An Econometric Analysis of the Impact of the Self-Sufficiency Project on the Employment Behaviour of Former Welfare Recipients

Jeffrey Zabel
Saul Schwartz
Stephen Donald

An Econometric Analysis of the Impact of the Self-Sufficiency Project on the Employment Behaviour of Former Welfare Recipients

Jeffrey Zabel
Tufts University

Saul Schwartz
Carleton University and IZA Bonn

Stephen Donald
University of Texas at Austin

Discussion Paper No. 2122
May 2006

ABSTRACT

An Econometric Analysis of the Impact of the Self-Sufficiency Project on the Employment Behaviour of Former Welfare Recipients[*]

The Self-Sufficiency Project (SSP) was a Canadian research and demonstration project that attempted to "make work pay" for long-term income assistance (IA) recipients by supplementing their earnings. The long-term goal of SSP was to get lone parents permanently off IA and into the paid labour force. The purpose of this study is to evaluate the impact of SSP on employment and non-employment durations and its overall effect on employment rates. We focus on generating estimates of the "effect of the treatment on the treated" (TOT) where the "treated" are those in the program group who qualified for the earnings supplement by finding a full-time job during the qualifying period (a group we call the "take-up" group). To obtain a consistent estimate of TOT we follow the work of Ham and LaLonde (1996) and Eberwein, Ham and Lalonde (1997) in estimating a joint model of non-employment and employment durations that controls for unobserved heterogeneity and non-random selection into work and into the take-up group. We find evidence of significant impacts of SSP on non-employment and employment durations. Simulation results show a TOT on the employment rate at 52 months after baseline of approximately 4 percentage points; a 10 percent increase compared to the control group. Further, this estimate of TOT using the results from our econometric model is 5 percentage points higher than the estimate from the raw data.

Corresponding author:

Saul Schwartz
School of Public Policy and Administration
Carleton University
Room 1005 Dunton Tower
1125 Colonel By Drive
Ottawa, Ontario, K1S 5B6
Canada
Email: saul_schwartz@carleton.ca

---

1. Introduction

The Self-Sufficiency Project (SSP) was a Canadian research and demonstration project that attempted to "make work pay" for long-term income assistance (IA) recipients by supplementing their earnings.[1] Lone parents on IA qualified for a generous earnings supplement if they took up full-time work and left the welfare rolls within twelve 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 (the "take-up group") could return to IA 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.

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. 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.

Program designers theorized that the SSP earnings supplement would induce women who would otherwise have stayed on IA to enter and remain in 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. Of particular interest, then, is the "long-term" impact of SSP on those program group members who were eligible for the SSP supplement. One purpose of this study is to evaluate the "short-term" and "long-term" impacts of SSP on the duration of employment and non-employment.[2] In doing so, we generate estimates of the "effect of the treatment on the treated" (TOT) where the

---

[1] The project took place in two Canadian provinces, British Columbia and New Brunswick, between November, 1992 and December 1999 and was funded by the federal department now known as Human Resources Social Development Canada (HRSDC) which was responsible for Canadian labour market programs and policy. 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 MDRC, a New York-based research organization and by SRDC. See Lin et al (1998) for a full description of SSP program.

[2] Here "short-term" means "while the take-up group was eligible for the earnings supplements" and "long-term" means "beyond the period of supplement eligibility". The period over which participants were followed is limited to the 54 months after random assignment, when the last survey was administered.

"treated" are the take-up group (i.e., those in the program group who found full-time work in the first twelve months after random assignment and hence received the 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 of participants is independent, statistically, 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, in the 52nd month after random assignment, we find that the employment rate of the program group was 2.2 percentage points higher than the control group.

Despite the power of experimental evaluation, some policy-relevant issues cannot be addressed with simple program-control comparisons. Among these are the effects of the program on wages, non-employment 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 a non-employment spell) while paying relatively little attention to strategies that enable welfare recipients to stay in their new jobs (lengthening the duration of employment). Perhaps most importantly, if SSP program group members neither experienced wage growth nor increased their 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 for a number of reasons. First, 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 since those who worked are non-random sub-samples of the program and control groups. Second, in SSP and in other programs for disadvantaged women, the "treatment" takes the form of an offer rather than being applicable to all program group members. In SSP, only a fraction of program group members (37.3%) became eligible to receive the earnings supplement by finding a full-time job within 12 months of random assignment. Those who choose to be "treated" in this way are a non-random subset of the program group and selection into the treatment must also be addressed.

In the next section, we discuss two previous papers that use similar methods to study similar issues using experimental data from two earlier evaluations of programs for disadvantaged adult women. In Section 3, we

discuss some of the basic information about SSP that has been released in previous analyses (see, for example, Lin *et al.,* 1998, Michalopoulos *et al.,* 2000, Morris *et al.,* 2000, and Michalopoulos *et al.,* 2002) and we derive other results from our own calculations. In Section 4, we describe the data and present some descriptive analysis of the non-employment and employment spells of the program and control groups. We develop an econometric model of non-employment and employment duration that allows us to evaluate the short- and long-term impacts of SSP in Section 5. In Section 6, we estimate this model and present the empirical results. In Section 7, we use the results obtained in Section 6 to carry out a simulation of the overall impact of SSP on the employment rate of the take-up group. Section 8 summarizes our sense of how SSP affected the employment behaviour of the take-up group and hence whether it was successful in meeting its long-term goals.

2. Two Previous Non-experimental Analyses of Experimental Data

Ham and LaLonde [hereafter HL] (1996) studied the effects of the National Supported Work (NSW) Demonstration; a program that provided work experience to a randomly-assigned group of disadvantaged women (called "trainees" by HL). All program and control group members were in the midst of a spell of unemployment at baseline. To control for the potential bias introduced by the non-random selection into work by both program and control group members, HL allowed for individual-specific error terms (representing unobserved heterogeneity) that were correlated across the unemployment and employment duration equations.

HL assumed that all program group members received a "treatment" consisting of work experience in a supportive environment.[3] This feature of NSW created a particularly difficult problem because all program group members would be starting "fresh" spells of unemployment or employment at the end of their period of supported work.[4] By contrast, control group members starting fresh spells would be a rather select group since they would first have to end the spell of unemployment ongoing at baseline. This programmatic feature forced HL to pay particular attention to an "initial conditions" problem – how to handle the spell of unemployment that was "interrupted" by random assignment.

---

[3] HL note, however, that a small number of program group members (14 individuals) were randomly assigned to the program group but were "no shows" and never received any work experience.

[4] A "fresh" employment spell is one that began after random assignment. In SSP, all employment spells are fresh except for those that began before random assignment for about 10 percent of the sample.

The second type of non-random selection that HL addressed occurred because some program group members entered a fresh employment spell directly from supported work while other trainees left supported work for a fresh spell of unemployment. HL account for this endogenous choice by specifying and estimating separate equations for the trainees who entered fresh employment and unemployment spells and by allowing the unobserved heterogeneity terms to be correlated with those in the other equations.

NSW differs from SSP in several relevant ways. First, SSP program group members were not removed from the labour force upon random assignment. The spells of non-employment that were on-going at baseline continued for both the program and the control group. Indeed, a large proportion of both groups never found full-time work within the 54-month follow-up period. Second, SSP did not involve any specific activities for the program group so there was no single point at which they left program activities to start fresh spells of employment or unemployment. However, there was a fixed point — twelve months after random assignment — after which program group members could no longer become eligible for the earnings supplement.

As was true for HL in their analysis of NSW, we must account for the non-random selection of SSP program and control group members into employment. And, like HL, we must also account for a second kind of selection. In SSP, however, this second selection is different; by program group members who chose to become eligible for the SSP earnings supplement by finding full-time work within 12 months of random assignment.

Eberwein, Ham, and LaLonde [hereafter EHL] (1997) also attempted to estimate the effect of the treatment-on-the-treated (TOT) in an experimental setting. In the National Job Training Partnership Act (JTPA) evaluation, one of the treatments was "classroom training" (CT). EHL were concerned with the studying the effect of those who received CT on employment and unemployment durations.[5] Like HL, EHL must account for non-random selection into work for both the program and the control groups. They did so in a similar way, specifying unobserved heterogeneity as part of the error terms in equations representing the propensities to exit employment and unemployment and allowing the unobserved heterogeneity to be correlated across the two equations.

The JTPA experiment was similar to SSP in that the program group was not removed from the labour force. The program group members who were assigned to CT could remain in whatever labour force status they occupied at baseline. Not all program group members actually received the treatment, however, since it was not

_____

[5] The "treatment" was actually the package of job-related services in which one component was classroom training (EHL, 1997, p. 658).

mandatory. The choice of whether or not to receive CT is thus endogenous. To estimate the TOT, EHL replaced the endogenous binary training indicator with an equation specifying the transition rate into training as a function of the exogenous binary program-control indicator. They allow the unobserved heterogeneity term in this equation to be correlated with those in the other equations. In estimating the effect of the SSP treatment on those who received it, we adopt a strategy that is similar in spirit to the EHL model yet differs in its actual implementation. We discuss our estimation procedure in detail in Sections 5 and 6.

3. The SSP Data

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. The number of 54-month survey respondents, excluding those who were assigned to SSP Plus, was 4,852; 2,460 respondents were in the program group and 2,392 were 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 been in any calendar month from November 1992 to March 1995. Participants who had missing values on one or more of the survey questions used in the multivariate models that we estimate were also excluded. The resulting sample size for our analysis is 4,769 (2,415 in the program group and 2,354 in the control group).

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 are available for only 51 months after random assignment. We use the employment status in month t-1 to determine the value of the employment status variable assigned in month t. Since employment status in the 51st month determines employment status in the 52nd month, there are 52 months over which spells of employment and non-employment can be calculated.

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 interview and were asked about every new job held since the baseline.[6]

An interesting and somewhat unexpected feature of the SSP sample is that a significant proportion of participants — about 10 percent — were already working full time in the month before they were randomly assigned. For those in this 10 percent who were randomly assigned to the program group, the SSP supplement was a "windfall" gain since they were already working full time and could qualify for the supplement without changing their employment status (although they would have to leave income assistance).

We provide the definitions of a number of characteristics of SSP participants used in our analyses in Table 1a. We give summary statistics for these variables for all program and control group members in columns (1) and (2) of Table 1b. One can see that the mean values are very similar across the two groups. We conducted t-tests for the equality of the population means for each variable. Only for MONTHIA (the number of months on income assistance in the three years prior to the baseline interview) do we reject the null hypothesis at the 5 percent significance level. Even for that variable, the difference in the means is relatively small. These results suggest that random assignment "worked" and that the two groups were quite similar in terms of these characteristics.[7]

4. Descriptive Measures of Non-employment and Employment Spells

We next describe the duration of non-employment and employment spells with only minimal modeling of the processes that determine the durations. The development and estimation of econometric models that account for self-selection, unobserved heterogeneity and duration dependence appears in the following sections.

"Employment", in this paper, 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 "non-employed" in this paper when they are unemployed, out of the labour force or working part-time. We will use the terms "employment" and "non-employment" to simplify terminology.

In Figure 1, we compare the employment rates for the entire program and control groups in our sample over the 52-month follow-up period. We see that SSP had a significant short-run experimental impact that peaked

---

[6] 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 used the adjusted data in this paper.

13-15 months after random assignment. After 52 months, the difference in full-time employment rates between the two groups was much smaller; 2.2 percentage points.[8] The gradual diminution of the program-control gap is sometimes referred to as "control group catch-up" since, over the follow-up period, the program group employment rates remained roughly constant, at the 13-15-month level, as the control group rates rose.

In this analysis, we are interested in determining the impact of SSP on the take-up group; those members of the program group who found work in the first 13 months since baseline.[9] In Figure 2, we show the employment rates for the self-selected take-up group. The appropriate comparison group is not the full control group but those control group members who found full-time work in the first 13 months. We refer to this group as the Control_13 group. Given that both the take-up and Control_13 groups have selected into full-time work, differences in their employment rates better reflect the impact of SSP than if we compared the take-up group to the full control group. The difference between the employment rates of the take-up and Control_13 groups measures the TOT without any econometric adjustment for heterogeneity (observed or unobserved) and the non-random selection into full-time work or into the take-up group. The "raw" non-experimental TOT peaks at month 14 with a value of 16.6 percentage points. The estimate of TOT remains above 10 percentage points until month 35. It then declines and actually is -1.4 percentage points in month 52 (46.5% versus 47.9%). These "raw" effects, however, might well be affected by the various selection biases discussed in Section 2. To determine the TOT controlling for observed and unobserved factors and for the non-random selection into full-time work and into the take-up group requires the econometric analysis that we carry out later in the paper.

Note that the overall program-control difference is a positive 2.2 percentage points despite the fact that the TOT is -1.4 percentage points. This is because the take-up group comprises 37.3% of the program group while Control_13 makes up only 23.3% of the control group. Further, the non-take-up group had an employment rate of

---

[7] See Lin et al. (1998), p.125, for similar analyses that also confirm that random assignment "worked".
[8] The program-control difference reported in the final SSP report is slightly different because, in that report, the full program and control groups were used. On average, over the first quarter of the fifth year after random assignment (i.e., months 49-51), the program-control difference was 3.3 percentage points. In the second quarter of the fifth year, the average difference dropped to an insignificant 1.6 percentage points.

[9] As described above, those randomly assigned to SSP had exactly 12 months to qualify for the earnings supplement by finding a full time job. The employment variable measured in the surveys, however, is based on calendar months. Participants whose 12-month qualifying period started on January 21st, for example, would qualify for the supplement if they left income assistance and found a full-time job by January 21st of the following year. However, in the survey data, "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

only 16.0% in month 52 while the employment rate for the control group members who did not find full-time work in the first 13 months was only 18.4%. That is, despite the fact that the employment rates for the two sub-groups who did and did not find work in the first 13 months are lower for the program group, the employment rate is 2.2 percentage points higher for the program group compared to the control group because of the compositional effect.

We now describe the initial interrupted non-employment spell and the initial fresh employment spell for SSP participants by calculating empirical hazard functions using the Kaplan-Meyer product-limit estimator. Some basic information about the incidence of subsequent spells of employment and non-employment is also provided.

Empirical hazards for the initial interrupted non-employment spell: For the initial interrupted non-employment spell, Figure 3 shows the empirical hazards of exiting non-employment for the program and control groups in each of the 51 months for which hazard rates can be calculated.[10] Since program group members had a much greater incentive to find full-time work in the first thirteen months, it is not surprising that their empirical hazards are higher than those of the control group during this period. After the end of the thirteen-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 and Figure 3 shows that their empirical hazard rate is no longer consistently higher than that of the control group.

Empirical hazards for the first fresh employment spell: There are several important differences between the analysis of fresh employment spells and the analysis of the initial interrupted non-employment spell. The first interrupted non-employment spell began at the point of random assignment for all participants (except the 10 percent who were working full-time in the month prior to random assignment) and the program and control groups differed only in that the program group had a stronger incentive to find full-time employment. By contrast, spells of interrupted full-time employment exist only for a minority of all SSP participants and can begin at any point between random assignment and the end of the 51st month after random assignment.

---

coded as starting in "Month 13." For that reason, we actually define the qualifying period to be the first 13 months after random assignment.

[10] 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 4 shows the empirical hazard functions for the program and control groups, where the hazard is the conditional probability of exit from the first fresh full-time employment spell. Control group members were 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 of the spell.

In describing the employment spells, it is crucial to recognize that while there are still two groups that face different incentives, the two groups are no longer the entire program and control groups. During their 36-month supplement eligibility period, take-up group members have a stronger incentive to stay employed than do control group *or* non-take-up group members. Thus, while 53.3% of take-up group members had a first fresh employment spell that was at least 12 months in duration, only 31.5% of non-take-up group members and 32.7% of control group members had spells of similar lengths.

Multiple Spells of Non-employment and Employment: Many SSP participants had more than one non-employment 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 group after their 36-month supplement eligibility period had ended. The first employment spell lasted less than 36 months for almost 80 percent of the take-up group so spells beyond the first must be considered.

About one-third of the control group and over 40 percent of the program group had *more* than one spell of non-employment, meaning that they found a full-time job after random assignment and then left that job within the follow-up period. Smaller percentages (27 percent of the program group and 22 percent of the control group) had more than one spell of full-time employment.[11]

5. A Multivariate Model of Non-employment and Employment Durations

In this section, we develop an econometric model of non-employment and employment durations that will allow us to obtain estimates of the impact of the SSP incentive that are not subject to the selection biases discussed earlier. We can estimate the impact prior to take-up, during the incentive period, and after the incentive period is

---

[11] Note that spells of employment and non-employment are period of continuous work or non-work. Hence an employment spell can include more than one job as long as there was no break between jobs.

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 analysis of the previous section for at least two reasons. First, an important assumption in estimating the empirical hazard rate functions in Figures 3 and 4 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 non-employment 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 non-employment and into the take-up group, unobserved heterogeneity, and duration dependence. Our econometric analysis provides such a framework.

Our model is in the spirit of the training model presented in EHL (1997, especially p. 666). Following both HL and EHL, we estimate equations for the first, interrupted, non-employment and employment spells that are distinct from the equations for subsequent fresh non-employment and employment spells. We do this for two reasons. First, the initial non-employment and employment spells are experimental in nature since they include all the program and control group members. Second, as discussed in HL and EHL, the first non-employment and employment spells are interrupted spells (that is, they started before baseline but are measured from the point of random assignment) and hence are different from subsequent fresh non-employment and employment spells.

Define y^{*}_{uidt} and y^{*}_{eidt} as latent measures of the non-employment and employment exit propensities for person i, spell duration d, and in (relative) time period t. We specify the equation for the first, interrupted, non-employment and employment spells as:

$$y^{*}_{uidt} = \alpha_{0u} + X_{it}\alpha_{1u} + \alpha_{2u}\text{PRE - INCENTIVE}_{it} + \alpha_{3u}\text{NO - INCENTIVE}_{it} \\ + h_{u}(d,\alpha_{4u}) + \theta_{u1i} + \varepsilon_{u1it}$$

(1)

and

$$y^{*}_{eidt} = \alpha_{0e} + X_{it}\alpha_{1e} + \alpha_{2e}\text{INCENTIVE}_{it} + \alpha_{3e}\text{POST - INCENTIVE}_{it} \\ + h_{e}(d,\alpha_{4e}) + \theta_{e1i} + \varepsilon_{e1it}$$

(2)

We observe:

$$y_{uidt} = \begin{cases} 0; \text{remain non-employed if } y_{uidt}^* \le 0 \\ 1; \text{exit non-employment if } y_{uidt}^* > 0 \end{cases}$$

and

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

where X_{it} is a vector of observable covariates. The h(.) functions are general functions that capture duration dependence and ε_{j1it} (j = u, e) are unobserved i.i.d. error terms.

The key program variables in equation (1) are the time-varying binary indicators PRE-INCENTIVE and NO-INCENTIVE. PRE-INCENTIVE is a 0-1 indicator that takes the value 1 for program group members during the first 1**3** months after random assignment; when they could become eligible for the supplement by finding a full-time job. NO-INCENTIVE is a 0-1 indicator that takes the value 1 for program group members after the 13-month "window" to qualify for supplement eligibility had closed. These variables are included to measure, respectively, the impact of being in the program group during and after the 13-month qualifying period. They are always zero for control group members. Because of the experimental nature of SSP, these two variables are defined for all program group members who were non-employed at baseline and hence are exogenous. During the first thirteen months of their initial non-employment spell, these program group members had a substantial financial incentive to find full-time work. The coefficient on PRE-INCENTIVE will capture the effect of this incentive on the exit propensity from the first non-employment spell. The coefficient on NO-INCENTIVE will capture the effect of SSP on those who did not qualify for the supplement.

Equation (2) is estimated for the 10 percent of the sample who were employed full-time at baseline. The key program variables in equation (2) are the time-varying binary indicators INCENTIVE and POST-INCENTIVE. INCENTIVE and POST-INCENTIVE are, respectively, 0-1 indicators that take the value 1 for program group members during the first 36 months after random assignment when they are eligible for SSP benefits and after 36 months when benefits have expired. These variables are always zero for control group members. We use these two separate indicators since we believe that the impact of SSP on employment duration differed during the incentive and post-incentive periods. Also, we are particularly interested in determining if SSP continued to have an impact on program group members once the eligibility period ended.

Next, we specify the equations for fresh non-employment and employment spells:

$$y^*_{uidt} = \beta_{0u} + X_{it}\beta_{1u} + \beta_{2u}\text{INCENTIVE}_{it} + \beta_{3u}\text{POST - INCENTIVE}_{it}$$
$$+ h_u(d,\beta_{4u}) + \theta_{u2i} + \varepsilon_{u2it}$$
(3)

and

$$y^*_{eidt} = \beta_{0e} + X_{it}\beta_{1e} + \beta_{2e}\text{INCENTIVE}_{it} + \beta_{3e}\text{POST - INCENTIVE}_{it}$$
$$+ h_e(d,\beta_{4e}) + \theta_{e2i} + \varepsilon_{e2it}$$
(4)

In equations (1) and (3), θ_{u1i} and θ_{u2i} represent unobserved heterogeneity in individual propensities to exit non-employment. In equations (2) and (4), θ_{e1i} and θ_{e2i} represent unobserved heterogeneity in individual propensities to exit employment. Conditional on the observables, individuals with relatively large values of θ_{u1i} and θ_{u2i} will be more likely to exit non-employment, leaving those who remain non-employed with relatively low values of θ_{u1i} and θ_{u2i}. Again conditional on the observables, individuals in the latter group will have a lower average probability of exiting non-employment. If we do not account for θ_{u1i} and θ_{u2i}, this decline in the exit probability over the duration of the spell will be erroneously identified as negative duration dependence. We will control for the sample-selection bias that arises from the non-random selection into fresh non-employment and employment spells by allowing the unobserved heterogeneity terms to be correlated when the equations of the model are jointly estimated.

Note that the coefficients in equations (1)-(4) measure the change in the probability of exiting non-employment and employment. For example, positive values for the α's and β's associated with X_{it} and the program variables will increase the probability of exit and hence *decrease* the duration of non-employment and employment.

The key program variables in equations (3) and (4) are the time-varying binary indicators INCENTIVE and POST-INCENTIVE. These variables are the same as in equation (2) with one important difference. They only take on the value of 1 for program group members who had qualified for the supplement – the take-up group. These variables are always zero for control group members and for program group members who did not qualify for the supplement (the non-take-up group). Since program group members must choose to take up the 36-month SSP incentive program, the variables INCENTIVE and POST-INCENTIVE in equations (3) and (4) are endogenous. In the next section, we discuss how we control for this endogeneity in the estimation process.

6. Empirical Results

In this section, we present the empirical results for the econometric model presented in Section 5. In the first sub-section, we present the model estimator and discuss some special issues that arise in its development. In the second sub-section we present the estimation results.

6.1 Model Estimator

Two important features of the model presented in the last section must be reflected in our estimator if it is to provide consistent coefficient estimates. First, the variables INCENTIVE and POST-INCENTIVE in the fresh non-employment and employment spell equations (3 and 4) are endogenous. Second, whereas the interrupted non-employment and employment spell equations (Equations 1 and 2) are experimental due to the randomized nature of SSP, this is not the case for the fresh spells because individuals non-randomly chose to exit the interrupted spells for the fresh spells. While they may seem similar, these two issues are actually quite different. The first results from program group members who chose to take-up the supplement by becoming employed full-time during the 13-month qualifying period. The second results from all SSP participants — in both the program and control groups — who chose to enter fresh spells of employment or non-employment. Our estimator controls for the first problem by creating instruments for the program participation variables and controls for the second problem by allowing the unobserved heterogeneity terms in Equations (1)-(4) to be correlated. Our estimator is developed formally in the Appendix. Here we describe, in a less formal way, how we handled each problem.

Endogeneity of SSP Program Variables: Since program group members non-randomly chose to take up the SSP incentive, the take-up variable that defines this sub-group is endogenous. In order to account for this endogeneity, we specify another equation for whether or not an individual found full-time employment in the first 13 months after random assignment, using PROGRAM — a 0-1 indicator of whether or not the person was randomly assigned to the regular SSP group — as an instrument:

$$\text{WFT13}_i^* = \pi_0 + X_i \pi_1 + \pi_2 \text{PROGRAM}_i + \varepsilon_i \qquad (5)$$

For all participants, we observe

$$\text{WFT13}_i = \begin{cases} 0; \text{remain non-employed for the first 13 months if } \text{WFT13}_i^* \leq 0 \\ 1; \text{employed in at least one of the first 13 months if } \text{WFT13}_i^* > 0 \end{cases}$$

One can view equation (5) as the instrumenting equation for the endogenous take-up variables in equations (3) and (4). The consistent estimator involves the joint estimation of equations (1)–(4) and (5). Details are provided in the Appendix. Note that INCENTIVE and POST-INCENTIVE in equations (3) and (4) represent the subset of the program group who took up the SSP earnings supplement, defined in different periods after random assignment and hence represent only one endogenous variable. That is, only PROGRAM is used in the construction of the two variables. Hence the one instrument — PROGRAM — is enough to identify the model.

Non-Random Selection into Full-time Work: In order to control for the non-random selection into subsequent employment and non-employment spells, Ham and LaLonde (1996, hereafter HL) allow the individual-specific error terms (i.e., the θ's representing unobserved heterogeneity) to be correlated across the non-employment and employment duration equations. Whereas standard selection models include a selection and a behavioural equation that occur contemporaneously, in this case the non-employment and employment equations occur sequentially (i.e., dynamically). EHL call the bias that can arise in this framework as "dynamic selection bias." Following HL, we model the unobserved heterogeneity terms as:

$$\theta_{uj} = \alpha_{1uj} + \alpha_{2u}\theta^* \qquad \text{and} \qquad \theta_{ej} = \alpha_{1ej} + \alpha_{2e}\theta^*, \quad j=1,2 \tag{6}$$

where θ^* is drawn from a two-point distribution. Normalizing $\alpha_{1u1} = 0$ and $\alpha_{2u} = 1$ reduces the number of parameters to estimate by two. We allow for only two points of support because of the computational complexity of the model.[12] 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.

6.2 Empirical Results

In this sub-section, we present the empirical results for the econometric model presented above. The data consist of non-employment and employment spells for the 2,415 program and 2,354 control group members for 52 months after baseline. This includes up to 6 non-employment and 6 employment spells for each individual. There are a total of 6,956 spells; 4,442 interrupted non-employment, 1.045 fresh non-employment, 347 interrupted employment, and 1,142 fresh employment spells.

---

[12]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.

We include in the non-employment and employment equations a set of covariates, measured in the baseline survey and described in Tables 1a and 1b. We include a series of dummy variables that indicate if an individual: was single and not previously married; did not have a high school degree; lived in British Columbia; reported limited activity due to a long term emotional, psychological, nervous or mental health condition or problem; had a child at baseline who was less than or equal to 4 years old; fell in specific age ranges between 19 and 40 or more. Continuous variables include the number of children, the number of months on income assistance in the 3 years prior to the baseline interview, and the average monthly income assistance benefit received in the four quarters prior to baseline interview.[13]

We allow for duration dependence by including the log of time, its square, and its cube. Higher order terms were generally not significant. This is a restriction on including a dummy variable for each period but the restriction greatly reduces the number of parameters in the model.[14] We estimate the model with one common baseline hazard so that the dummy variables representing the program effects capture the full differences between the program and control groups.[15] HL recommend controlling for pre-baseline duration in the interrupted spells, claiming that this better identifies these equations. Thus, in the interrupted non-employment equation, we include PRE-BASE UNEMP, a variable measuring the number of months of non-employment in the 12 pre-baseline months. In the interrupted employment equation, we include PRE-BASE EMP, the number of months of full-time employment in the 12 pre-baseline months.

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 non-employment spells. We decided to update only age, at the beginning of each spell. It might have been useful to update other variables — 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, responses on the 18, 36 and 54 month surveys tell us only whether there has been a change in these variables since

---

[13] Initially, we also included the time-varying, province-specific monthly unemployment rate. Generally, this variable was not significant. Hence we did not include it in the final model because its presence complicated the estimation procedure.

[14] Further, Baker and Melino (2000) showed that models with nonparametric specifications for both duration dependence and unobserved heterogeneity do not perform well.

[15] We could have calculated the program effect by estimating separate hazards for the program and control groups, 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 seemed unnecessarily complicated.

the last interview. Thus, at best, we can only update on the survey dates and not necessarily at the month when the change actually occurred. Second, it may well be that fertility, educational attainment, and marital status are not only endogenous but are also affected by program take-up. In order to get the full program effect, these impacts would need to be taken into account. By not updating the covariates, any impact of SSP on these variables will end up in the overall program effect as measured by the coefficients on the program variables, as is appropriate for our analysis.

In estimating equations (1)-(4), we allow the coefficients for the program variables PRE-INCENTIVE, INCENTIVE and POST-INCENTIVE to vary over time, thus allowing the impact of SSP to vary over time within these three periods. For the sake of parsimony, we grouped months for which the coefficient estimates were similar (in an economic and a statistical sense). For example, we know there is a spike in the exit rate for the first non-employment spell near month 13 as program group members strive to become eligible for the supplement. We therefore allow the impact of PRE-INCENTIVE to vary across two periods: the first nine months and months 10-13. For fresh non-employment spells, we allow two different impacts of INCENTIVE in months 2-8 and 9-36.[16] The impact of the same variable is allowed to vary in four periods — months 1-8, 9-16, 17-32, and 33-36 — in the employment spell equation. Finally, we allow the impact of POST-INCENTIVE to differ across the three possible four-month intervals.

In Table 2, we present three different sets of results: (1) the four equations (interrupted non-employment and employment and fresh non-employment and employment spells) estimated separately with no individual heterogeneity and without accounting for the endogeneity of the take-up variable; (2) the same four equations estimated jointly with correlated individual heterogeneity but without accounting for the endogeneity of the take-up variable; and (3) the four equations estimated jointly allowing for correlated individual heterogeneity and accounting for the endogeneity of the take-up variables. With few exceptions, the three sets of results are very similar. Apparently, including individual heterogeneity and accounting for the endogeneity of taking up the SSP supplement has little impact on most of the parameter estimates. Given such similar estimates, we focus our discussion of the third set of results for the estimator that includes correlated individual heterogeneity and accounts for the endogenous take-up variable (columns 5 and 6 of Table 5). We will mention the few cases where the results differ across estimators.

Apart from the coefficient estimates for the demographic and program variables, the model produces several other parameter estimates that bear on the importance of unobserved heterogeneity and on the possible cross-equation correlation of the heterogeneity terms. These parameter estimates are given at the bottom of Table 2. The parameters μ_1 and μ_2 are the two points of support for the unobserved heterogeneity. The estimates are -0.9386 and -1.0532. Though statistically significant, these estimates are relatively close in magnitude and are not significantly different from each other. The probability that μ_1 will occur is prob $= e^\gamma/(1+e^\gamma)$ (if $\gamma = 0$ then prob $= 0.5$). With $\hat{\gamma} = 0.2907$, prob $= 0.572$. The standard error for $\hat{\gamma}$ is large, so one cannot reject the null hypothesis that there is equal probability that each point of support will be realized. The parameters α_{1u2}, α_{1e1}, α_{1e2}, and α_{2e} determine θ_{u2}, θ_{e1}, and θ_{e2} in equation (6). Note that none of these parameter estimates are significantly different from zero. Overall, there is little evidence of significant unobserved heterogeneity in the model.

The variables in the model are of two types. The first is the set of demographic characteristics that we use as covariates and that are described above.[17] The second is the set of program variables, described in the last section, that capture the effects of random assignment to SSP at different points in the follow-up period. Given the structure of the model, the coefficient estimates must be transformed to have the useful interpretation of the change in the probability of exiting non-employment or employment for a unit change in the variable in question. In the case of the binary program variables, we calculate the difference in the estimated probability of exiting non-employment or employment with the variable set at 1 and 0 for each person. Results are given at the mean of the difference over the sample observations and are presented in Table 2 along with the actual coefficient estimates.

Results for the interrupted baseline spell of non-employment: For the interrupted baseline spell of non-employment, represented by equation (1) and shown in columns (5) and (6) of Table 2, the coefficients corresponding to the two pre-incentive variables show that a program group member had a 0.0115 higher probability of exiting non-employment in any given month within the first nine months after random assignment and a 0.0270 higher probability of exiting non-employment in months ten through thirteen, than the control group (holding all observable characteristics constant). Both impacts are highly significant. The latter represents a 215

---

[16] Note that the first month of the incentive period must be part of an employment spell so the earliest that an individual can be in her second non-employment spell is month two of the incentive period.

[17] The coefficient estimates for the demographic variables are not shown but are available from the authors.

percent increase in the probability of finding employment (0.040 vs 0.014).[18] This result is comparable to the

higher estimated hazard rate implied in Figure 3 for the program group. It is also consistent with the theoretical

model in Card and Hyslop (2005) in which the reservation wage of program group members is declining over the

duration of the pre-incentive period. Finally, note that the coefficient for NO-INCENTIVE is small and not

significant.[19]

Many of the demographic covariates in this duration equation are significant at the five percent level (not

shown, but available on request). Participants with young children, less education, or emotional problems at

baseline were less likely to exit the baseline non-employment spell (holding constant the program indicators). The

results also imply a probability of gaining full-time employment that declined with each age category, as indicated

by declining coefficients on the age variables. Greater baseline IA benefits and more months on IA in the three

years prior to baseline also decreased the likelihood of exiting non-employment. The coefficients for the log of

duration, its square, and its cube are not significant.

Results for fresh spells of non-employment: For the fresh non-employment spells, the impact of being eligible for

the supplement on the exit rate from non-employment is positive but not significantly different from zero when

both heterogeneity and endogeneity are taken into account in columns (5) and (6) of Table 2. This is one place

where there is a substantial difference between the first two models that do not control for the endogeneity of take-

up and the one that does. The impact of SSP supplement on the exit rate from non-employment is positive and

significant in months 2-8 and 9-36 (see columns 1 and 3) with an exit probability of roughly 0.01, similar to the

impact during months 1-9 of the pre-incentive period. Once the endogeneity of take-up is accounted for, these

impacts decrease in size by about two-thirds and are no longer significant. An implication of this result is that the

average take-up group member was more likely to exit non-employment than a randomly chosen program group

member.

---

[18] We calculate the probability of exit by taking the mean of the predicted probability over the full sample of
observations. First, we calculate the mean probability for the control group, a mean that is 0.014. Then we
calculate the mean predicted probability for the program group in months 10 -13 of the pre-incentive period. In
this case the mean is 0.040.

[19] We had speculated that those program group women who did not find a job in the first 13 months (NO-
INCENTIVE=1) might have been more likely to exit non-employment than the control group because they were
more intensively searching for work during the pre-incentive period but this does not appear to be the case.

None of the three coefficients on the POST-INCENTIVE variables is statistically different from zero. This evidence suggests that once the incentive program ends, the take-up group members are no more likely to exit non-employment than are the control group members. This is the first important evidence that SSP did not seem to have a long-term impact on the labour force behaviour of the take-up group.

The effects of the other baseline covariates (not shown) are generally less important in the fresh spells of non-employment than in the first interrupted spell. Still, participants with less education, emotional problems, or higher IA benefits at baseline were significantly less likely to exit fresh non-employment spells at the 1% level. The coefficients on the log of duration, its square, and its cube are once again not significantly different from zero at the 1 percent level.

Results for interrupted baseline spell of employment: The results for the interrupted baseline spell indicate that, for the take-up group, there was a negative and significant impact on the exit rate from employment in the first 8 months. None of the other program indicators are significant even though some of the magnitudes are large. This is not too surprising since there are few individuals who were employed at baseline and who survived in their baseline spell until the end of the incentive period and then into the post-incentive period.

Results for fresh spells of employment: As the coefficients on the INCENTIVE variables in column 5 of Table 2 show, there is a general monotonic decline (in magnitude) in the impact of the subsidy on the duration of the fresh employment spells of the take-up group. In months 1-8, a take-up group member had a 0.0171 lower probability of ending their employment spell in any given month than a control group member, a 31 percent decrease in the probability of exiting employment relative to the control group (0.042 vs. 0.059).[20] The effect of being in the take-up group is still relatively large in months 9-16; a 0.0122 lower probability of exiting employment. The impact is not statistically different from zero in the final twenty months. Given this latter rsult, there does not appear to be a significant change in the behaviour of the take-up group in the last months (months 33-36) of the incentive period; there is no evidence of a "cliff" effect.[21]

---

[20] See note 19 for the method used to calculate these estimates from the coefficients in Table 5.

[21] The "cliff" effect was the name given in previous SSP reports to the possibility that many of the take-up group would work up to the end of the eligibility period and then face a dramatic drop in their earnings, possibly causing a rapid drop in their employment rates.

For the first 8 months of the post-supplement period, there is no difference in the exit rate from employment between the take-up group and the control group. In months 9-12, however, the impact is negative and quite large in magnitude; the probability of exit in a given month is 0.0251 lower for a take-up group member as compared to a control group member. This provides some evidence of a long-term impact of SSP on employment duration; SSP increases the duration of employment after supplement eligibility has ended. Given this trend in the coefficients, it would be useful to obtain data beyond the 54-month period analyzed in this study.

Looking at the demographic covariates; there is a decreasing probability of exiting full-time employment as participants' age and those without a high school degree, with emotional problems, or higher benefits at baseline were significantly more likely to leave full-time employment. There is evidence of positive duration dependence for the first five months and then negative duration dependence thereafter. This is similar to what is seen in the Kaplan-Meier hazard rate estimates in Figure 4.

Overall, there does appear to be an impact of SSP on the probability of exiting non-employment during the pre-incentive period and on the probability of exiting employment during the incentive and post-incentive periods. To gauge the full impact of SSP, we need to be able to estimate the employment rates over the 52 months for the program and control groups. We do this in the next section.

7. Simulation

Using the econometric model estimated in Section 6, 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 to the entire control group since the take-up group is the self-selected subset of the entire program group that found eligible full-time work and received the SSP supplement. The appropriate comparison group is not the full control group but Control_13 group, those members of this group who found full-time employment in the first 13 months.

Recall that the unadjusted non-experimental difference in the employment rate between the take-up group and Control_13 is -1.4 percentage points in month 52 (see Figure 2). To determine the extent of the bias in Figure 2 due to the non-random selection into work and into the take-up group, we carry out a simulation based on the econometric results presented in Section 6.2. This allows us to get an unbiased estimate of the effect of the treatment on the treated (TOT). For each individual in the sample, we randomly generate an error term, drawn

from the normal distribution, for each of 52 months. In addition, for each individual, we randomly generate one of the two values of θ_u and θ_e estimated in Section 6.2 and shown in Table 5. The generated error and unobserved heterogeneity terms (and the values of the variables and coefficients) determine if the individual is non-employed or employed in any given month. We also adjust for duration dependence depending on the number of months in a given employment status. We carry out this exercise twenty times for each individual so that we can get an average picture of each individual's simulated employment behaviour. Using the entire sample of 4,769 individuals, each simulated 20 times, there are a total of 95,380 observations for each simulation. Note that the same 4,769 individuals are used when we simulate results for both the control and program groups. This means that the results are not affected by the choice (albeit random) of which individuals are in the program and control groups.

First, we simulate the employment rates using the estimates for the final model in column (5) of Table 2 and assuming each of the 4,769 individuals are members of the control group. To do so, we set all the program variables to zero for each individual. To be in the simulated Control_13 group, an individual must either have been employed full-time at baseline or predicted to "find a job" (i.e., be randomly given values of the normal error and unobserved heterogeneity terms that lead to a prediction of full-time employment) in the first thirteen months. Next, we simulate the employment rates assuming each individual is a member of the program group. In this case, the program variables are set as if all 4,769 individuals are in the program group. Again, to be in the simulated take-up group, an individual must either have been employed full-time at baseline or predicted to "find a job" in the first thirteen months.

Given these two simulations, we first look at the employment rate differential for the full set of 95,380 observations for the program and control groups. This gives an idea of how well the simulation is calibrated to the actual employment rates for the 2,415 program and 2,354 control group members. In month 52, the simulated program and control group employment rates are 26.1 and 21.8 with a difference of 4.3 percentage points. The actual employment rates are 27.5 and 25.3 with a difference of 2.2 percentage points. Thus the simulation shows a treatment effect that is 2.1 percentage points higher than the actual treatment effect. This occurs for at least two reasons.

First, we run the simulations using all 4,769 individuals for both the program and control group simulations. When we use only the 2,415 program group members to get the program group employment rates and the 2,354 control group members to get the control group employment rates, the difference at month 52 is 3.6

percentage points. Thus there is a 0.7 percentage point reduction in the treatment effect due to the differences in the covariates across the actual members of the program and control groups. Second, the simulation does not pick up the full extent of the control group "catch up" at the tail end of the follow-up period. When we use only program group members to get the program group employment rates and the control group members to get the control group employment rates, the employment rate differential in month 47 is 4.1 percentage points. This is exactly the same differential in the actual data. In month 52, this differential falls to 3.6 percentage points using the simulated data whereas it falls to 2.2 percentage points using the actual data; this is a 1.4 percentage point differential. This final increase in the monthly employment rate of the control group in the actual data is mostly due to those members who did not find a full-time job in the first 13 months and is not picked up in the simulated data.

Turning now to a comparison of the take-up group and the Control_13 group, we present the simulated employment rates for these two groups in Figure 5. The difference in employment rates at month 52, the estimate of TOT, is 3.6 percentage points. This is 5 percentage points higher than the estimate of TOT from the actual data, -1.4 percentage points. Thus accounting for the selection into work and into the take-up group is important for getting an accurate estimate of TOT. Recall from the previous section that we found that SSP led to a higher probability of exiting non-employment during the pre-incentive period and a lower probability of exiting employment during the incentive and post-incentive periods. The combination of these impacts results in an employment rate for the take-up group that is 3.6 percentage points higher than the control group.

8. Conclusion

The power of social experiments such as the Self-Sufficiency Project lies in their ability to answer some kinds of questions without the fear that self-selection bias is influencing the answers. Simple comparisons of the program and control groups provide unbiased estimates of the treatment effect. However, questions involving the experiences of non-random subsets of the program and control groups — for example, the effect of SSP on those who found full-time work and actually received earnings supplements — cannot necessarily be answered with simple program-control comparisons.

The kind of econometric modeling that we have undertaken in this paper attempts to answer the latter type of question by estimating the parameters of a joint model of non-employment and employment durations that (1)

includes an explicit specification of the unobserved heterogeneity that might create dynamic selection bias and (2) accounts for the bias due to the non-random selection into the take-up group. This should allow us to obtain more accurate and detailed estimates of the impact of SSP than would otherwise be possible.

Our ability to control for the non-experimental aspects of the data depends on specifying the unobserved individual heterogeneity. This allows us to control for unobservable individual factors that affect employment decisions in two ways. First, when we estimate the fresh non-employment (employment) duration equation using multiple spells, we control for the fact that the second fresh spell is conditional on there being a first fresh spell through the common unobserved heterogeneity term, θ_{u2} (θ_{e2}). 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_{u2} (θ_{e2}). Second, we estimate the interrupted and fresh non-employment and employment equations jointly by allowing the unobserved heterogeneity terms in the four equations to be correlated.

An important finding is that there is little evidence of unobserved heterogeneity in the model. Given this lack of unobserved heterogeneity in our model, it is not surprising that we did not find evidence of dynamic selection bias when we estimated the equations jointly. This result is consistent with the finding in EHL that programs, like SSP, that do not remove the experimental group from the labour force are less likely to exhibit dynamic selection bias.

Our data covers 52 months after random assignment for the full program and control groups and this allows us to determine the short-run and long-run impact of SSP; that is, before and after the eligibility period ended. We find that SSP had a significantly positive impact on the exit rate from non-employment for program group members during the pre-incentive period, presumably because program group members were required to find a full-time job in the first twelve months after random assignment in order to receive the supplement. For the program group members who actually qualified — the take-up group — there is a significant and positive impact of receiving the supplement on the probability of exiting fresh spells of non-employment by finding a new job when we do not control for the endogeneity of the take-up variable. Once we account for this endogeneity, the impact is positive but only one-third the magnitude and no longer significant. This result highlights the need to control for the non-random selection into the take-up group. Once the 36-month period of supplement eligibility ends, there is no significant impact on the probability of exiting non-employment.

We find that the receipt of the SSP supplement by take-up group members had a negative and significant impact on the likelihood of their leaving a full-time job. We also find evidence of a negative and significant impact on the exit rate from employment once the supplement payments end. In summary, we find some evidence that SSP has a long-term impact on the employment behaviour of the take-up group.

We use the econometric results to simulate the employment rates for the take-up and Control_13 groups. The difference in these simulated employment rates in month 52 is our unbiased estimated of the effect of the treatment on the treated (TOT). We estimate TOT to be 3.6 percentage points. This is 5 percentage points higher than the -1.4 percentage point unadjusted estimate of TOT using the actual data. Thus, we find that the unadjusted estimate of TOT from the actual data can significantly distort the estimate of the TOT. The use of the econometric model developed in this paper is necessary to get a more accurate estimate of this important experimental effect.

What do these results imply about the effectiveness of SSP to promote long-term movement away from IA? The observed experimental treatment effect of 2.2 percentage points implies that program group members are employed at roughly a 10% higher rate at the end of the follow-up period compared to the control group. Further, our simulated TOT of 3.6 percentage points implies a roughly 10% higher employment rate for the take-up group relative to those control group members who found full-time work in the first 13 months. These differences do not seem large enough to claim that SSP had a substantial impact on the long term behaviour of the program group.

That said, two reasons might underlie the 10% higher employment rate of the program group. First, the higher employment rates of the program group during the pre-incentive and incentive periods implied more time working which might, in turn, have fostered an increased attachment to work. If so, this could have led to a higher employment rate for the program group even after the supplement ended. Gottschalk (2005) analyzed the 36-month SSP data and he found that increased work led to significantly more positive attitudes about work. This was particularly true for younger workers. Second, the increased work experience of the program group could have led to relatively higher wages compared to the control group. Card, Michalopoulos, and Robbins (2001) and Zabel, Schwartz, and Donald (2005) both find little evidence of relative wage progression for the "non-incentivized" group; those members of the take-up group who would have found full-time work in the first 13 months even without the SSP incentive. Unlike Card, Michalopoulos, and Robbins, Zabel, Schwartz, and Donald find relative wage progression on the order of 10 percentage points for the "incentivized" group; those members of the take-up

group who would not have found full-time work in the first 13 months without the SSP incentive. An increase in

wages for the incentivized group might have contributed to the higher program group employment rates.

Appendix
Maximum Likelihood Estimator for the Model of Employment and Non-employment Durations

In this Appendix, we develop the maximum likelihood estimator of the model of employment and non-employment durations as given by equations (1) - (5) in the paper. This method is an extension of the method for dealing with an endogenous dummy variable in a Probit model of discrete choice as discussed in Wooldridge (2002, Section 15.7.3). As in that model we want to allow for possible correlation between the residual in (5), relating to unobserved factors that determine $\mathrm{WFT13}_i^*$, and the residuals in equations (3) and (4) – note that in those equations the incentive variable for the program group is endogenous since it is formed using interactions of the variable $\mathrm{WFT13}_i$ and other variables. To deal with endogeneity we derive the likelihood conditional on ε, the error term in equation (5). This means that equations (3) and (4) become respectively,

$$
\begin{aligned}
y_{u2idt}^* = {} & \beta_{0u} + \mathrm{X}_{it}\beta_{1u} + \beta_{2u}\mathrm{INCENTIVE}_{it} + \beta_{3u}\mathrm{POST\text{-}INCENTIVE}_{it} \\
& + h_u(d,\beta_{6u}) + \theta_{u2i} + \rho_{u2}\mathrm{PROGRAM}_i \cdot \varepsilon_i + \varepsilon_{u2it}^*
\end{aligned}
\tag{3'}
$$

and,

$$
\begin{aligned}
y_{e2idt}^* = {} & \beta_{0e} + \mathrm{X}_{it}\beta_{1e} + \beta_{2e}\mathrm{INCENTIVE}_{it} + \beta_{3e}\mathrm{POST\text{-}INCENTIVE}_{it} \\
& + h_e(d,\beta_{6e}) + \theta_{e2i} + \rho_{e2}\mathrm{PROGRAM}_i \cdot \varepsilon_i + \varepsilon_{e2it}^*
\end{aligned}
\tag{4'}
$$

Note that ε only enters equations (3') and (4') for program group members. For those observations ε_{u2it}^* (respectively, ε_{e2it}^*) is normally distributed with mean zero and variance $1-\rho_{u2}^2$ (respectively $1-\rho_{e2}^2$). Thus, for instance, for completed fresh non-employment spells of duration t_u the contribution to the likelihood will be,

$$
(1-\Phi(\mu_{u2it_{u2}}^*))\prod_{t=1}^{t_{u2}-1}\Phi(\mu_{u2it}^*)
$$

where,

$$
\begin{aligned}
\mu_{u2it}^* = {} & \left(1-\rho_{u2}^2 PROGRAM_i\right)^{-1} \cdot \\
& \begin{pmatrix} \beta_{0u} + \mathrm{X}_{it}\beta_{1u} + \beta_{2u}\mathrm{INCENTIVE}_{it} + \beta_{3u}\mathrm{POST\text{-}INCENTIVE}_{it} \\ + h_u(d,\beta_{6u}) + \theta_{u2i} + \rho_{u2}\mathrm{PROGRAM}_i \cdot \varepsilon_i \end{pmatrix}
\end{aligned}
$$

The contribution for a censored spell will be,

$$
\prod_{t=1}^{t_{u2}}\Phi(\mu_{u2it}^*)
$$

Similar contributions can be found for fresh employment spells. This then gives a complete likelihood function conditional on the individual heterogeneity terms as well as the error term ε_i. The heterogeneity terms are integrated out by summing across the two points of support and then we integrate the resulting likelihood over the distribution of ε_i -- this is only done for observations for program group members. For observations where $\mathrm{WFT13}_i = 1$ we integrate over the normal density truncated to the region where $\mathrm{WFT13}_i^* > 0$ which is the interval $(\pi_0 + \mathrm{X}_i\pi_1 + \pi_2 PROGRAM_i, \infty)$. Similarly, we integrate over the complement of this interval for observations for which $\mathrm{WFT13}_i = 0$. Since these integrals are one dimensional they can be computed accurately using Gaussian quadrature.

We then combine the likelihood so constructed with the likelihood function for the variable $WFT13_i$ and our estimates that take account of both heterogeneity and endogeneity are obtained by maximizing this complete likelihood function over all parameters. Identification is achieved provided that the coefficient on PROGRAM in the model for $WFT13_i$ is not zero. This holds since, although these variables also appear to enter the equations determining the fresh non-employment and employment spells, they only do so through their interaction with $WFT13_i$ (and other variables relating to when the incentive is still in place and when it no longer applies).

References

Baker, M. and A. Melino. 2000. "Duration Dependence and Nonparametric Heterogeneity, A Monte Carlo Study," *Journal of Econometrics*, 96: 357-393.

Card, D. and D.R. Hyslop. 2005. "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers." SRDC Working Paper, 05-02.

Card, D., C. Michalopoulos, and P.K. Robbins. 2001. "The Limits to Wage Growth: Measuring the Growth Rate of Wages for Recent Welfare Leavers." NBER Working Paper 8444.

Eberwein, C., J.C. Ham, and R.J. LaLonde. 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.

Gottschalk, P. 2005. "Can Work Alter Welfare Recipients' Beliefs?" *Journal of Policy Analysis and Management*, 24: 485-498.

Ham, J.C. and R.J. Lalonde. 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. and M.L.Moeschberger 1997. *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer

Lin, W., P.K. Robins, D. Card, K. Harknett and S. Lui-Gurr. 1998. When Financial Incentives Encourage Work: Complete 18-Month Finding from the Self-Sufficiency Project. Ottawa:SRDC.

Meyer, B.D. 1990. "Non-employment Insurance and Non-employment 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., D. Card, L. Gennetian, K. Harknett, and P.K. Robins. 2000. The Self-Sufficiency Project at 36 Months: Effects of a Financial Work Incentive on Employment and Income. Ottawa: SRDC.

Michalopoulos, C., D. Tattrie, C. Miller, P.K. Robins, P. Morris, D. Gyarmati, C. Redcross, K. Foley, and R. Ford. 2002. Making Work Pay: Final Report on the Self-Sufficiency Project for Long-Term Welfare Recipients. Ottawa: Social Research and Demonstration Corporation.

Morris, P. and C. Michalopoulos. 2000. The Self-Sufficiency Project at 36 Months: Effects on Children of a Program that Increases Parental Employment and Income. Ottawa: SRDC.

Ondrich, J. and S.E. Rhody, 1999. "Multiple Spells in the Prentice-Glockler-Meyer Likelihood with Unobserved Heterogeneity," *Economics Letters*, 63: 139-144.

Wooldridge, J.E., 2002. *Econometric Analysis of Cross Section and Panel Data, Cambridge, MA, MIT Press.*

Zabel, J., S. Schwartz, and S. Donald. 2005. "An Analysis of the Impact of SSP on Wages." Mimeo.

Table 1a – Variable Descriptions

Variables that do not change over time

PROGRAM	1 if program group member; 0 if control group member
TAKE-UP	1 if program group member who found work in qualifying period (13 months following baseline interview; 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 4 years; 0 otherwise
LESS THAN HS	1 if respondent has less than a high school education at baseline; 0 otherwise
NEVER MARRIED	1 if respondent is single, never married at baseline; 0 otherwise
MONTHS IA	Number of months on IA in the 3 years prior to the baseline interview
EMOTIONAL PROB	1 if respondent had limited activity due to a long term emotional, psychological, nervous or mental health condition or problem at baseline; 0 otherwise
BENEFIT	Average monthly IA amount received in 4 quarters prior to baseline interview (in hundreds)
AGE1922	1 if respondent's age is 23-25 at baseline; 0 otherwise
AGE2325	1 if respondent's age is 23-25 at baseline; 0 otherwise
AGE2629	1 if respondent's age is 26-29 at baseline; 0 otherwise
AGE3034	1 if respondent's age is 30-34 at baseline; 0 otherwise
AGE3539	1 if respondent's age is 35-39 at baseline; 0 otherwise
AGEGE40	1 if respondent's age is more than 40 at baseline; 0 otherwise
PRE-BASE UNEMP	Number of months not employed full time in 12 months before baseline
PRE-BASE EMP	Number of months employed full time in 12 months before baseline

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
INCENTIVE	1 during the 36-month incentive period for program group members who found full-time work within qualifying period; 0 otherwise
NO-INCENTIVE	1 starting in month 14 for program group members who did not find work within qualifying period; 0 otherwise
POST-INCENTIVE	1 following the 36-month incentive period for program group members who found full-time work within qualifying period; 0 otherwise
Monthly non-employment rate	The non-employment rate for women, 15+, in the respondent's province, for each calendar month corresponding to each month after random assignment for the respondent

	Control Group	Program Group		
	Full Sample	Full Sample	Take-Up	Non-Take-Up

Table 1b

Variable Means and Standard Deviations (in parentheses)

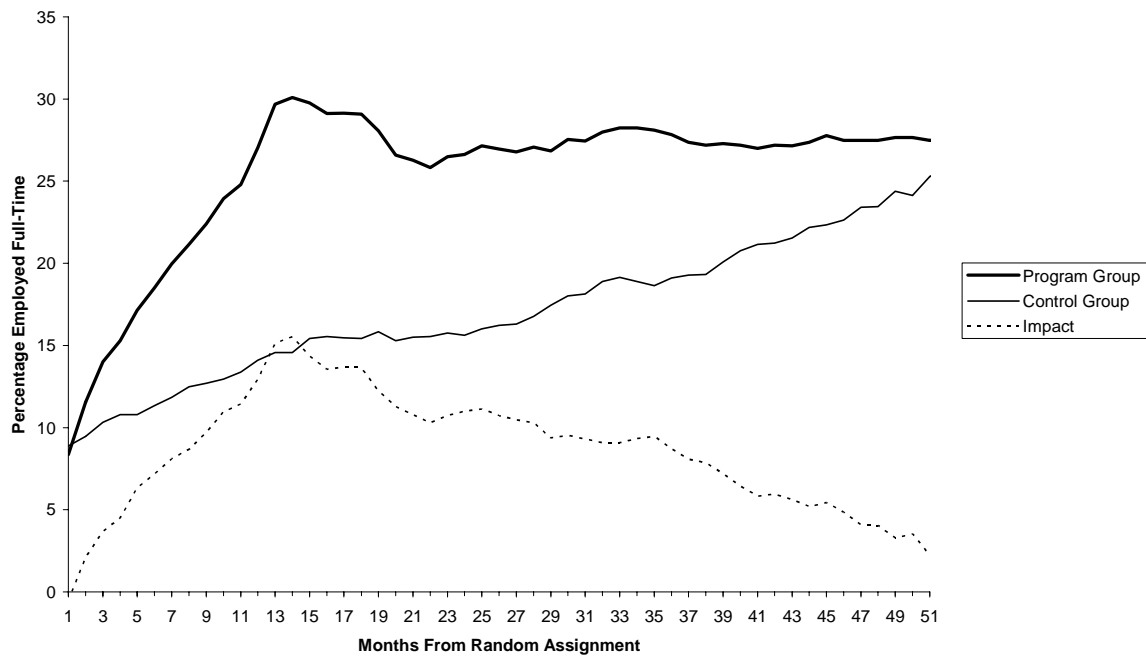|	Control Group	Program Group		
	Full Sample	Full Sample	Take-Up	Non-Take-Up
IDBC	0.52	0.53	0.50	0.54
	(0.50)	(0.50)	(0.50)	(0.50)
NKIDS	1.71	1.70	1.62*	1.75
	(0.89)	(0.84)	(0.77)	0.87
YGCHLE4	0.49	0.47**	0.47	0.47
	(0.50)	(0.50)	(0.50)	(0.50)
LESS THAN HS	0.54	0.53	0.44**	0.59*
	(0.50)	(0.50)	(0.50)	(0.49)
NEVER MARRIED	0.49	0.49	0.50	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
MONTHS IA	29.53	30.15**	29.40	30.56**
	(8.02)	(7.72)	(7.94)	(7.57)
EMOTIONAL PROB	0.07	0.08	0.06	0.098*
	(0.26)	(0.27)	(0.23)	(0.30)
BENEFIT($100s)	8.31	8.47	7.99**	8.7**
	(2.61)	(2.52)	(2.50)	(2.49)
AGE2325	0.14	0.12	0.13	0.12*
	(0.35)	(0.33)	(0.34)	(0.32)
AGE2629	0.17	0.17	0.18	0.17
	(0.38)	(0.38)	(0.39)	(0.37)
AGE3034	0.22	0.23	0.24	0.22
	(0.41)	(0.42)	(0.42)	(0.42)
AGE3539	0.17	0.17	0.17	0.18
	(0.37)	(0.38)	(0.37)	(0.38)
AGEGE40	0.17	0.17	0.14*	0.18
	(0.38)	(0.38)	(0.35)	(0.39)
Number	2354	2415	744	1511

*,** - test for the difference in population means between control group and program (take-up, non-take-up) group is significant at the 5%,1% level

Table 2: Equation for the Hazard of Exiting Non-Employment Spells

Variable	Single Equation Estimation No Corrections for Heterogeneity nor Endogeneity		Joint Estimation Correction for Heterogeneity None for Endogeneity		Joint Estimation Correction for Heterogeneity and Endogeneity	
	Coefficient (Std Error)	Partial Derivative	Coefficient (Std Error)	Partial Derivative	Coefficient (Std Error)	Partial Derivative
	(1)	(2)	(3)	(4)	(5)	(6)
	Interrupted Non-Employment Spell					
PRE-INCENTIVE (Months 1-9)	0.2579*** (0.0297)	0.0116	0.2593*** (0.0298)	0.0116	0.2563*** (0.0568)	0 .0115
PRE-INCENTIVE (Months 10-13)	0.4741*** (0.0339)	0.0268	0.4762*** (0.0347)	0.0269	0.4763*** (0.0649)	0.0270
NO-INCENTIVE	-0.0001 (0.0226)	0.000003	0.0024 (0.0235)	0.0001	0.0022 (0.0238)	0.0001
Number of Spells	4422		4422		4422	
	Fresh Non-Employment Spells					
INCENTIVE (Months 2-8)	0.1278** (0.0637)	0.0117	0.1285** (0.0642)	0.0116	0.0484 (0.0793)	0.0041
INCENTIVE (Months 9-36)	0.1224*** (0.0256)	0.0112	0.1228*** (0.0260)	0.0110	0.0464 (0.0525)	0.0040
POST-INCENTIVE (Months 1-4)	0.1023* (0.0578)	0.0092	0.1040 (0.0581)	0.0092	0.0293 (0.0734)	0.0025
POST-INCENTIVE (Months 5-8)	-0.0557 (0.0736)	-0.0044	-0.0569 (0.0738)	-0.0044	-0.1296 (0.0866)	-0.0095
POST-INCENTIVE (Months 9-12)	0.0534 (0.0784)	0.0046	0.0575 (0.0789)	0.0049	0.0142 (0.0890)	-0.0012
Number of Spells	1045		1045		1045	
	Interrupted Employment Spell					
INCENTIVE (Months 1-8)	-0.1881** (0.0823)	-0.0246	-0.1887** (0.0827)	-0.0247	-0.1915** (0.0833)	-0.0251
INCENTIVE (Months 9-16)	-0.0759 (0.1245)	-0.0107	-0.0783 (0.1251)	-0.0110	-0.0794 (0.1261)	-0.0112
INCENTIVE (Months 17-32)	0.1052 (0.1083)	0.0167	0.1061 (0.1090)	0.0169	0.1045 (0.1093)	0.0166
INCENTIVE (Months 33-36)	-0.3983 (0.2976)	-0.0451	-0.3871 (0.3004)	-0.0443	-0.3653 (0.3035)	-0.0426
POST-INCENTIVE (Months 1-4)	0.0267 (0.2282)	0.0040	0.0225 (0.2298)	0.0034	-0.0152 (0.2300)	0.0023
POST-INCENTIVE (Months 5-8)	-0.1247 (0.2930)	-0.0170	-0.1278 (0.2950)	-0.0174	-0.1336 (0.2954)	-0.0182
POST-INCENTIVE (Months 9-12)	0.1312 (0.2582)	0.0211	0.1329 (0.2600)	0.0215	0.1354 (0.2611)	0.0220
Number of Spells	347		347		347	
	Fresh Employment Spells					
INCENTIVE (Months 1-8)	-0.1772*** (0.0336)	-0.0177	-0.1764*** (0.0340)	-0.0175	-0.1726*** (0.0594)	-0.0171
INCENTIVE (Months 9-16)	-0.1200*** (0.0382)	-0.0125	-0.1223*** (0.0386)	-0.0126	-0.1177* (0.0610)	-0.0122
INCENTIVE (Months 17-32)	-0.0250 (0.0299)	-0.0028	-0.031- (0.0318)	-0.0034	-0.0263 (0.0558)	-0.0029
INCENTIVE (Months 33-36)	0.0365	0.0043	0.0313	0.0036	0.0348	0.0040

	(0.0553)		(0.0568)		(0.0721)	
POST INCENTIVE (Months 1-4)	0.0368	0.0043	0.0326	0.0038	0.0357	0.0042
	(0.0571)		(0.0586)		(0.0732)	
POST-INCENTIVE (Months 5-8)	0.0047	0.0005	-0.0008	-0.0001	0.0022	0.0002
	(0.0640)		(0.0656)		(0.0793)	
POST-INCENTIVE (Months 9-12)	-0.2675***	-0.0249	-0.2732***	-0.0251	-0.2729***	-0.0251
	(0.0949)		(0.0962)		(0.1052)	
			Model Parameters			
μ_1					-0.9386***	
					(0.2501)	
μ_2					-1.0532**	
					(0.4464)	
γ					0.2907	
					(11.254)	
α_{1u2}					-0.1511	
					(0.1495)	
α_{1e1}					-0.1599	
					(2.7025)	
α_{1e2}					-1.3768	
					(2.6845)	
α_{2e}					1.0202	
					(2.7117)	
Number of Spells	1142		1142		1142	
Likelihood Function	-27371.338		-27370.732		-5774.022	
*,**,*** - significant at the 10, 5, and 1 percent levels						

Figure 1: Percentage Employed Full-Time, by Month From Random Assignment, in the SSP Recipient Study

Figure 2
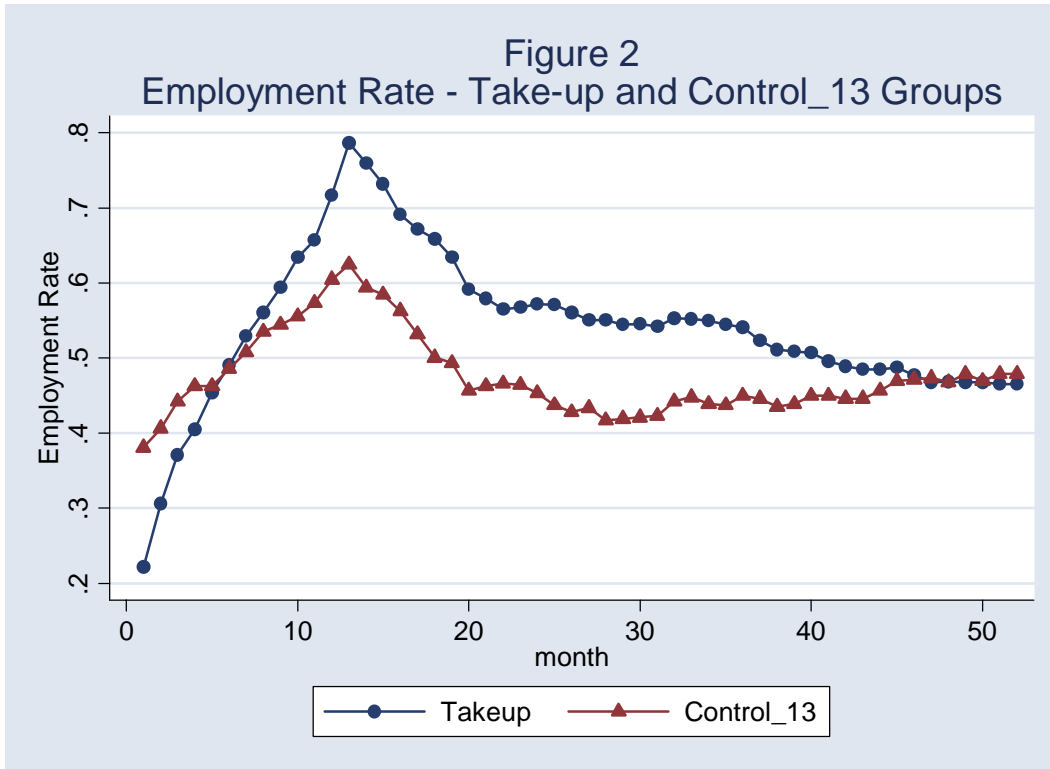Employment Rate - Take-up and Control_13 Groups

Figure 3: Empirical Hazard Rates for First Spell of Unemployment: Program and Control Groups
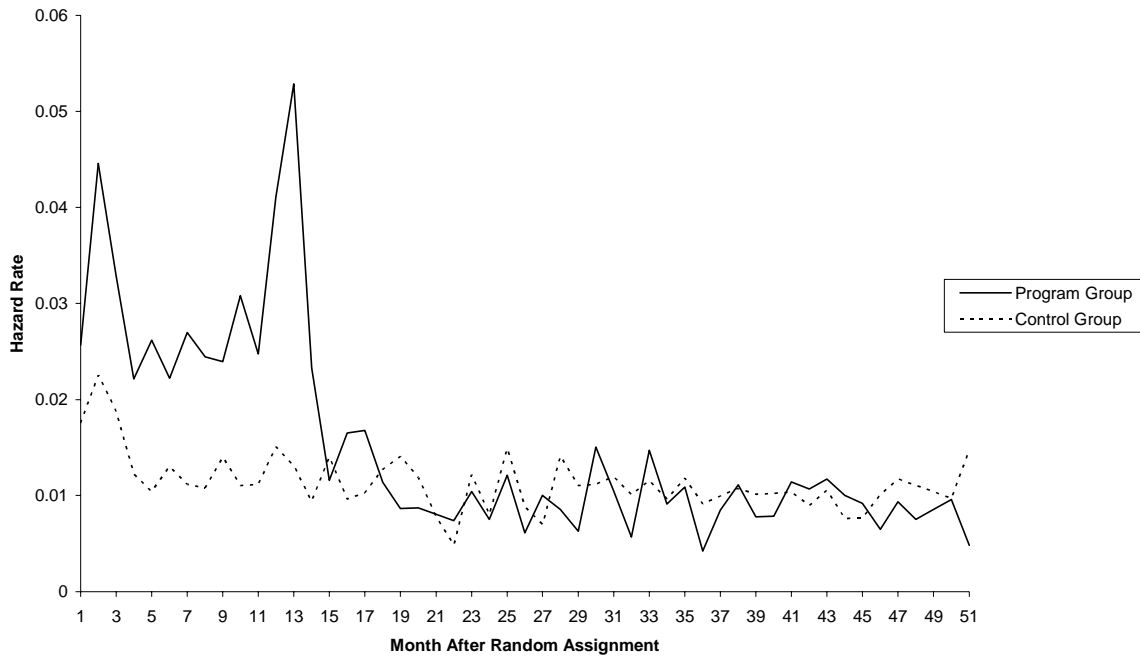
Figure 4: Empirical Hazard Rates for First Spell of Employment: Program and Control Groups

Figure 5
Simulation Results
Employment Rate - Takeup and Control_13 Groups