This includes up to six unemployment and six employment spells for each individual. We use information for 52 months so there are a total of 229,840 observations.

We include the following covariates, measured in the baseline survey, in the unemployment and employment equations: dummy variables that indicate if the individual was employed in at least one of the 12 months prior to baseline (EEMPFTBR), was single and not previously married (NVRMAR), did not have at least a high school degree (BLTHS), lived in British Columbia (IDBC), had limited activity due to a long term emotional, psychological, nervous, or mental health condition or problem (EMOPROB), had a child at baseline who was less than or equal to four years old (YGCHLE4), if age was between 23 and 25, 26 and 29, 30 and 34, 35 and 39, or 40 and older (BAGE2325, …), the number of children (NKIDS), the number of months on IA in the three years prior to the baseline survey (MONTHIA), and the average monthly IA amount received in four quarters prior to baseline survey (BENEFIT). And, as noted above, we also include the monthly unemployment rate (UNRATE).

We allow for duration dependence by including the log of time, its square, and its cube. Higher order terms were not significant. This is a restriction on including a dummy variable for each period (which is essentially what happens in the Cox proportional hazard model) but the restriction greatly reduces the number of parameters in the model. We estimate the model with one common baseline hazard plus program effect dummy variables. Thus, the program variables (PRE-INCENTIVE, TAKEUP, and POST-INCENTIVE) pick up the "average" difference in the baseline hazards between the program and control groups.[16]

One issue is whether or not to update the covariates at the beginning of each employment spell and at the beginning of the second and higher unemployment spells. We update age at the beginning of the spell. It might be useful to update variables such as the number of children in the household, whether or not the person had a high school degree, and the person's marital status but two complications arise. First, at the 18-, 36-, and 54-month interviews, we know only whether there has been a change since the last interview. Thus, at best, we can only update at these points and not necessarily at the month the change occurred or even at the beginning of each spell. Second, it may well be that fertility, educational attainment, and marital status are not only endogenous but are affected by program uptake. In order to obtain the full program effect, these impacts would need to be taken into account. By not updating the controls, any

[16]We could have calculated the program effect by estimating separate hazards for the program and control group, calculating the difference in the baseline hazard at fixed values for the covariates for each month, and then taking the appropriate weighted average but this seems unnecessarily complicated.

impact of SSP on these variables would end up in the overall program effect as measured by the coefficients on the program variables, as is appropriate for our analysis.

The empirical strategy is to estimate the unemployment and employment hazards with unobserved heterogeneity — first individually and then jointly in the next section. In this section, we begin by estimating separate hazards for the initial spell and for the second and higher spells of unemployment and employment. This allows us to see if the behaviour is different for these higher spells. We then combine the data and estimate unemployment and employment hazards using all spells. To estimate the individual unemployment and employment equations, we use the panel logit model with unobserved heterogeneity and estimate using Gauss-Hermite quadrature ("xtlogit" in *Stata*).

Given that the model is panel logit, the coefficient estimates must be transformed to have the useful interpretation of the change in the probability of exiting unemployment (employment) for a unit change in the variable in question. In the case of the binary program variables, we take the difference in the estimated probability of exiting unemployment (employment) with the variable set at 1 and 0.[17] Results are given as the mean of the difference over the sample observations.

The estimation results for the unemployment and employment equations are given in Tables 13 and 14. The results when the initial spell and the second and higher spells are estimated separately are given in Columns 1 to 4 of each table. The results when all spells are estimated jointly are given in Columns 5 and 6. We test to see if the variables other than the program dummies are equal across the two equations. Given the structure of the experiment, it is not surprising that we find that unemployment behaviour is different in the second and higher unemployment spells (the equality of covariate coefficients is rejected at the 5 per cent significance level) but that there is no difference in behaviour for the different employment spells (the equality of covariate coefficients is not rejected at the 5 per cent significance level). Thus, in the case of unemployment duration, we allow for separate effects for the covariates in the second and higher spells when all spells are estimated jointly (not shown). Finally, note that for unemployment, the variable NON-TAKEUP is included in both the initial spell and second and higher spell equations. We test and do not reject the hypothesis that the coefficients for NON-TAKEUP are the same across these two equations (5 per cent significance level). Given the outcomes of these tests, we focus on the results for all spells estimated jointly (Columns 5 and 6 of Tables 13 and 14).

For unemployment, during the pre-incentive period a program group member had a 0.022 higher probability of exiting unemployment in any given month than the control group (holding all observable characteristics constant). This impact is highly significant. It represents a 134 per cent increase in the probability of finding employment (0.039 vs. 0.017 — not shown). This result is comparable to the higher estimated hazard rate for the program group in Figure 1. On the other hand, the exit rate from unemployment for the non-take-up program group is significantly lower than the control group (as seen by the statistically negative and significant coefficient on NON-TAKEUP) though the derivative is relatively small (0.004).

[17]Note that for duration, even though we include the log of duration, its square, and its cube, only one derivative is reported that is the change in the probability of exiting from a one-month increase in duration.

For program group members in the pre-incentive period, the probability of exiting their first unemployment spell is 0.022 higher than for control group members (Table 13, Column 6, line 1). For take-up program group members (program group members who found a job in the pre-incentive period) the probability of exiting their second or higher spell of unemployment is 0.008 higher than for all other participants (Table 13, Column 6, line 2).

Still, this represents a 47 per cent increase in the probability of exiting unemployment (0.025 vs. 0.017 — not shown). The negative and significant coefficient for POST-INCENTIVE indicates that once the incentive program ends, the take-up program group members are more likely to stay unemployed relative to the incentive period.

Table 13: Estimation Results for Unemployment Model

	1st Spell		2nd + Spells		All Spells			
Variable	**Coefficient (1)**	**Derivative (2)**	**Coefficient (3)**	**Derivative (4)**	**Coefficient (5)**	**Derivative (6)**	**Coefficient (7)**	**Derivative (8)**
PRE-INCENTIVE	0.873**	0.017			0.875**	0.022	0.875**	0.022
	(0.06)				(0.06)		(0.06)	
TAKEUP			0.468**	0.020	0.393**	0.008	0.394**	0.008
			(0.08)		(0.06)		(0.06)	
NON-TAKEUP	-0.310**	-0.003	-0.287**	-0.009	-0.293**	-0.004	-0.293**	-0.004
	(0.06)		(0.10)		(0.05)		(0.05)	
POST-INCENTIVE			-0.838**	-0.020	-0.716**	-0.008		
			(0.14)		(0.13)			
POST1-4							-0.510**	-0.007
							(0.17)	
POST5-8							-0.993**	-0.010
							(0.26)	
POST9-12							-0.683*	-0.008
							(0.29)	
UNEMPL RATE	0.041	0.0009	-0.062*		0.041		0.041	
	(0.02)		(0.03)		(0.02)		(0.02)	
LNDURATION	-0.890**	0.0002	-0.437*		-0.885**		-0.885**	
	(0.19)		(0.21)		(0.19)		(0.19)	
$LNDURATION^2$	0.611**		0.298		0.607**		0.608**	
	(0.11)		(0.15)		(0.11)		(0.11)	
$LNDURATION^3$	-0.106**		-0.074*		-0.105**		-0.105**	
	(0.02)		(0.03)		(0.02)		(0.02)	
ρ	0.000		0.091		0.000		0.000	
p-value for ρ=0	1.000		0.000		1.000		1.000	
Observations	155,315		31,530		186,845		186,845	
Individuals	4,420		1,658		4,420		4,420	

Notes: Standard errors are in parentheses.

* Indicates significant at five per cent; ** indicates significant at one per cent.

Table 14: Estimation Results for Employment Model

	1st Spell		2nd + Spells		All Spells			
Variable	**Coefficient (1)**	**Derivative (2)**	**Coefficient (3)**	**Derivative (4)**	**Coefficient (5)**	**Derivative (6)**	**Coefficient (7)**	**Derivative (8)**
TAKEUP	-0.375**	-0.018	-0.074	-0.002	-0.348**	-0.017	-0.348**	-0.017
	(0.060)		(0.089)		(0.060)		(0.060)	
NON-TAKEUP	0.105	0.006	-0.012	0.0004	0.105	0.006	0.105	0.006
	(0.068)		(0.121)		(0.069)		(0.069)	
POST-INCENTIVE	0.188	0.012	-0.204	-0.006	-0.038	-0.002		
	(0.227)		(0.138)		(0.113)			
POST1-4							0.126	0.008
							(0.141)	
POST5-8							-0.107	-0.006
							(0.193)	
POST9-12							-0.432	-0.021
							(0.302)	
UNEMPL RATE	0.027		0.033		0.026		0.026	
	(0.024)		(0.036)		(0.021)		(0.021)	
LNDURATION	2.975**		3.570**		3.154**		3.171**	
	(0.286)		(0.472)		(0.239)		(0.240)	
$LNDURATION^2$	-1.394**		-1.744**		-1.447**		-1.460**	
	(0.167)		(0.289)		(0.139)		(0.140)	
$LNDURATION^3$	0.172**		0.240**		0.187**		0.189**	
	(0.029)		(0.054)		(0.024)		(0.024)	
ρ	0.000		0.000		0.078		0.078	
p-value for ρ=0	1.00		1.00		0.000		0.000	
Observations	30,348		12,647		42,995		42,995	
Individuals	2,216		933		2,216		2,216	

Notes: Standard errors are in parentheses.

* Indicates significant at five per cent; ** indicates significant at one per cent.

The overall impact in the post-incentive period — the sum of the coefficients for TAKEUP and POST-INCENTIVE — is also negative (-0.323) and significant. This indicates that the take-up program group members are more likely to remain unemployed relative to the control group. This contrasts with the greater likelihood of exiting unemployment while the earnings supplement is available. We have data for about one year of the post-incentive period for a large enough number of take-up program group members, so we can investigate whether the average impact masks a trend in the unemployment behaviour of the take-up program group. We estimate a model where we include dummy variables for the first four months (POST1-4), months 5 to 8 (POST5-8) and months 9 to12 (POST9-12) of the post-incentive period. The results are given in Columns 7 and 8 of Table 13. The estimates are fairly similar (and not significantly different from each other) so there does not appear to be a declining trend in the effect of SSP. This negative effect might be picking up the fact that the take-up program group

includes a set of women who would not have found a job without the SSP incentive and so once the incentive is over, they are more likely to remain unemployed (relative to the control group as a whole).

Many of the other covariates in the first unemployment spell are significant at the one per cent level (not shown but available on request). Participants who reported young children, less education, and emotional problems at the time of the baseline survey are less likely to exit unemployment (holding constant the program indicators). The results also imply a probability of gaining full-time employment that declines with each age category, as indicated by the declining coefficients on the age variables. Greater income assistance benefits and longer spells on welfare at the baseline also decrease the likelihood of exiting unemployment. Those who have been employed full time in at least one of the 12 months prior to the baseline have a greater probability of exiting unemployment than those who were not employed in the year before random assignment.

There is evidence of duration dependence since the log of duration, its square, and its cube are jointly significant at the one per cent level. The shape of the duration dependence generally follows the estimate of the hazard rate for the control group in Figure 1; after a decline in the exit rate in the first few months, it is basically flat for the remainder of the period. The unemployment rate is not significant at the five per cent level. It appears that different economic conditions (as represented by the unemployment rate) did not affect the duration of the first spell of unemployment.

These other covariates generally have much less important impacts on the second and higher spells of unemployment. There is significant evidence of negative duration dependence in the second and higher spells. Also, the unemployment rate has a negative impact on exiting unemployment. While British Columbia has a lower unemployment rate than New Brunswick, a sample member in British Columbia is less likely to exit unemployment than one in New Brunswick.

Information on unobserved heterogeneity is given by ρ, the proportion of the total error variance that is attributable to individual unobserved heterogeneity. We provide an estimate of ρ and the p-value for the likelihood ratio test that ρ is zero in Table 13. There is no evidence of unobserved heterogeneity in the first unemployment spell (in fact, $\hat{\rho} = 0.000$) but there is significant unobserved heterogeneity for the second and higher spells. But even in the latter case, ρ is only 0.091. There is no evidence of unobserved heterogeneity in the combined model but this is not surprising given that the bulk of the observations are from the first spell.

The results for the employment duration equation are given in Table 14. Other than the program effects indicators, the variables are the same as those that appear in the model for unemployment. The main result is that (holding all observable characteristics constant) a take-up program group member had a 0.017 lower probability of ending their employment spell in a given month, in any period, than a control group member. This impact is highly significant. It represents a 28 per cent decrease in the probability of exiting employment relative to the control group (0.045 vs. 0.063 — not shown). Non-take-up program group members were not significantly more or less likely than control group members to end their employment spell.

Another important result is that there is no change in the employment behaviour of the take-up program group once the incentive period ends (the coefficient on POST-INCENTIVE is not significant). This is evidence that SSP had a positive long-term impact on employment duration for the take-up program group. Again, we investigate the possibility of a trend over the post-incentive period by including POST1-4, POST5-8, and POST9-12 in the employment model. This shows a positive impact during the first fourth months of the post-incentive period and then an increasingly negative impact over the next eight months (Columns 7 and 8 of Table 14). While these estimates are not significant individually or jointly, it does provide some weak evidence that the positive long-term impact of SSP on employment duration is increasing over time.

In contrast to the estimates of the unemployment model, few of the other variables are significant. There is a monotonically decreasing probability of exiting full-time employment as participants age, and those without a high school degree are more likely to leave full-time employment. There is evidence of positive duration dependence for the first seven months and then negative duration dependence thereafter. This is similar to what is seen in the Kaplan–Meier hazard rate estimates in Figure 4. The unemployment rate has a positive effect on the exit from full-time employment but it is not significant.

Interestingly, there is no evidence of unobserved heterogeneity when the first and second and higher employment spells are estimated separately but there is significant unobserved heterogeneity in the joint model. It appears that the combination of information about all employment spells enables the identification of the unobserved heterogeneity. However, it only comprises 8.5 per cent of the total error variance. When the model is estimated without unobserved heterogeneity, the main difference is that the coefficient on TAKEUP drops (in magnitude) from -0.348 to -0.279.

EMPIRICAL RESULTS – EQUATIONS ESTIMATED JOINTLY

In this sub-section, we analyze the results from estimating the unemployment and employment equations jointly. In particular, we assume that the unobserved heterogeneity terms θ_u and θ_e are drawn from a two point distribution as given in equation (8). Given that we find no strong evidence that the unemployment rate is significant in either the unemployment or employment duration equation, we exclude it here. This simplifies the estimation procedure since we no longer have a covariate that varies every month. The results are presented in Table 15.

As expected, given the general lack of evidence of unobserved heterogeneity in the models estimated thus far, they do not differ dramatically from the single equation results.[18] α_{1e} and α_{2e} are the parameters that determine θ_e in equation (8), μ_1 and μ_2 are the two points of support, and $e^{\gamma}/(1+e^{\gamma})$ is the probability that μ_1 will occur (if $\gamma = 0$ then the probability is 0.5).

[18]We do not include the unemployment rate in this model because of the difficulty in dealing with time-varying covariates but we note that it is not significant in the previous results so this should have little impact.

While μ_1 and μ_2 are highly significant, they are not significantly different from each other so there is no strong evidence of two-point heterogeneity in this model.

Table 15: Estimation Results for Unemployment and Employment Model — Correlated Heterogeneity

Variable	Unemployment (Coefficient)	Employment (Coefficient)
PRE-INCENTIVE	0.8837**	
	(0.06)	
TAKEUP	0.4063**	-0.3099**
	(0.06)	(0.06)
NON-TAKEUP	-0.288**	0.118
	(0.05)	(0.07)
POST-INCENTIVE	-0.7455**	-0.040
	(0.14)	(0.11)
LNDURATION	-0.9241**	3.1567**
	(0.19)	(0.25)
$LNDURATION^2$	0.625**	-1.4822**
	(0.11)	(0.14)
$LNDURATION^3$	-0.158	1.9166**
	(0.08)	(0.25)
α_{1e}	0.4522	
	(3.10)	
α_{2e}	2.0706	
	(1.34)	
μ_1	-2.0943**	
	(0.22)	
μ_2	-2.4439**	
	(0.19)	
γ	-0.3939	
	(1.10)	
Observations	229,840	
Individuals	4,420	

Notes: α_1 and α_2 are the parameters that determine θ_e in equation (8).
μ_1 and μ_2 are the two points of support and exp(-γ)/(1+exp(-γ)) is the probability that μ_1 will occur.
* Indicates significant at five per cent; ** indicates significant at one per cent.

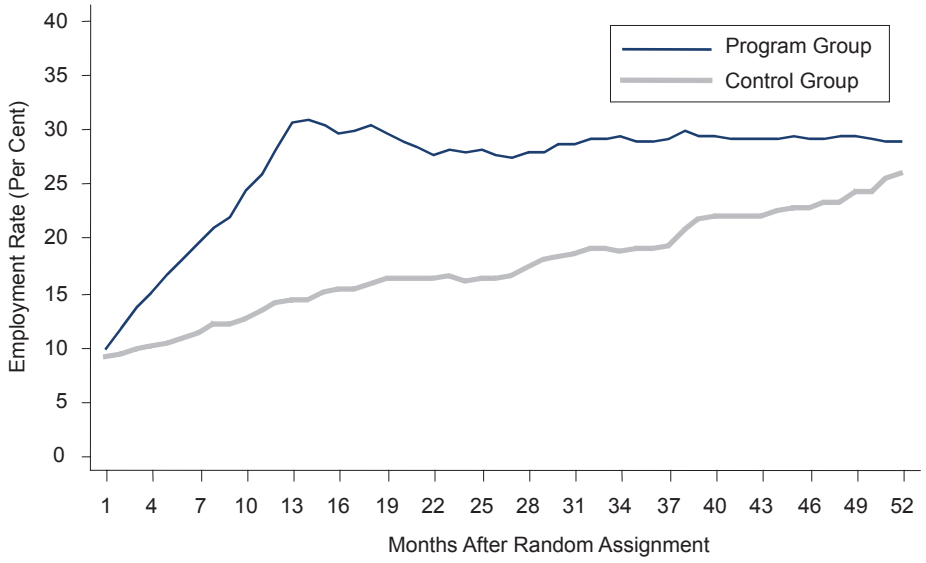
These results provide little evidence of dynamic selection bias in the model. This is consistent with our expectations given the previous discussion of Eberwein, Ham, and LaLonde (1997). That is, the fact that SSP does not result in the program group being removed from the labour force means that the control and program groups will not be substantially different across unemployment and employment spells.

Simulation

Using the econometric model estimated in the fourth section, we can now obtain an unbiased comparison of the employment rates for the take-up and control groups. Note again that this comparison is not the same as the standard experimental comparison of the entire program group — all those randomly assigned to the group that is offered the Self-Sufficiency Project (SSP) earnings supplement — to the entire control group. The take-up program group is the self-selected subset of the entire program group that found eligible full-time work and received the SSP supplement.

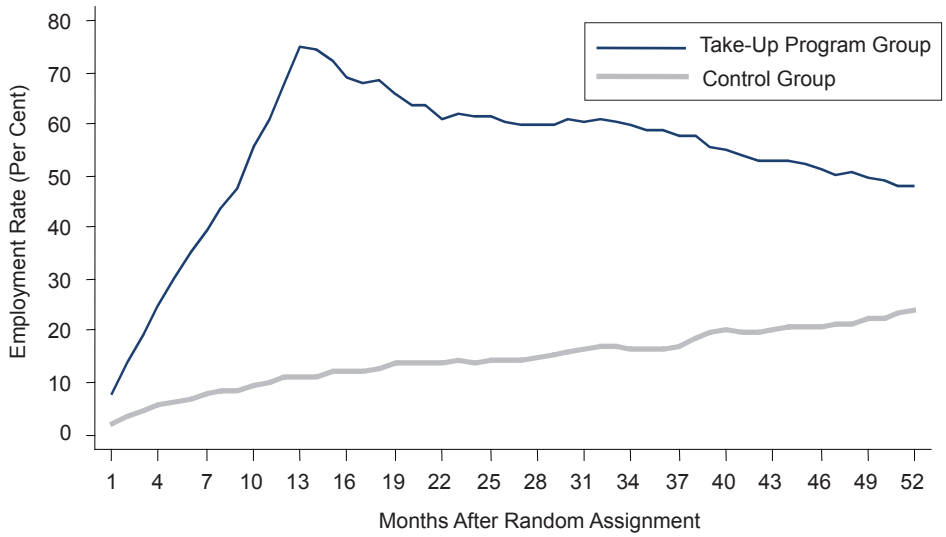
Before using our econometric model to simulate the difference in full-time employment rates between the take-up program group and the control group, we compare the full-time employment rates for the entire program and control groups over the 52-month period (Figure 10). In this figure (and in previously published evaluations), we see that SSP had a significant short-run impact that peaked 12 to 13 months after random assignment. After 52 months, however, there was little difference in full-time employment rates between the two groups; we calculate a difference, after 52 months, of 1.6 percentage points. The gradual diminution of the program–control gap is sometimes referred to as "control group catch-up" since, over the 52 months, the program group employment rates remained roughly constant, at the 12- to13-month level, as the control group rates rose.

Figure 10: Monthly Full-Time Employment Rate



In Figure 11, we show the employment rates for the self-selected take-up program group and the entire control group, calculated from the unadjusted survey data. This difference measures the so-called "effect of the treatment on the treated" without any econometric adjustment for heterogeneity. While this "raw" non-experimental effect declines over time, it remains at 22.5 percentage points in Month 52.

Figure 11: Monthly Full-Time Employment Rate — Take-Up and Control Groups



To determine the extent of the selection bias due to unobserved heterogeneity in Figure 11, we carry out a simulation based on the econometric results presented in the fourth section. This allows us to illustrate the difference in the magnitude of the "effect of the treatment on the treated" when heterogeneity is and is not taken into account. The simulation assumes a particular set of observed characteristics. We simulate results for an SSP participant who:

- is from British Columbia at a time when the unemployment rate was 8.8 per cent;
- has two children (at least one being four years old or younger);
- has less than high school education;
- has 29 months on income assistance (IA) in the three years prior to baseline;
- has a monthly IA benefit at baseline of $1,000;
- was previously married;
- has reported no long-term emotional problems; and
- was not employed and working full time in any of the 12 months preceding the baseline survey.

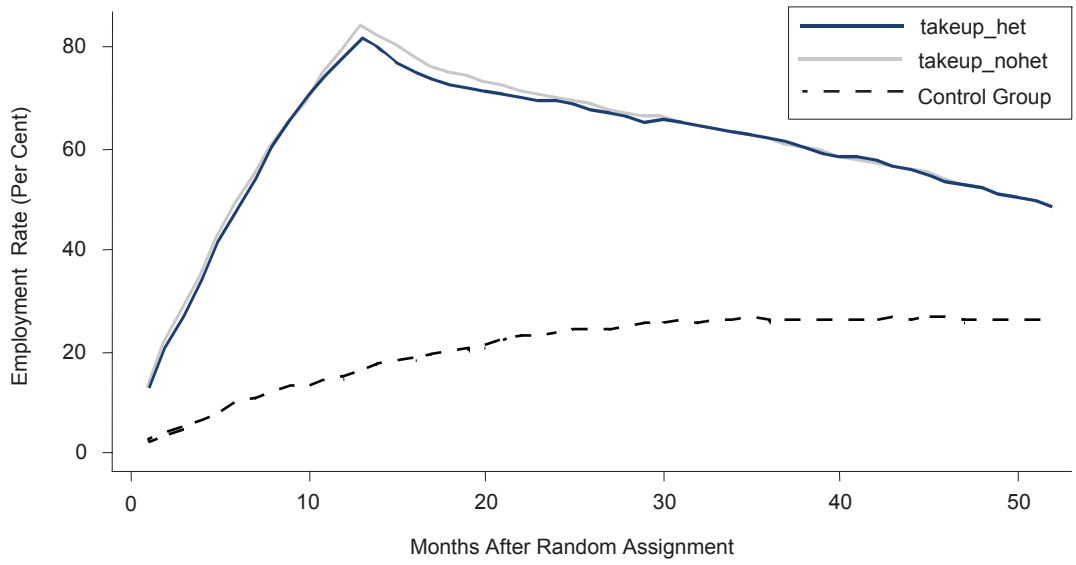
We also adjust for duration dependence depending on the number of months in a given state. We then combine these "observed" characteristics with the parameter estimates of the models presented in the fourth section. The idea of the simulation is to create simulated employment histories for a large number of individuals, all of whom have the above observed characteristics. These simulated individuals will differ from each other only because of a randomly generated error term, drawn from the logit distribution, for each of 52 months. This error term (and the values of the variables and coefficients) determines if the simulated individual is unemployed or employed in any given month.[19] Simulated individuals in the program group differ from those in the control group only in that the program group variables (e.g. PRE-INCENTIVE and POST-INCENTIVE) and their coefficients are added into the equation. To be in the simulated take-up program group, a simulated program group

[19]We adjust the constants in the unemployment and employment equations to give probabilities that are similar to the raw data.

member must "find a job" (i.e. be randomly given values of the logit error term that lead to a prediction of full-time employment) in the first 13 months. Because of this, approximately half of the simulated individuals in the program group are not used in our comparisons of the employment rates of the take-up program group and the control group (since they are simulated members of the non-take-up program group). Arbitrarily, we carry out this exercise for 10,000 program group members (in order to have about 5,000 take-up program group members) and 5,000 control group members.

Using the simulation, we first generate results for the control group and for the take-up program group with two-point heterogeneity, implying that there are two types of individuals in the data, holding constant all of the observed covariates. That is, in addition to the standard covariates listed above, and in addition to the randomly generated logit errors, each of the 15,000 simulated individuals is given one of the two values of θ_u and θ_e estimated in the fourth section. With the heterogeneity added in, this scenario is similar to looking at the unadjusted data in the sense that we believe that the take-up program group is a non-random subset of the entire program group (i.e. that unobserved heterogeneity exists). In particular, the existence of unobserved heterogeneity will result in a take-up program group that is more likely to exit unemployment because, for example, those with high values of θ_u will be more likely to leave unemployment. The results for the control group and the take-up program group with heterogeneity (the line labelled "takeup_het") are given in Figure 12. They replicate quite well the patterns observed in the unadjusted data in Figure 11. The treatment effect at 52 months is 22.1 percentage points.

Figure 12: Simulation Results — Take-Up Program Group (With and Without Heterogeneity) and Control Group



Next, we simulate the employment rate for the take-up program group without heterogeneity. Instead of the two points of support (θ_u and θ_e), we add in the expected value of these two heterogeneity terms to each observation. In this case, the selection of the take-up program group is random (it depends only on the random logit error) so the comparison with the control group is unbiased. The results (takeup_nohet) are given in Figure 12. Note that

the employment gap (treatment effect) is virtually identical to the case with heterogeneity; the lines labelled takeup_het and takeup_nohet are virtually the same. This implies that the selection into the take-up program group caused by unobserved heterogeneity does not appear to bias the estimates of the comparison of the employment probabilities. The primary lesson to be learned from our simulation (and, implicitly, from our econometric model) is that a simple unadjusted comparison of the employment rates of the self-selected take-up program group and the entire control group gives an estimate of the "effect of the treatment on the treated" that is not greatly biased by unobserved heterogeneity.

In our model, there are two "channels" through which heterogeneity might have affected the simple comparison. First, we might have found evidence of substantial unobserved heterogeneity. In fact, however, we found only weak evidence of such heterogeneity. We consistently find that unobserved heterogeneity accounts for less than 10 per cent of the error variance. So one might accept the simple unadjusted comparison in Figure 11 on the assumption that the take-up program group does not differ in unobserved ways from the rest of the program group.

The simulation, however, also includes the limited heterogeneity that we did in fact observe. The effect of that heterogeneity, small to begin with, is further muted by the positive correlation between the unobserved heterogeneity (between θ_u and θ_e) across the unemployment and employment equations that was mentioned at the end of the fourth section. That is, assuming that there is some unobserved heterogeneity, the *form* of the heterogeneity is such that those who are most likely to leave unemployment are also more likely to leave employment (Ham and LaLonde (1996) report the same result). Thus, an effect of heterogeneity that would reduce the gap between the take-up and control groups is counter-balanced by an effect that would increase that gap. This was somewhat unexpected since one might have thought that those most likely to leave unemployment would be least likely to leave employment. If that had been true the line labelled "takeup_nohet" would lie beneath that labelled "takeup_het" and would indicate the presence of selection bias. One explanation for the positive correlation is that those individuals who are more likely to leave unemployment would be more likely to leave a job they did not like given their ability to find a new job, whereas individuals who are less likely to find work would be less willing to give up a bad job.

Overall, the simulation provides additional evidence of a true long-term impact of SSP on those who received the supplement. Of course, the period for which we have employment histories that post-date the incentive period is relatively short. To see a true long-term effect it would be useful to see the results after another 18 months.

Conclusion

We have conducted an econometric analysis of the impact of the Self-Sufficiency Project (SSP) on unemployment and employment durations. We have gone beyond the simple experimental comparisons of the labour force behaviour of the entire program and control groups by controlling for the non-random selection into work and second and higher spells of unemployment. Assuming our model is correctly specified, this allows us to obtain more accurate and detailed estimates of the impact of SSP. We have data for 52 months for our full set of sample members and this allows us to determine the short-run and long-run impact of SSP, after the eligibility period ended. Our econometric framework is based on the work of Ham and LaLonde (1996). In one sense, our analysis is simpler than theirs since SSP is a randomized experiment and because there is no training period with which to contend. SSP, however, involved a more complicated incentive structure to model and analyze. SSP is more comparable to the experimental data analyzed by Eberwein, Ham, and LaLonde (1997) since classroom training occurs while individuals are working.

In terms of unemployment behaviour, we find that SSP had a strong pre-incentive impact, presumably because program group members were required to find a full-time job in the first 12 months after random assignment in order to receive the supplement. For the program group members who actually qualified — the take-up program group — there is a significant and positive impact of receiving the supplement on the probability of exiting second and higher spells of unemployment. But once the 36-month period of supplement eligibility ends, there is a negative and significant impact on the probability of exiting unemployment. We estimate that the monthly probability of exit from unemployment in a given month decreases by 0.031 after eligibility has ended. This represents a 79 per cent decrease in the monthly probability of exit for the take-up program group relative to the eligibility period. Relative to the control group, there is a 50 per cent lower probability of exit from unemployment. This indicates that, on average, there is a negative impact of SSP on the take-up program group once they stop receiving the supplements.

There is a negative and significant impact on exiting employment from receiving the SSP supplement on the take-up program group. On average, there is no change in this impact on employment duration once the supplement payments end. Thus, SSP appears to have a long-term impact on the employment behaviour of the take-up program group.

Our ability to control for the non-experimental aspects of the data derives from the specification of unobserved individual heterogeneity. In essence, this allows us to control for unobservable individual factors that affect their employment decisions, which if not accounted for, can bias the estimates of the model parameters. This works in two ways. First, when we estimate the unemployment (employment) equation using multiple spells, we control for the fact that the second spell is conditional on there being a first spell through the common unobserved heterogeneity term, θ_{ui} (θ_{ei}). That is, we do not treat the multiple spells as if they are actually single spells for different people but as multiple spells for the same person since each includes the same value for θ_{ui} (θ_{ei}). Second, we estimate the

unemployment and employment equations jointly by allowing the unobserved heterogeneity terms θ_{ui} and θ_{ei} to be correlated.

There are numerous ways for specifying unobserved heterogeneity. We could follow Meyer (1990) and use the gamma distribution. Instead, we use the "xtlogit" command in *Stata* when estimating the unemployment and employment equations individually. This procedure employs the method of Gauss-Hermite quadrature for estimating the unobserved heterogeneity (that is, the unobserved heterogeneity is assumed to be normally distributed). When estimating the joint model, we follow Ham and LaLonde (1996) and use a fully nonparametric two-point one-factor distribution. This latter specification is justified by the complexity of the estimation problem. Generally, we do not find much evidence of the existence of unobserved heterogeneity in either the unemployment or employment equations. And when it does exist, in the full model of employment and in the unemployment model for second and higher spells, the proportion of the variance attributable to unobserved heterogeneity is less than 0.10.

We use the econometric results to simulate the effect on simple estimates of the difference in employment rates for the take-up and control groups. The simulation illustrates the important point that, in SSP, unobserved heterogeneity does not seem to create bias in the simple estimates of the "effect of the treatment on the treated" generated by a comparison of the employment and unemployment rates of the self-selected take-up program group to the entire control group.

There is one large benefit of the kind of non-experimental econometric modeling undertaken here — we can answer questions about the effect of SSP on the labour force behaviour of the non-random subsets of the program and control groups who worked full time after random assignment. But there is a perhaps equally large caveat — as with most econometric analyses, the validity of our answers depends crucially on the correctness of the specification of the model.

References

Eberwein, C., Ham, J. C., & LaLonde R. J. (1997). The impact of being offered and receiving classroom training on the employment histories of disadvantaged women: Evidence from experimental data. *Review of Economic Studies*, *64*, 655–682.

Ham, J. C., & LaLonde, R. J. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica*, *64*, 175–205.

Klein, J. P., & Moeschberger, M. L. (1997). *Survival analysis: Techniques for censored and truncated data.* New York: Springer.

Lin, W., Robins, P. K., Card, D., Harknett, K., & Lui-Gurr, S. (1998). *When financial incentives encourage work: Complete 18-month findings from the Self-Sufficiency Project.* Ottawa: Social Research and Demonstration Corporation.

Meyer, B. D. (1990). Unemployment insurance and unemployment spells. *Econometrica*, *58*, 757–782.

Meyer, B. D. (1996). What have we learned from the Illinois Reemployment Bonus experiment? *Journal of Labor Economics*, *14,* 26–51.

Michalopoulos, C., Card, D., Gennetian, L., Harknett, K., & Robins, P. K. (2000). *The Self-Sufficiency Project at 36 months: Effects of a financial work incentive on employment and income.* Ottawa: Social Research and Demonstration Corporation.

Michalopoulos, C., Tattrie, D., Miller, C., Robins, P. K, Morris, P., Gyarmati, D., Redcross, C., Foley, K., & Ford, R. (2002). *Making work pay: Final report on the Self-Sufficiency Project for long-term welfare recipients.* Ottawa: Social Research and Demonstration Corporation.

Morris, P., & Michalopoulos, C. (2000). *The Self-Sufficiency Project at 36 months: Effects on children of a program that increased parental employment and income.* Ottawa: Social Research and Demonstration Corporation.

Ondrich, J., & Rhody, S. E. (1999). Multiple spells in the Prentice-Glockler-Meyer likelihood with unobserved heterogeneity. *Economics Letters*, *63*, 139–144.