Explaining and Forecasting Results of
The Self-Sufficiency Project

Christopher Ferrall
Department of Economics
Queen's University

December 19, 2008

Abstract:
This paper models the Self-Sufficiency Project ( SSP), a controlled randomized experiment concerning welfare. The model of household behavior includes stochastic labor market skill, job opportunities, and value of non-labor market time. All the variation within and between treatment groups, jurisdictions (provinces), demographic groups, and sub-experiments is derived from four underlying sources: policy variation, endogenous selection into the experimental samples, the SSP treatments themselves, and different mixtures over 4 underlying types. Using the variation within the treatment group is quantitatively important for identifying the complex model. Parameters are estimated precisely using efficient GMM. Variation within the treatment group is much more important for identification than either variation within the control group or between treatment and control groups. The model tracks the primary moments well, in and out of sample, except for under-estimating trends in the sample of new applicants. Predictions from counterfactual experiments suggest that details of the SSP’s design are critical for interpretation of results. And it appears that the small SSP+ treatment may have longer lasting impacts than the in-sample impacts suggest.

JEL Classification: I3, C9, J0, C5

Keywords:
Dynamic Household Behavior, Welfare Policy, Controlled Experiments, GMM

Data access was based on a research contract with Social Research and Demonstration Corporation (SRDC). Research support from SSHRC and computing resources from HPCVL, Westgrid and SHARCNET are gratefully acknowledged. Full acknowledgements and supporting material are available at the link in the Appendix.
I. Introduction

The Self-Sufficiency Project (SSP) was a controlled randomized experiment performed in two Canadian provinces designed to study whether long-term recipients of income assistance (i.e. welfare) respond to earnings subsidies. The main SSP treatment group, single parents on income assistance (IA) for at least one year, was offered a large supplement to earnings if and when a full-time job was acquired within one year (and the parent went off IA). The premise of the SSP, that it would induce sustained full-time employment which would generate skills and thereby substantial wage gains, was partly fulfilled. Approximately 35 percent of the treatment group qualified for the supplement and at the peak that group had a 100 percent increase in full-time work and a 70 percent increase in earnings relative to the controls. However, most of the impact disappeared soon after the supplement expired (Michalopoulos et al. 2002).\(^1\) The hoped-for self-sufficiency through endogenous wage gains failed to appear. Despite this, the careful and ambitious design of the experiment provides a unique opportunity to study labor market dynamics among low-income households.

This paper describes a forward-looking model of single parent households that predicts labor market outcomes before, during, and after the SSP experiment. The SSP creates exogenous variation in budget constraints and expected future income that is used to identify a rich model. The model is estimated using GMM on the first 36 months of the (roughly) 54 months of experimental data. Results from the final 18 months are compared to out-of-sample forecasts from the model, and counter-factual experiments are computed.

The SSP incentive scheme attempted to forge a path out of the “welfare trap” for single-parent households. The depth of the welfare trap is difficult to measure because it depends on how a person’s current skill relates to their labor market history. An inference

---

\(^1\) Most of the work analyzing the SSP is discussed and summarized by the authors themselves in SRDC 2006.
is required on what wages would be now if the parent had (counter-factually) worked more and more steadily in the past. Papers that address dynamic in welfare policy using non-experimental data include Miller and Sanders (1997), Swan (1998), Kennan and Walker (2003), Keane and Wolpin (2002), Fang and Keane (2004) and Fang and Silverman (2004). This paper provides new evidence on what labor market policies can do to affect welfare dependency by confronting a model that captures many elements of the previous papers with the large and complex variation in static and dynamic incentives created by the SSP. Exploiting this variation yields identification of the time-varying and heterogeneous effects of income and opportunity cost that underly patterns of welfare receipt. The explanation for the SSP that the estimated model provides is a combination of several factors that are not usually considered jointly. Four aspects of the data drive the explanation.

1. Going on and off Income Assistance. The welfare trap is not as deep for some parents as others. As seen from control group data presented below, a large fraction of households on long-term IA eventually leave without incentives. This aspect of the data requires a churning in the status of households. Some churning comes from exogenous layoffs, but households also classify many separations as quits. The model includes a reduced-form specification for non-linear costs of labor market time. The curvature captures the effects of child care arrangements, commuting and other factors affecting the income-leisure tradeoff for a single parent. This leisure-income tradeoff is subject to persistent shifts which can induce the parent to change work hours or quit altogether and move back to IA. Second, changes in IA status also occur without changes in working status. Citing early U.S. research, Moffitt (1992) concludes, "most exits from AFDC are not a result of an increase in earnings by the female head, but are instead the result of a change in marital status that results in the loss of AFDC eligibility." To capture this type of transition, households have an outside support option that, if accepted, makes it ineligible for IA. The unobserved value of this option also includes any stigma associated
with accepting IA. The value of the outside support option shifts over time, which explains changes in IA status with no change in labor market status.

2. Gradual Response to Static Incentives. The SSP treatment is a static short-term incentive designed to alter long-term behavior. That is, it subsidizes labor market attachment so that over time skills develop, earnings rise, and work becomes more attractive than government transfers. The response to the SSP was gradual. Given a year to qualify for the SSP, households did so steadily. Without labor market frictions, a short-term response would happen immediately if it were to happen at all. Thus, like many explanations of the labor market, frictions in the form of costly job search play an important part of the estimated model. Non-working households given an incentive to work full-time must search for a suitable job. Some succeed, some do not. Households who were on IA and working part-time at the start of the treatment would not be subject to search friction. But many of these households did not immediately qualify by increasing their hours. Thus, the explanation includes work hours that are not fully flexible. Households search for jobs based on wage and jobs that allow full-time work while balancing the value of non-market time covered in point 1 above.

3. Lack of Dynamic Response. After the gradual short-term response, self-sufficiency failed to happen. On average, wage growth is observed in the sample, but the SSP failed to induce a substantial difference in growth between treatments and controls. Once the subsidy ended the treatment groups fell back to rates of receiving IA similar to the control groups (which was itself evolving). The estimates capture the short-term success and long-term failure of the SSP to alter behavior. If a model were used that automatically relates wage growth to employment it could not combine a strong short-term response in working with no long term response in wages. The explanation here adds a third important element: heterogeneity in the growth potential of jobs. In the estimated model, households search for jobs characterized by works hours and both an initial offer and a rate at which wages respond to skill accumulation. Households respond to the short-
term incentives by finding and accepting full-time jobs but as it turns out they do not discriminate enough between "dead-end" jobs and "stepping-stone" jobs. Legislated minimum wages play an important role in whether skills are tied to wage growth on particular jobs.

4. Expectations and Heterogeneity. So far the explanation has related broad patterns in the data with key elements of the model as if the patterns were independent of each other. In the data the average effects described so far are surrounded by complex distributions in outcomes. Heterogeneity across households will also be an important element of the story to match these distributions. In addition, forward-looking households do not view these responses as separate. The movements on and off IA outside the experiment, the gradual short-term response, and the minimal long-term response are interconnected. The glue that combines them is the degree of foresight among households, captured as usual by a time discount factor. A myopic household facing short-run incentives to work full time will do so, even if it may take time to find such a job, but it has weak incentives to hold out for a good job that will pull them out of the welfare trap years in the future. The SSP provides strong experimental evidence on the patience of single-parent households. This experimental variation is used to estimate a distribution of discount factors in the population. The four-fold explanation concludes with the result that a large segment of the population sampled (endogenously) in the SSP is myopic. The road to self-sufficiency paved by the SSP is too long for them to take account of the post-treatment exit from the welfare trap.

One of the motivations for developing this complex explanation of the SSP is to apply it outside the experiment. Subsequent research will indeed use the estimated model to study a larger set of counter-factual policies and experimental designs. But given the number of elements of the explanation and the richness of the SSP environment itself, the applications reported in this paper focus on the experiment itself. These forecasts include actual forecasts of outcomes in the data beyond the interval available for estimation. In
addition, several experiments are run at the estimated parameters to ask whether the
impacts observed in the SSP are robust to modest changes in the design of the experiment.
In some dimensions the parameters of the experiment do not have a great effect on the
model outcomes. In other dimensions the results are found to be sensitive to the details.
The lesson is that it appears unwarranted to extrapolate the experimental impact to related
welfare policies without reference to a model of behavioral response to the experiment.

Another use of the model considers the value of using the treatment outcomes for
estimation versus the alternative of using the experimental variation to validate a model
estimated on the control groups only. To quantify the tradeoff, standard errors are re-
computed within experimental groups. For many parameters of the model the standard
errors explode when based only on the control samples. This shows that in this case exper-
imental variation helps identify a richer model than identified by control variation alone.
The richer model, if validated in other ways, may then be applicable in environments
farther afield than a weaker model. Somewhat surprisingly, re-scaled standard errors are
smaller when based only on the treatment groups. The implication of this outcome is
developed below.

II. Overview of the Experiment and the Analysis

This section describes the SSP experiment, the model, and the estimation procedure
with minimal notation. Only the details that are pertinent to this studied are described in
this section. The reader interested in data quality, the randomization process and other
details is referred to the many papers and reports based on the SSP and summarized in

II.A Static Elements of IA and the SSP Treatment

Low-income single parents face a number of constraints when trying to establish
a long-term attachment to the labor market. The well-known static tradeoffs between
income and leisure created by the welfare are illustrated in Figure 1. A household is
eligible for income assistance in the amount $IAB as long as they do not take forms of non-government outside support that make them ineligible. The value to the single-parent of this non-government, non-labor market transfer is denoted OS. For example, if the parent decides to cohabitate with someone with significant income this reduces the value of IAB substantially, possibly to $0. The model is separable in income and leisure, so other forms of outside income that do not affect eligibility for IA can be ignored. Changes in OS correspond to point 1 of the introduction. Throughout the 1990s earnings up to $200 per month (denoted SA on the figure) could be set aside without a reduction in benefits. Thereafter benefits were replaced by earnings (in the figure 1-for-1). An indifference curve with optimal point A illustrates the disincentive to work more than part-time under this budget. The curvature of the indifference curve is captured by the non-linearities mentioned in point 1 of the introduction.

**Figure 1** also displays how the main SSP treatment alters the budget. The parent always has the option of receiving IA, but to get the earnings supplement they cannot be on IA that month. Thus, SSP income starts at OS and has a slope equal to the wage until the full-time work requirement is met. At this point, earnings are supplemented to be half the difference between actual earnings and 3.9 times full-time earnings on a minimum wage job. The resulting discontinuous change in the budget constraint is the solid red line on Figure 1. As drawn this particular parent is indifferent between staying on IA at point A (which is still a feasible choice during treatment) and working full time with the supplement at point B. This is the short-term response to the incentives discussed in point 2 of the introduction.

Moving from point A to point B for many periods could shift the budget line (the wage rate) if the induced shift to full-time work generates wage growth. This is shown as a mixed blue line in **Figure 1**. After treatment ends the supplement budget disappears but the wage rate is now shifted and a new optimal choice is C, which has the parent working part-time at a high wage rate and not receiving IA. In this case the temporary SSP treatment
Figure 1. Earnings, Income, Work Hours under IA and the SSP

IA payments reduced 1-for-1 by earnings above an amount $SA combined with a low-wage results in limited work hours (point A). Dynamically, lack of work experience keeps wages low and results in a trap. The SSP treatment attempts to smooth the path out of the trap by encouraging full-time work (point B) for a limited three-year period. If during this treatment wages respond to steady work then the budget shifts endogenously. Once the treatment ends the induced wage growth makes IA less attractive to labor market earnings (point C).

has eliminated the welfare trap through employment-induced wage growth. But if the wage on the job underlying point B does not respond to experience then the shift to the mixed line budget does not occur by the end of treatment. Instead the parent moves back to A. Whether wage growth occurs among households who qualify for the SSP depends on heterogeneity in wage growth and patience (points 3 and 4 in the introduction).

---

2 Keane and Moffitt (1998) use an estimated model of income maintenance programs in the U.S. to predict that reforms without such a full-time work requirement would significantly increase total transfer payments to poor households and shift many away from full-time work.
II.B Sample Selection and Program of Treatment

The SSP experiment was not as simple as randomly selecting single-parent households, changing their static budget constraints in Figure 1 for one period and then measuring the impact relative to control households. This paper models all the details of how the SSP sample was selected and how treatment progressed for up to 60 months. There were several experiments within the SSP. First, two distinct samples were drawn and studied in the SSP: the Recipient Study and the Applicant Study. The Recipient Study was carried out in both British Columbia (BC) and New Brunswick (NB). The Applicant Study was conducted only in British Columbia. The SSP treatment depends on minimum wages which itself differs across provinces. Thus, there are 3 treatment groups and 3 control groups covering these studies (BC Recipient control and treatment, BC Applicant control and treatment, NB Recipient treatment and control). Second, within the Recipient Study there was a small separate treatment called SSP Plus (or SSP+) conducted only in New Brunswick. The SSP+ treatment group shared the control group with the Recipient Study. So this brings to 7 the distinct groups. Finally, because IAB (shown in Figure 1) depends on the number of children, this paper controls for whether there are one or two-or-more children in the household. That brings to $7 \times 2 = 14$ the total number of distinct groups consider in the analysis below.

Except for the SSP+ sample, the SSP has a single program of treatment with three phases. The treatment illustrated in Figure 1 is just one of those phases. However, with forward-looking households that were fully informed about the experiment each phase alters the reality of their situation even when no supplement is involved. The estimation procedure fully exploits this variation to identify a rich model.

Figure 2 illustrates the phases of treatment in the SSP. Begin with the whole population of single-parent households (in a given province and with either 1 or 2+ children). Since the model is stochastic and tracks many state variables a given household may or not be eligible in a given month. To be eligible for the Recipient Study, the main sample,
the household must have been on IA for twelve or more months, which graphically is represented as a subset of full set of households.\(^3\) Among households on IA for 12+ months, a sample was drawn and then randomized into treatment and control groups. One of the key aspects of the estimated model is that this selection criterion is endogenous to the model. Thus the estimated parameters pertain to the broader population not just the subset eligible for random selection. The treatment group was not just provided the supplement. They had twelve months to qualify for the supplement by accepting a full-

\(^3\) This is the condition in the model. It simplifies the actual condition that the Recipient Study selected single-parent households that had been on IA for 12 out of the last 13 months.
time job and going off IA. If they could not satisfy this condition within twelve months they failed to qualify and treatment ended. Figure 2 puts the treatment group on a timeline at month 0, the start of the twelve month qualifying phase. If the parent qualified they moved to the eligible phase. This phase lasted 36 months.4

At any point in this time the household can go back on welfare and/or stop working. They receive no supplement in such months and the 36-month clock continues to count down. But once they again work full time and go off IA they get the supplement again. Thus, the full complicated SSP budget constraint in Figure 1 applies during the eligibility phase, including the segment that involves IAB. After 36 months, this eligibility phase ends and the treated sample returns to the status quo. The separate SSP+ treatment followed the same schedule as the Applicant Sample which is why there is no separate control group for them. But in addition to the supplement illustrated in Figure 1 these households were given job search counseling and other support for efforts to go off IA.

The Applicant Sample is based on a completely different selection criterion. It was drawn from households that had not received IA for at least six months and were now applying for a period of IA. Households meeting this criterion were sampled and then randomized into treatment and control. This is also graphically represented as a subset of all possible household states that does not overlap with the Recipient Sample subset. The estimation procedure treats both the Applicant and Recipient samples as parts of one population differing only on their behavior in the year leading up to randomization.

The Applicant Sample was drawn for a particular purpose: to study whether offering the SSP to long-term IA recipients might induce short-term recipients to stay on IA, sometimes called the entry effect of the program. To mimic this within an experiment, the Applicant households were told (truthfully) that if their spell on IA lasted one year they

---

4 Actually, each twelve-month period of eligibility was a distinct phase because in at most two months per year the recipient could receive the supplement when hours fell below the full-time requirement (to smooth within-month interruptions in employment). This is not included in model for simplicity, but in the notation the 36-month eligibility phase is still broken into three phases (3-5). In this way all phases of treatment last at most 12 months.
would then be able to qualify for the SSP. Thus, in Figure 2 we see that this sample was assigned to phase 1: waiting for a chance to qualify. Because the Recipient Sample was considered the main sample, we follow the SSP reports and label their entry point as time 0. This means that in a unified program of treatment the Recipient Sample starts at time -11. If households in this sample went off IA before time 0 was reached their treatment ended.

**II.C Innovations for Analyzing Welfare and Experimental Data**

In many respects the analysis is like many empirical applications of dynamic programming. Namely: the single parent is either unemployed or employed coming into a period (a month). If employed they choose hours of work or to quit the job (set hours to 0). If unemployed they can choose to actively search for a job or not. Accepted job offers start the next month. Each period the household also chooses whether to accept income assistance or not. If they do not accept IA they forgo any stigma felt while on welfare and can accept non-government transfers that preclude or severely reduce their level of benefits. These choices help determine the state of the household next period which also depends on stochastic elements of the environment. These elements include the type of job offers received and exogenous job loss (layoff). The model is parameterized and includes permanent unobserved heterogeneity (in parameters of the dynamic programming problem) and transitory unobserved heterogeneity (hidden variation in state variables that affect the household’s utility). The model is estimated using GMM including an efficient second-stage weighting matrix.

Now consider the elements that are not as standard or are new to the literature. First, as emphasized in the introduction, the problem the SSP attempts to address, the welfare trap, is not simplistic in the estimated model. The single parent is buffeted by shocks within the household (in terms of opportunity costs of labor market time), within relationships (in terms of ‘stigma’ and the value of private transfers that can substitute for IA), and
within the labor market (in terms of exogenous job loss, costly job search and stochastic skill development). Jobs are not simply wage offers. Parents search for jobs that differ in hours flexibility, initial wages, and wage growth. Skills both accumulate and depreciate depending on labor market activity, so that inactivity reduces wage offers and steady work increases them. In response to increased skill, jobs are heterogeneous in their potential for wage growth. Some jobs offer no growth based on skill accumulation. Further, minimum wage legislation alters the growth path of wages based on skills and job quality. Thus, there are multiple reasons why households end up on IA for long periods of time even while some apparently similar households come and go. The SSP treatment is not a silver bullet. While it mechanically alters the non-convex budget constraint in Figure 1 it does not ensure that single parents who take it up step onto a ladder of growing wages and self-sufficiency. The estimated model uses the experimental outcomes to parcel out how big a role each element of the model plays.

The approach to unobserved heterogeneity, discussed more fully in Ferrall (2003), differs from more common approaches. It is related to previous specifications used by Ferrall (1997) and Eckstein and Wolpin (1999). Like many applications, a finite mixture over types is used, but as with Eckstein and Wolpin types differ in many dimensions rather than in one or two parameters. In addition, as in Eckstein and Wolpin there is no interaction between observable aspects of the household that are considered exogenous and the type-specific parameters. As with Ferrall (1997) the mixture across types in the population is allowed to differ across groups defined by lagged exogenous variables (numbers of children and province). And as with Ferrall, a type-specific discount factor is estimated.

Outside the experiment the household’s decision horizon is infinite. In part this is because decisions are month-to-month (so lifetime concerns are not of primary impor-

---

5 Of the 92 parameters Eckstein and Wolpin (1999) estimated, over half (40) were specific to each of their four types, including the type probabilities.
tance). In part this reflects the lack of complete labor market histories in the SSP. Unlike applications that use data that follow every agent from some pre-determined first period of decision making (e.g. age 16), conditions at the start of the data are endogenous. Transitions in this infinite horizon are made to be ergodic (all states will be visited in the future) which closes out the predictions of the model before and after the SSP.\(^6\) The GMM estimation procedure explicitly models the full process for becoming eligible for either SSP sample (Applicant or Recipient). In particular, on each evaluation of the objective the exact stationary distribution over 2304 states is computed 16 times: once for each combination of agent type (4) and each policy environment (2 provinces and 2 household sizes). Then from these distributions the entry conditions for each study are computed (not simulated), conditioning on being on IA or off IA for the appropriate number of months. The result is a fraction of the population share of each type. So the in-sample proportions are endogenously different than the population proportions. The same is true of the distribution over transitory unobserved households states. To my knowledge no previous research has used a model of forward-looking agents to both draw from the stationary distribution to deal with unobserved histories but also endogenize sample selection using the exact same selection rule as the data.

The phases of treatment are built into the behavior of the treated households. This creates valuable exogenous variation in budget constraints and expectations that are exploited in estimating the model. For example, the “waiting” and “qualifying” phases in Figure 2 do not treat household utility, only expectations about future utility. Thus, if households were not forward looking, data from these phases would be identical in the treatment and control groups.\(^7\) This variation reveals how patient the households are, allowing

---

\(^6\) This means \textit{a priori} the SSP has no permanent effect in the sense of lasting for eternity. This may seem like a negative aspect of the model, particularly when one is used to considering models of behavior counting back from death. But nothing stops the half-life of the treatment effects from being, say, 40 years. This would be the stationary equivalent to a permanent lifetime impact.

\(^7\) Card and Hyslop (2005) discuss a stylized model of job search which is a very special case of the one estimated here, but all of their empirical results are based on a non-structural model of transition probabilities which are not the reduced-form of their stylized theoretical model. (They state, "Our model of
type-specific discount factors to be estimated. But the amount of exogenous variation produced by the SSP, and useful for estimation if the rules are built into the agent problem, is so large that it is difficult to specify all the ways it is useful for identifying parameters. How much variation? The bottom of Table 1 summarizes the model size. The rules of the SSP expand the state space from 2304 elements outside the experiments to 138,240 states within the experiment for a given household. Together with the combinations of agent types, policy regimes and treatment groups the total number of states is 6,672,384. When combined with the 12 feasible actions the number of distinct outcomes is 80,068,608.

III. The Environment

III.A States and Parameters

In the model a household’s situation each month outside the experiment is described by nine endogenous state variables:

$$\theta_{end} = (l \ p \ n \ x \ b \ s \ h \ d \ k).$$

(1)

The state variables are listed in Table 1 and described in more detail in the Appendix. Briefly the variables are indicators or indices for: the parent lost their previous job; the parent worked in the previous month; the earnings offer in the current job; the parent’s

SSP is not structural.” They model state transitions and do not account for how welfare and employment affect household income. They estimate non-structural parameters specific to each phase of SSP treatment whereas in the model estimated here only parameters of the household’s problem are estimated. Differences in behavior across phases is endogenous to the environment. Within phases of treatment they allow for trends in their non-structural model of transitions, including a fourth-order polynomial trend. In the model estimated here all reactions to time remaining in each phase is determined endogenously with no additional parameters. Card and Hyslop include unobserved heterogeneity in their non-structural parameters, but the distribution of heterogeneity is conditional on being eligible for random assignment (i.e. already being on welfare for twelve or more months). As emphasized in this paragraph, the model here makes eligibility for random assignment endogenous, so the estimated heterogeneity is not conditional on welfare participation, which is endogenous to the policy reform being considered. Finally, their non-structural model is rejected by reported goodness-of-fit statistics based on a small subset of the data dimensions considered here.

Identification is discussed in the Appendix, but it is a somewhat common belief that discount factors cannot be estimated well, leading many applications to fix the value a priori. Unlike many applications with infinite or very distant finite horizons, the exogenous finite deadlines created by the experimental environment drives identification of the discount factor. Wolpin (1995) provides a formal proof for this result for a simple job search model.

C. Ferrall. Explaining The SSP

III. The Environment Page 14
skill based on previous experience; the upper bound on working hours in the current job; the level of outside non-governmental support; the opportunity cost of time spent outside the household; the observed demographic group, and the parent’s unobserved type (k). The first two variables, \( l \) and \( p_b \), do not affect the household’s decisions directly and are tracked as state variables in order to match SSP results on job loss and quits.

The endogenous state vector \( \theta_{\text{end}} \) is contained in the overall state vector \( \theta \), which concatenates five sub-vectors:

\[
\theta \equiv \left( \theta_{\text{clock}} \quad \theta_{\text{exp}} \quad \theta_{\text{end}} \quad \theta_{\text{exog}} \quad \theta_{\text{pol}} \right).
\]

(2)

Each state variable belongs exclusively to one of these sub-vectors. The other sub-vectors are described as needed.

The demographic index \( d \) varies across households but not over time for a given household. Characteristics treated as demographic in the SSP model are indicators for province of residence and whether the parent has two or more children. Each of the \( D = 4 \) demographic groups has a vector of policy parameters,

\[
\Psi_p[d] \equiv (\text{IAB}_d \quad \text{SA}_d \quad \text{CB}_d \quad \text{MW}_d).
\]

(3)

The parameters are the maximum level of income assistance benefits, the income set-aside before benefits are clawed back, the claw-back rate on benefits, and full-time earnings on a minimum wage job. The parameter values were illustrated in Figure 1 and listed below in Table A.2.

The index for unobserved type, \( k \), is also fixed for a household and determines which of \( K = 4 \) vectors of exogenous parameters pertains to the household:

\[
\Gamma[k] \equiv (\Gamma_k \quad \Pi_k \quad \delta_k \quad \rho_k).
\]

(4)

Exogenous parameters that shift utility are contained in \( \Gamma_k \). Parameters that shift the evolution of state variables are contained in \( \Pi_k \). The scalar \( \delta_k \) is the discount factor, and \( \rho_k \)
controls smoothing of choice probabilities. The elements of \( \Upsilon_k \) and \( \Pi_k \) and the exact roles of \( \delta_k \) and \( \rho_k \) are described in the rest of this section.

Within demographic group \( d \) household types are distributed according to \( \Lambda[d] \), a vector with elements \( \lambda[d,k] \). For example, \( \lambda[1,2] \) is the second element of \( \Lambda[1] \) and equals the proportion of type 2 households in group 1. Only the proportions and policy parameters \( \Psi_p[d] \) vary with \( d \). For example, no province dummy appears in the earnings-offer distribution. Instead, inter-provincial differences are generated by a different mixture of types across provinces. Also, the opportunity cost of labor market time does not depend directly on the number of children. Instead, one-child households can be a different mixture across unobserved types than 2+ households.

The exogenous vector contains all the estimated parameters:

\[
\theta_{\text{exog}} = (\Lambda[1] \quad \cdots \quad \Lambda[4] \quad \Gamma[1] \quad \cdots \quad \Gamma[4]).
\]  

There are \( N = 19 \) parameters in \( \Gamma[k] \) leading to a total of \( K(D + N) = 4(4 + 18) = 88 \) exogenous parameters. Free parameters are fewer because three parameters in \( \Gamma[k] \) are constrained to be equal across type (on the presumption that they are the least likely to be identified for observed variation). Accounting for this and the fact that elements of \( \Lambda[d] \) sum to 1 results in \( 3 + 4(19 - 3) + 3(4) = 79 \) parameters estimated from the data.

**III.B Actions**

Each month the parent chooses an action vector,

\[
\alpha = (m \quad a \quad i) \in \mathbf{A}(\theta),
\]  

containing three variables: labor market hours, active job search, and acceptance of income assistance. The feasible set \( \mathbf{A}(\theta) \) imposes two restrictions. First, active job search while working is ruled out: \( m > 0 \) or \( a = 1 \), but not both. Second, the parent faces an upper bound on work hours: \( m \leq u(b) \) where \( b \) is a state variable specific to the current job. When the parent has no job, \( b = 0 \). With a part-time job they can only work less than \( b = PT < 1 \).
hours relative to full-time. PT equals the value in the SSP treatment of 75%. When holding a full-time job \( b = 1 \) the parent can chose to work fewer hours. A parent holding a job who does not work at all effectively quits and loses the option to work until a new job is found and accepted.

### III.C Outcomes and Results

The combination of an action and a state, \((\alpha, \theta)\), is referred to as an outcome. The state next month, \( \theta' \), is randomly determined by the transition \( P \{ \theta' | \alpha, \theta \} \) (fully described in the Appendix). Not all aspects of a household’s outcome can be observed by outsiders: some states and some actions are unobserved. Understanding welfare and the incentive to work involves many hidden states, including skill, job quality, and leisure-income tradeoffs. The vector of measurements made from an outcome \((\alpha, \theta)\) is denoted \( Y(\alpha, \theta) \). Only some endogenous state variables in the SSP model are directly available in \( Y(\alpha, \theta) \): \( l, p, \) and \( d \). The action variables \( i \) and \( m \) are observed but active job search \( (\alpha) \) is not. The moments drawn from the data to estimate the model are based on \( Y(\alpha, \theta) \) and are described below.

### III.D Utility

Utility equals income plus outside support minus the opportunity cost of labor market time:

\[
U(\alpha, \theta) = \text{Income}(\alpha, \theta) + \text{OS}(\alpha, \theta) - C(\alpha, \theta). \tag{7}
\]

In turn, income is the sum of earnings, income assistance payments, and SSP payments:

\[
\text{Income}(\alpha, \theta) = \text{IA}(\alpha, \theta) + \text{TrueEarn}(\alpha, \theta) + \text{SUP}(\alpha, \theta). \tag{8}
\]

As seen from Figure 1 households on IA have an incentive to hide earnings because their benefits are clawed back 1-for-1 after a small setaside each month. Underreporting of income also helps the model explain the sizeable fractions of households working while receiving IA. So in the analysis measured earnings are a fraction \( \beta \) of TrueEarn when \( i = 1 \).
The components of (8) are defined in the Appendix. The second term in (7) is the sum of non-government transfers and additional utility (in dollar equivalent) from forgoing IA:

$$OS(\alpha, \theta) = (1 - i)s\xi IAB.$$  \hspace{1cm} (9)

The transfer component of OS is support that, if accepted, disqualifies the parent for IA. Outside support varies from month to month based on the endogenous variable, $s$. When $s$ changes the parent may go off welfare and rely on other sources of support with or without any change in labor market status. A drop in $s$ may push the parent back to receiving IA. Because OS includes foregone stigma of the static form in Moffitt (1983), maximum OS is expressed as a factor of $IAB$, the maximum amount of IA the household is entitled to. The parameter $\xi$ is a positive exogenous value dependent on type $k$.

Three possible sets of feasible work hours (depending on $b$) are shown in Figure 3 as ranges along the x axis starting from the right at zero work hours ($m = 0$). The x-axis is non-market time expressed as a fraction of full-time employment. The y-axis is dollars per month, and the discrete values of $m$ are indicated by vertical lines. The value of $b$ changes from one month to the next for various reasons. A non-working parent finds a job with probability $p_j(\alpha, \theta)$, which will have an upper bound on $m$ of either PT or 1. A working parent loses a job permanently with probability $\pi_1$ and results in an upper bound of 0 next month. A working parent can quit by setting $m = 0$. A parent who quits or is laid-off can immediately engage in job search ($a = 1$), but a job offered that month begins the next month. Thus leaving or losing a job is matched to cases where the parent experiences at least one month not working. Job-to-job transitions are treated as the same job. The model attributes growth in full-time equivalent earnings between contiguous jobs as skill acquisition.

The cost of labor market time,

$$C(\alpha, \theta) = W_{max}v [m + \kappa a]^{d(k)},$$

is expressed as a fraction of $W_{max}$ maximum possible earnings (defined later). Costs
of time depend on working hours and search time when not working. Job-search is converted to work time by the exogenous parameter $\kappa$. The cost of full-time work is then $\nu W_{\text{max}}$.9

The curvature of costs is determined by $c(h)$. It shifts with $h$ as described in the Appendix and is illustrated in Figure 2 by three dotted lines. When not working the parent can choose to search actively for a job and incur cost $\nu c(d)$, which is shown on the graph along the mixed (red) line.

A shifting preference for full and part-time work hours is represented by three different costs depending on the state variable $h$, which jumps to a new value each month with probability $\pi_h$. Costs rise slowly with $m$ when, for example, children are in school and part-time work has a low opportunity cost. Costs rise quickly with $m$ when, for example, children are young or sick or part-time care arrangements break down. When the value of $h$ jumps to a new value a working parent may change hours or quit and drop out of the labor market. Either change may induce a change in welfare receipt. A non-working parent may respond to a change in $h$ by beginning or ending active search.

### III.E Skill, Job Search and Wages

Skill is expressed as a fraction of full-skill: $x \in \{1/4, 1/2, 3/4, 1\}$. From month to month $x$ either remains constant or changes by $1/4$ with a probability that depends on labor market status. While working, skills accumulate with a probability $m \pi_a$. While not working, skills decrease with probability $\pi_d$. When $\pi_a = \pi_d = 0$ endogenous skill accumulation and depreciation are eliminated and $x$ becomes a permanent random effect for the parent.

The Mincer earnings function that relates skill to accumulated labor market experience assumes $\pi_a = 1$ and $\pi_d = 0$. That is, the stock of skill accumulates linearly with experience and does not depreciate while not working. For other values of $\pi_a$ and $\pi_d$ welfare spells

---

9 In the case of $C(\alpha, \theta)$ and OS$^*(\alpha, \theta)$ their range depends on other values, IAB and $W_{\text{max}}$, respectively. Any outside support above IAB would be equivalent to IAB in terms of observed choices. In addition, household costs are bounded by $W_{\text{max}}$ because any value above that would be equivalent in terms of work choices. This scaling keeps the time-varying and stochastic terms of the utility function within their relevant ranges.
A single-parent household faces a shifting tradeoff between income and leisure (non-labor market time) driven by the state variable $h$. When $h$ is large costs are convex to the origin. The parent prefers part-time work, perhaps because children are in school or family members are available to provide part-time daycare. If these arrangement breaks down costs become concave. The parent now faces fixed costs of work, for example paid day care plus commuting time. Now the parent prefer either full-time work or no labor market activity (all else constant). The ability to respond the SSP or to spread fixed costs across more hours depends on whether the parent has a job that allows full-time work ($b = 2$) or only has a part-time job ($b = 1$).

caused by transient conditions can last longer than those conditions. The longer a parent
is out of work the more likely skill has fallen. Wage offers fall and become less valuable
relative to time spent in the household. If a job were taken, $x$ would eventually increase.
But in the presence of IA (even with forward-looking behavior) the rate of endogenous
wage growth may be too slow to make work pay.
Wages are expressed as full-time equivalent monthly earnings, denoted \( W(\theta) \). Jobs have two characteristics, \( b \) described above and an earnings offer \( n \) that takes on 6 values. The offer \( n = 0 \) is a "dead-end" job that does not depend on skill and pays MW regardless of skill. Such job offers are a fraction \( \pi_m \) of all jobs. Job offers with \( n > 0 \) come from a discretized log-normal distribution with log-mean \( \mu \) and log variance \( \sigma \).

The wage function allows for an interaction between the minimum wage, skill, the distribution of offers and the subsequent growth of earnings. To explain, start with the simple case of \( MW = 0 \). Then, \( W() \) collapses to a familiar log-linear form:

\[
\ln W^0(\theta) = \mu + \sigma \Phi^{-1}(n) + \eta \ln x, \quad n > 0. \tag{11}
\]

The offers are percentiles of the distribution with skill shifting the distribution. Each offer is equally likely and the parameter \( \eta \) corresponds to the return to experience. In the \( MW=0 \) case \( n = 0 \) is not a real offer. When \( MW > 0 \) it is assumed that regular offers \( (n > 0) \) are not each associated with its own level of earnings. Instead, for a given \( x \) let \( \phi_x \) denote the fraction of the underlying distribution below \( MW \):

\[
\phi_x = \Phi\left[\frac{\ln(MW) - \eta \ln(x) - \mu}{\sigma}\right]. \tag{12}
\]

For the lowest \( x \) the lowest two regular offers produce a wage of \( MW \). Each offer occurs with probability \( (1 - \pi_m)\phi_x/2 \). For the next skill level \( (x = 2/4) \) only the \( n = 1 \) offer is at \( MW \) with probability \( (1 - \pi_m)\phi_x \). For greater skill levels no wages other than \( n = 0 \) are at the minimum wage. So \( W(\theta) = MW \) if any of three mutually exclusive indicators are true:

\[
M(n,x) = B[n = 0] + B[x \in \{1/4, 2/4\} \land n \in \{1, 2\}] + B[x = 2/4 \land n = 1]. \tag{12}
\]

For other combinations of \( n \) and \( x \) the wage exceeds the minimum wage. Each offer is equally likely given \( x \). Let \( \tilde{n}(x) \) be the number of offers above \( MW \),

\[
\tilde{n}(x) = 3 + B[x > 1/4] + B[x > 2/4]. \tag{12}
\]
We arrive at the general expression for full-time earnings (including $W_{\text{max}}$ that appeared above):

$$W(\theta) = M(n, x)MW + (1 - M(n, x)) \left( x^n \exp \left\{ \mu + \sigma \Phi^{-1} \left( \phi_x + \left( 1 - \phi_x \right) / \tilde{n}(x) \right) \right\} \right)$$

$$W_{\text{max}} = \exp \left\{ \mu + \sigma \Phi^{-1} \left( \phi_1 + \left( 1 - \phi_1 \right) / 5 \right) \right\}.$$  \hspace{1cm} (13)

For a low-skill parent, minimum wage jobs differ in their growth potential. For some offers they will wait for two increases in skill before pay increases. For others only one increase is required. For high-skilled workers only $n = 0$ offers start at MW. All other offers will increase wages with the first accumulation of skill. Skilled workers may choose to quit a low-offer job to search for a better one (depending on such things as the job offer probability and the risk of losing skill). A dead-end job is not really a dead-end since it is assumed skills still accumulate on them (and can be transferred to other jobs). This allows for a pattern in which people respond to the SSP subsidy in terms of employment but may not be on track to self-sufficiency because it does not create a strong incentive to differentiate low-wage jobs with growth potential from those without.

**III.F SSP Plus**

The SSP Plus Sample was a small treatment group selected with the same criteria as the Recipient Study in New Brunswick. Subjects were given the same supplement under the same terms, but in addition they were offered a set of employment services. This additional treatment is not represented in Figure 1 which presumes a single wage is already available. Instead, the potential impact of these services is captured by a change in the job offer probability relative to otherwise identical households. The model assumes that these services enhance active search by raising the probability of a job offer each

---

10 This can (loosely) be interpreted as the employer over-paying a worker whose productivity is below MW and then eventually under-paying them once their skills increase to recoup the lost. However, no explicit bargaining or contracting model such as Flinn (2006) is included.
period:

\[ p_j(\alpha, \theta) = a [\pi_j + B[g = 1] \pi_+ (1 - \pi_j)] . \]  

(14)

This leads to a change in the decision to search actively for a job and what is an acceptable job offer. Active job search is treated as unobserved, and no attempt is made to measure whether an eligible subject took advantage of the services. Thus, \( \pi_+ \) is a measure of how effective the offer of the services is not how effective the services are given they are used.

The model assumes that these services enhance active search by raising the probability of a job offer each period:

\[ p_j(\alpha, \theta) = a [\pi_j + B[g = 1] \pi_+ (1 - \pi_j)] . \]  

(15)

III.G Value and Choices

To recap, the exogenous parameters that determine utility and transitions are gathered into two vectors,

\[
\Upsilon \equiv (\beta \ \eta \ \kappa \ \mu \ \nu \ \sigma \ \zeta \ \xi) \quad (16)
\]

\[
\Pi \equiv (\pi_j \ \pi_m \ \pi_f \ \pi_h \ \pi_i \ \pi_d \ \pi_l \ \pi_s \ \pi_+) .
\]

Where: \( \beta \) is the rate of income reporting; \( \eta \) is the curvature in skill; \( \kappa \) converts job search into work time; \( \nu \) is the (scaled) income-equivalent cost of full-time work; \( \mu \) and \( \sigma \) determine the location and spread of wage offers; \( \zeta \) determines the variance in the curvature of time-costs over time; and \( \xi \) is the factor on outside support. The \( \Pi \) vector includes all parameters that enter the transition from one period to the next: \( \pi_j \) is the probability that active job search generates a job offer (in the absence of job-finding support); \( \pi_m \) is the proportion jobs that are pure minimum wage jobs; \( \pi_f \) is the proportion of job offers that are full-time jobs; \( \pi_h \) is probability that the curvature in time-costs change; \( \pi_a \) is the probability that skills accumulate while working; \( \pi_d \) is the probability that skills decline while not working; \( \pi_l \) is the probability that a working parent loses their job exogenously;
\( \pi_s \) is the probability that outside support changes; and \( \pi_+ \) is the parameter that determines the effectiveness of SSP+.

The value of an outcome and the value of a state satisfy Bellman’s equation:

\[
v(\alpha, \theta) \equiv U(\alpha, \theta) + \delta E[V(\theta')] = U(\alpha, \theta) + \delta \sum_{\theta'} P\{\theta' | \alpha, \theta\} V(\theta') \tag{17}
\]

\[
\forall \theta \in \Theta, \quad V(\theta) = \max_{\alpha \in A(\theta)} v(\alpha, \theta). \tag{18}
\]

State-contingent choice probabilities are smoothed with a logistic kernel\(^{11}\) with parameter \( \rho \geq 0 \):

\[
\tilde{v}(\alpha, \theta) \equiv B[\alpha \in A(\theta)] \exp\{\rho[v(\alpha, \theta) - V(\theta)]\}
\]

\[
P\{\alpha | \theta\} = \tilde{v}(\alpha, \theta) / \sum_{\alpha'} \tilde{v}(\alpha', \theta). \tag{19}
\]

Given \( \theta \) and the choice probabilities the expected result vector is

\[
E[Y | \theta] \equiv \sum_{\alpha \in A(\theta)} P\{\alpha | \theta\} Y(\alpha, \theta). \tag{20}
\]

Combining endogenous choice probabilities with exogenous outcome-to-state transitions generates the state-to-state transition, \( P_s\{\theta' | \theta\} \). Based on this transition there exists an ergodic (stationary) distribution over the endogenous variables, conditional on the non-ergodic values \( d \) and \( k \) (see Ferrall 2003). Let \( P_\infty(\theta) \) denote this distribution, which is the starting point for modeling the selection into the experiment.

IV. The SSP Experiment

IV.A Conditioning Variables, Impact and Predictions

Since \( Y(\alpha, \theta) \) does not include the full outcome, the analysis must condition measurements on less information than households have. A standard impact analysis would

---

\(^{11}\) In smoothing choice probabilities without adding an extreme value error term to utility follows Eckstein and Wolpin (1999), footnote 30. The smoothing parameter is allowed to differ across types because it is closely related to the discount factor \( \delta \), which also differs by type. One \( \rho \) applied to widely differing \( \delta \)'s would result in very different amounts of smoothing.
condition only on variables that are not functions of past behavior (given eligibility). That is,

\[ \theta_{\text{cond}} \equiv (t \ g \ e \ d). \]  \hspace{1cm} (20)

The variable \( t \) is experimental time. Usually this would simply be the number of periods since random assignment, but the SSP has two different studies in which subjects enter the same program of treatment at different stages. Thus, to coordinate measurements, \( t \) is set to be 0 at the beginning of phase 2, the later of the two points of entry. The expected measurement is

\[ E [Y (\theta_{\text{cond}})] = \sum_{\theta} \lambda^* (k|\theta_{\text{cond}}) \Omega (\theta | k, \theta_{\text{cond}}) E [Y|\theta], \]  \hspace{1cm} (21)

where (20) defines \( E[Y|\theta] \) as the expected outcome conditional on the subject’s information. The Appendix defines \( \Omega(\cdot) \) as the distribution over states observed at \( \theta_{\text{cond}} \) among type \( k \) households and \( \lambda^*(\cdot) \) as the selected proportion of type \( k \).

Let \( \hat{Y}(\theta_{\text{cond}}) \) denote the vector of average observed (empirical) results conditional on the exogenous variables \( \theta_{\text{cond}} \). Observed impact is the difference in mean results between a treated group and its control:

\[ \hat{\Delta}(\theta_{\text{cond}}) \equiv \hat{Y}(\theta_{\text{cond}}) - \hat{Y}(\theta_{\text{cond}}|_G). \]  \hspace{1cm} (22)

The notation \( |_G \) means replace \( g \) in \( \theta_{\text{cond}} \) with \( G (=3, \text{the control group}) \). The model’s predicted impact is simply

\[ \Delta(\theta_{\text{cond}}) \equiv E [Y (\theta_{\text{cond}})] - E [Y (\theta_{\text{cond}}|_G)]. \]  \hspace{1cm} (23)

Some insights can be drawn from these expressions without reference to the particular model or experiment. While undergoing treatment the transitions are different from the real world, so the treatment group drifts away from its control group. Selection on un-observables is important if \( \lambda^* (k|\theta_{\text{cond}}) \) differs significantly from \( \lambda[k,d] \). Control groups are drifting as well, but they continue to follow the same transitions as outside the experiment.
Their state distribution converges back to $P_{\infty}$ but only given the underlying (permanent) type. Based on observables the control group outcomes converge to a different mean than outside the experiment due to selection on unobservables. Ultimately, treatment ends and treated households begin to converge to the same distribution as the controls. So the impact of treatment in a finite-lived experiment is relative to a non-stationary distribution that is converging to the same distribution as the treatment group but at a different rate.

V. Experimental Outcomes

V.A Measurements

As shown in Table 2, 8,898 people who took part in the SSP experiment are included in the analysis here. Roughly two-thirds were sampled from British Columbia, because the Applicant Study was conducted in BC alone. The SSP Plus Sample includes 292 people in New Brunswick. Roughly one-half of the households had more than one child at the baseline. The results here use 36 post-assignment months of data ($t = 1$ to $t = 36$) in the Recipient Study and 30 months ($t = -11$ to $t = 18$) in the Applicant Study. The result vector is computed for each value of demographic, experimental, and treatment group and by experimental time $t$.\textsuperscript{12}

The 12 contemporaneous variables chosen for study are summarized in Table 3. The monetary variables include mean monthly earnings, mean monthly IA benefits received, and mean monthly SSP supplements received (when applicable). The means of earnings squared and IA squared are also matched because higher moments of these distributions help identify the wage offer distribution. The remaining six results in $Y(\alpha, \theta)$ are indi-

\textsuperscript{12} Attrition from the sample after the baseline interview is treated as an exogenous result independent of the subject’s situation and the SSP treatment. According to this assumption it is valid to use either all individuals reporting results in a given month or use only those individuals who remained in the sample throughout the measurement period. Not all subjects entered the experiment in the same calendar month, so in the 36-month data file there are some observations beyond the 30 and 36 month cut-offs. For a cell’s values to be included in this analysis, there had to be at least 50 observations.

\textsuperscript{13} There is a lag in receiving SSP supplements and IA benefits. SSP benefits received and recorded in month $t = 2$ are, for the most part, based on outcomes in month $t = 1$. For IA the lag is often two
cators for labor market outcomes. The mean values are therefore proportions of subjects in the given situation, including receiving any IA in the month, earning a wage within $.10 of the current provincial minimum wage, not working this period due to quitting a job last period; not working this period due to losing a job last period, working full time (according to the SSP minimum), working part-time, and an interaction between receiving IA and working either full- or part-time.

V.B Experimental Impact

Table 4 reports relative impacts \( \Delta (\theta_{\text{cond}}) / Y (\theta_{\text{cond}} \mid_G) \) for selected variables at different values of \( t \). At \( t_0 + 1 \) relative impacts are small, as would be expected with random assignment. The only impacts that appear sizeable one month after assignment are 25% responses in earnings and full-time employment in the NB2+ and BC1 groups. By month 13 (one month after the qualification period ends) the earnings impact varies between 32% and 128%. By month 24 relative impacts are generally below the earlier maximum impact, but in many groups is still larger than the initial values. The relative impact on IA receipt is generally smaller than on earnings. By month 24 anywhere between 8% and 32% fewer subjects in the treatment groups are on IA than in the control group. The impact in the Applicant Study at month 13 is in the same range. The relative impact of the SSP treatments on the proportion of jobs at the minimum wage is typically negative and smaller than the other impacts. That is, conditional on working full or part time, a smaller proportion of the treatment groups are working at or near the minimum wage than in the control groups. The differences are small when compared to the impacts on full-time work itself, which range from 52% to 146% in the Recipient Sample.

The impact of the SSP treatment is not limited to mean values of the measured results. The co-relationship between the variables also differs across treatment groups. Table 5 reports the matrix of simple correlations in seven of the results. The SSP Plus Sample was

months. For this reason SSP and IA results are forwarded by one and two months so that they are (roughly) contemporaneous with the situation that generated them. This adjustment is not perfect, but it appeared to be the best fixed rule.
excluded and the four demographic groups were combined, leaving four entry/treatment groups. The main purpose of Table 5 is to compare the same correlation between treatment and control groups. In other words, to compare entries across the diagonal. In each of the four quadrants the signs of the correlations follow similar patterns, which is not surprising given that earnings must be strongly related to work hours and negatively correlated with IA receipt. When comparing correlations across treatments and controls we see only small differences in the Applicant Study. For example, the correlation between earnings and IA benefits among the treated is -0.356. Among controls the same correlation is -0.360.

The difference in the correlations is substantially larger in the Recipient Study, and the number of observations greater (however they are measured). For example, the same earnings/IA correlations are -0.409 and -0.317, respectively. This is consistent with the model since treatment is milder among applicants than recipients. For a minimum of twelve months there is no direct impact of treatment on utility for recipients. The impact is felt solely through the eventual opportunity to qualify for the supplement, and this forward-looking impact is the same as that felt in the Recipient Study from the start of their post-assignment period. For the applicants the impact is discounted by $\delta$ and by the uncertainty of finding a job. Thus, the applicant treatment group will on average appear closer to its control group than the recipient treatment group. The one caveat is that the two groups are created by nearly opposite criteria applied to IA receipt. As long as the underlying model exhibits positive correlation in IA receipt, the cross-treatment difference in correlations will indeed be smaller in the Recipient Study. The presence of skill accumulation and depreciation, along with persistence in the other household states and the IA rules themselves combine to ensure some measure of persistence in IA receipt.

Table 5 suggests that analyzing each measured result (and impact) separately is inefficient in a statistical sense. That is, earnings, IA, and full-time employment are not separate outcomes that each requires a separate sequence of impacts. More importantly, the SSP treatment is associated with differences not just in mean results, but also in correlations.
across contemporaneous results. Even when not using individual-level panel data, the
different movements in mean results across variables through experimental time contains
important information about the treatment.

VI. The Estimated Model

The model is estimated using Generalized Method of Moments by minimizing the
distance between observed and predicted values of the conditional moment vectors:

\[ \hat{\Delta}(\theta_{\text{cond}}) = \hat{E} [Y | \theta_{\text{cond}}] - E [Y | \theta_{\text{cond}}] = 0, \]

for all vectors \( \theta_{\text{cond}} \) post random assignment. The interaction of \( d, g, e, \) and \( t \) with the
twelve contemporaneous results contained in \( Y(\alpha, \theta) \) results in 4884 total moments. The
Appendix describes the estimation procedure including computation of the optimal
weighting matrix. It also discusses how variation across samples, treatments, provinces,
experimental time and elements of the measurement vector \( Y(\alpha, \theta) \) contribute to the ident-
ification of parameters of the model. Some technical aspects of the estimation are show
in Table 6.

VI.A Parameter Estimates

Table 7 reports the estimated parameter vector. Since there are no coefficients on
observed variables included in the parameters (as in, say, a Mincer earnings function)
many of the parameters are difficult to compare with other results. Many of the values
are probabilities, but their magnitudes depend on the number of values the state variables
take on. For these reasons the discussion of the parameters is short whereas as discussion
of the model predictions is extensive.

\footnote{One can treat a conventional impact analysis as estimating each difference between treatment and
control, \( \hat{\Delta}(\theta_{\text{cond}}) \), with a free parameter (the predicted impact is the observed impact). Meanwhile, the
estimated model generates impact as the difference between two of the model’s predictions without adding
new parameters. Thus, the model estimated using GMM can be seen as a nested hypothesis within the
unrestricted impact analysis. From this point of view, an impact analysis has as many parameters as
moments and has no power to predict out of sample. The estimated model is parsimonious, with only 79
free parameters.}
The estimated mixing probabilities in Table 7.1 show that two types predominate in BC and three types in NB, with NB1 primarily of one of those types. Type proportions vary more across provinces than between numbers of children. The dynamic programming parameters in Table 7.2 indicate that types have very different levels of patience. A period is one month. For only the first two types is δ̂ close to 1 and place a large amount of weight on the future. The other types make decisions close to a static manner: next year’s outcomes have essentially no impact on today. The income reporting parameter β is straightforward to interpret. Three types are estimated to report approximately 40% of their income when on welfare. One type reports 95%.

Wage offer distributions differ across types (Table 7.3) as does the stigma associated with welfare (captured by the coefficient on outside support, χ̂). Full-time work has a very similar cost across type (ν), but recall that this value is relative to maximal earnings for a given type. This contrasts with the cost of active job search, which is only large and precisely estimated for type 1 (and to lesser extent type 4). Returns to skill and the convexity in household costs are difficult to interpret beyond their effect on predictions.

The last panel of the parameter estimates (Table 7.4) reports the transition shifters. Here we see that type 1 is constrained by a low job offer probability. Most offers are full-time, so the high fraction of part-time work reflects a choice to work fewer hours than the job allows. Between 13% and 53% of job offers are true minimum wage jobs (with no on-the-job growth potential). Estimates of the home environment indicate that outside support is highly persistent (πs is small) but household costs of work and job search is not (πh is high). Type 1 workers accumulation skills each period (and have rapid on-the-job wage growth). For other types growth is slower, but still only type 3 has any significant chance of further growth after one year of working. Because average wages do not accumulate in the treatment group, this suggests that the return to skill (ν) reported in the previous table is not large. Thus, parents achieve modest wage growth early in an employment spell but not sizeable long-term growth. Only for type 3 is depreciation of
skills rapid while not working. Thus the impression the model gives for the SSP results is that the treatment requires long-term and persistent growth in skills. Skill persistence is much less of an issue than a predominance of jobs with no growth potential and a low wage elasticity to skill accumulation.

**VI.B Fit to Selected Moments: Earnings, OnIA, Total Transfers**

Figure 5a and Figure 5b present the observed and predicted moments for New Brunswick and British Columbia, respectively. For both family sizes in New Brunswick the predictions track the data quite well. Selection and the evolution of state variables together generates the upward trend in the control groups as they return to the ergodic distribution. The response to treatment generates an impact that mimics the data. The one aspect of the data that the estimates fail to capture qualitatively is the slope of change in the Applicant groups (Figure 5b). The starting level and impact are accurate but the selection effect in the Applicant study is larger than the model predicts. The fraction of each group on IA is shown in Figure 6a and Figure 6b. The match to the data is similar to that for earnings, although the mismatch in the Applicant study is of a different form. For OnIA the model impact is too large before time 0. From the government budget perspective, the SSP is valuable if additional transfers during treatment result in lower transfers later on. Figure 7a and Figure 7b show total transfers, IA + SSP. Since the impact fades, the policy is a failure in total transfers. In all groups and at each month the impact on transfers is non-negative. The subsidy never induces a substantial move to self-sufficiency. In some groups the model generates a larger impact than the data, but it captures the rise and then near constant impact until month 36.

**VI.C Variation from Policy, Selection and Heterogeneity**

Figure 8 illustrates the combined effects of all sources of variation. Each panel shows the behavior of a particular unobserved type in all four observed environments. The two most patient types, \( k = 1 \) and \( k = 2 \), are shown. Since preferences are held constant, the
effect of policy variation is illustrated by comparing the four panels within each type. And since the SSP is based on a selected sample the trends in the control groups capture how distant the selected group is from the population average. The ergodic mean is shown as a triangle. For type 1 we see that all groups are well below the average in earnings. By month 36 the control group has nearly returned to the ergodic distribution.

The most striking aspect of the top half of Figure 8 is the large response to SSP+, which is a combination of a large estimate of effectiveness ($\pi_\oplus$) and a low job offer probability ($\pi_j$). Type 1 is a small fraction of the NB population so the modest additional impact of SSP+ is a combination of a large individual response among a small part of the population. This same group is not particularly responsive to the SSP treatment; its households are constrained by a lack of job offers which the SSP+ alleviates.

The bottom half of the Figure shows type 2. For this type the selection effect is more extreme and even after 36 months the control group is still far from the stationary average. The impact under NB policies starts very small and then becomes negative. Apparently this group was induced to accept low wage jobs to qualify for the supplement while their control group counterparts held out for better jobs. Those who qualify tend to keep these jobs until the subsidy ends. This group illustrates one of the difficulties in designing incentive schemes for low-skill parents. The SSP encourages employment but not necessarily patience to wait for employment with high growth potential. The response of type 2 is itself heterogeneous, because the opposite pattern occurs under British Columbia policies. Here the expected impact in earnings appears and is long lasting. However, type 2 is estimated to be a vanishingly small fraction of the population in BC.

VI.D Treatment: Identification or Validation?

The estimated standard errors reported in Table 7 indicate that many parameters are precisely estimated by the variation in moments generated by the experiment. The
parameters are identified by restrictions on how the moments can vary across treatment groups, over time within a group, and across demographic groups. An alternative use of the exogenous variation generated by the experiment is to validate a model estimated only on the control group. Todd and Wolpin (2006) and Lise et al. (2003) follow this approach by estimating models of forward-looking agents on control groups within experiments (Progressa and the SSP, respectively) and then using the experimental data for out-of-sample validation. A major advantage of this approach is that behavior under the treatment does not have to be solved repeatedly while estimating the parameters. The potential cost is that the model that can be estimated from the control group alone may be not be as rich as one that can be estimated using the experimental data. Thus, the parameter estimates may be less applicable outside the sample and less reliable for understanding behavior in populations facing similar but not identical environments.\footnote{Todd and Wolpin (2006) suggest that if the model is validated then one might go ahead and estimate using the experimental data as well to increase efficiency of the estimates. This retains any limits on the identified model created by the limited variation in the control group.}

To quantify this potential cost of not using the experimental variation for estimation, the standard errors for the parameters were re-computed using only the moments within groups. Results were re-scaled to mimic a sample of the original size. Table 8 reports the results. Standard errors based on all the data are compared to those from the control and treatment groups alone.\footnote{The standard errors in the “all” column are slightly different that in Table 7 because in all calculations the columns for $\pi_{+}$ were eliminated. It was eliminated because it is not identified from the “Ctrl” group, an extreme example of limited identification from variation within a group.} First consider the “Ctrl” column. It is not surprising that throwing out the experimental variation increases the standard errors. However, for nearly all the parameters the estimated standard error is eight times larger than when based on all the data. Included among these are key parameters for understanding dynamic behavior of low-income households: the discount factor ($\delta$), the wage offer parameters ($\mu$ and $\sigma$), the return to skill ($\eta$) and many probabilities that determine persistence in wages and other states. Thus, if the validation strategy had be used here, a model estimated from the
control data alone would have been much simpler in form without the ability to capture some details in the experimental outcomes.

Another result is revealed in Table 8 when the “all” column is compared to the “Treat” column. This counter-factual throws out the variation between treatments and controls and replaces it with more information on the experimental variation. In nearly all the cases the re-scaled standard errors are smaller when based on the treatment groups alone and often the increased precision is not trivial. In many cases the standard error is reduced by 25% or more. The source of this extra precision is simply the experimental variation in incentives generated by the experimental design. Within the program of treatment the next month is quite different than the current month since one deadline or another is approaching. Within the control group no such deadlines exists.

This result has a somewhat surprising implication. When using social experiments solely to study impact (Δ(θ_{cond})) a control observation and a treatment observation have equal weight. Thus, absent other costs, splitting the overall sample evenly is a reasonable design. Table 8 suggests that this logic does not hold when experimental variation will be used to identify an underlying model. In this case impact is not the only outcome of interest and an additional treated observation may be more valuable than an additional control observation. When the fixed costs of running a experiment and it is designed to estimate an economic model then the treatment group should be larger than the control group.

VI.E Out of Sample Predictions

Figure 9a and Figure 9b compare the model predictions to data on earnings and OnIA by province from the end of the experiment that was not available for estimation.\footnote{In general, whichever group faces more exogenous variation is the key. Perhaps some experiments actually reduce variation in the subjects’ situations and it is the control group that retains. However, most experiments tend to be finite-lived and in all cases they create a ‘surprise’ that varies across individual states. So typically the experimental group will be the one with more in-sample economic variation than the control group.} Not

\footnote{Due to a finite research period promised to SSP participants, neither the micro data nor summary}
surprisingly, the impacts seen to fade in the previous figures continue to fade toward zero as treatment ends and all subjects return to the status quo. The model’s prediction are similar in trend but it continue to miss the the level of earnings in the applicant sample. One pattern that is intriguing is that the impact of the SSP+ continues to lie above the regular impact even after treatment ends. The impact decays more in the model, but it also produces a lasting impact of the extra help in the SSP+ program.

VII. Policy Experiments and Other Implications

This section conducts experiments that explore the implications of the SSP for counterfactual policy questions. In the figures the results of the hypothetical changes are compared not with the data but with the model predictions based on the SSP experiment.

VII.A Experiments on Sample Design of the SSP

The SSP+ treatment was available only in New Brunswick and the Applicant study was conducted only in British Columbia. Using the model the missing experiments can be run. Figure 10 shows total transfers for NB 1 Child and BC 2+. We see that an Applicant study in NB would have had a very large short-run impact but once treatment ended those who qualified would quit work and return to IA. The pattern makes it clear that in NB the predominant types can easily find a full-time job before time 0, stay on IA to remain eligible then at time 0 start collecting the subsidy. Since this figure extends to month 60 it also reveals an implication not shown earlier in Figure 7a. Namely, the model does produce a very modest negative impact on government transfers in the NB 1 Child group. It occurs only near month 48 when all supplements are ended.

Another interesting pattern emerges if SSP+ were run in BC where job offers are a major constraint. The extra job-finding help would not only have a major impact, but it is negative almost from the start, meaning that total government transfers are cheaper under statistics for the demographic categories used in estimation are any longer. The model predictions are averaged to compute province-level predictions which are available.
the SSP than welfare. In BC good jobs (low supplements) are available but hard to find. Impacts are long-lasting. This prediction is out-of-sample and is possibly an artifact of the model of heterogeneity. Perhaps the impact of SSP+ in job offers would not be so high in BC because with more types the concentrated effect in NB would not be shared by a large fraction of the BC population. This highlights the difficulty of drawing inferences from experiments out of sample whether it is based on a model or an atheoretic impact analysis. Together with the modest negative impact in NB it also shows that the hoped-for impact of the SSP and SSP+ are present in the model, but when accounting for all the outcomes the responsive households are not common.

Next, for the Recipient Study consider a sample of single parents who are on IA for exactly 6 months rather than 12 or more months. A practical reason for the or more clause is that it creates a large population to draw from and it includes long-term welfare recipients. On the other hand, if the SSP were implemented as policy it would not be long until the people qualifying for it would only be on IA for twelve months. The stock of long-term recipients without the benefit of the SSP would no longer exist. Perhaps an experiment on the flow into the long-term recipient pool would more closely reflect results of an SSP policy after an initial transition period. Because the long-term response is so low in the recipient sample an entry condition of just six months on welfare is used. This is an out-of-sample change since many parents meeting this condition would not meet the twelve-month rule. A reverse change is made to the Applicant Study. Parents newly applying to IA after one month or more off as opposed to six months or more are eligible.

The results of this switch in stock versus flow sampling is shown in Figure 11. We see that this slight change in experimental design might have had a very different pattern, at least in NB where the immediate impact is much larger although the impact still disappears rapidly once the supplement ends. The change to six-month flow sampling actually wipes out the small negative impact on total transfers in the NB 1 Child group. The conclusion is that, even if the SSP had encouraged real policy reform it may not have provided accurate
guidance for the ultimate response since the stock of long-term IA recipients appears to be much different than the flow, at least in NB. For BC 2+ children households, the difference with the actual SSP sampling scheme is modest, although we see that steeper slope in the data in the Applicant group is similar to the model when "six or more months of IA" is not enforced.

**VII.B Alternative Treatments**

Finally the model’s prediction for total transfers under two alternatives to the SSP treatment are shown. One is a simpler and larger “re-employment bonus.” This is a policy in which full-time employment is subsidized for just six months not three years. In addition, the subsidy is up to the full 3.9 times minimum wage earnings rather than “half-way” to that same target wage. Based on the previous result the outcome is not hard to guess, and Figure 12 shows the outcome for the BC 2+ group. This group jumps at the subsidy and drives total transfers way up. But the impact on long-run behavior is even worse than under the SSP. Thus there is some benefit to the longer subsidy period, but as illustrated earlier the SSP was not precise enough to encourage taking only jobs with good wage growth potential. This simple bonus is even worse in these terms.

Finally, consider an experiment that would be difficult to run but may reflect a policy that is ultimately behind most reforms to welfare. Namely, consider offering the SSP treatment while cutting IAB by 20%. Many parents who do not anticipate finding a job will be worse off in this treatment, but real policy changes might likely combine the carrot of the SSP with a stick of reduced IA levels. Figure 13 shows the effect for BC 2+ for OnIA and total transfers. Recipients respond strongly to the cut in benefits. Rates on IA are much lower during the qualifying and eligibility phases. And unlike the actual treatment the impact on total transfers are negative during the qualifying phase. But as implemented, those who failed to qualify leave treatment and return to regular IA benefits. Rates and transfers return to roughly what we see in the experiment.
VIII. Conclusion

Social experiments are designed to guide decisions based on a particular policy (the treatment). As a by-product they create exogenous variation which can be used to infer behavioral responses to other similar policies. That inference depends on a model, of course. This paper has found that results from the SSP can be modeled in a comprehensive way. During treatment the SSP generated sizeable impacts in key outcomes that the model captures quite well, but it failed to induce any obvious long-run move to self-sufficiency. Out-of-sample prediction of the model are validated on this score as well. The model confirms the difficulty in affecting long-term outcomes for low-income households through lack of job market opportunities, slow transitory skill acquisition, and short decision horizons generated by low discount factors in some parts of the population. Counter-factual experiments confirm that related policies could induce greater short-run response. Only in the case of the SSP+ treatment is there any hint of lasting impacts among a fraction of the population. These are parents who are forward-looking and can acquire skills but have trouble securing employment. With regard to the SSP+ the model has intriguing prediction that stronger results may have been detected if the SSP+ had been run in British Columbia where the population mix contains a higher proportion of this type.

Beyond welfare policies, this paper has explored an alternative approach to combining models and experiments. Estimated standard errors computed after removing groups demonstrate quantitatively that intra-group variation generated by the treatment is critical for identifying a rich and presumably more generally applicable model of household behavior. The literature emphasizes either inter-group variation without any model of behavior or a reliance on variation within the control group for identification. In the case of the SSP either of these strategies is highly inefficient in using the costly exogenous variation generated by the experiment.
IX. Technical Appendix

IX.A Details of the Model

Components of Income

\[
\text{TrueEarn}(\alpha, \theta) \equiv mW(\alpha, \theta) \\
\text{Earn}(\alpha, \theta) \equiv (1 - \beta i) \text{TrueEarn}(\alpha, \theta) \\
\text{IA}(\alpha, \theta) \equiv i \max \left\{ IAB - \beta CB \left( \text{Earn}(\alpha, \theta) - \min\{ SA, \text{Earn}(\alpha, \theta) \} \right), 0 \right\} \quad (A25) \\
Q(\alpha, \theta) = B \left[ i = 1 & 2 \leq f \leq 5 & m > PT & W(\theta) \geq MW \right] \\
\text{SUP}(\alpha, \theta) = Q(\alpha, \theta) \max \{ 0, (1 - TB) [UL \times MW - \text{Earn}(\alpha, \theta)] \}
\]

Policy parameter values that appear in these equations are reported in Table A.2. The components MW, IAB, and SA are defined in the text and illustrated on Figure 1. An exception is CB, which is the claw-back rate on benefits after the month earnings setaside. On Figure 1, and in all aspects of this paper, its value is 1. But subsequent policy experiments can consider changes in the claw-back rate. UL and TB are the parameters of the SSP which did not vary in the experiment but could in subsequent experiments.

Endogenous Variables. To describe the transition for each variable, let \( q' = q^*(\bar{q}, \{ \pi_j \}, \{ Q_j \}) \) denote a discrete variable \( q \) that has a default value of \( \bar{q} \) next period and can then jump into one \( j \) different sets of values with probability \( \pi_j \) (not the same as the model parameter). Conditional on jumping into \( Q_j \), each element of the set is equally likely.

S1. Unobserved Type: \( k \in \{1, 2, 3, 4\} \)

○ Role: index into \( \Gamma \) and the mixing distribution \( \Lambda \).

○ Transition: \( k' = k^*(k, 0, 0) \)

S2. Observed Type: \( d \in \{1, 2, 34\} \)

○ Role: index into the policy vector \( \theta_{pol} \) and the mixing distribution \( \Lambda \).

○ Transition: \( d' = d^*(k, 0, 0) \)

S3. Household Time Cost: \( h \in H = \{1/4, 2/4, 3/4\} \)

○ Role: determine the curvature of the time-cost function.

○ Auxiliary Equations

\[
c(h) = -\zeta \ln (1 - h) . \tag{A26}
\]

The right hand-side is the inverse exponential distribution with decay rate \( 1/\zeta > 0 \). The value of \( c(h) \) determines the convexity of costs for labor market activity less than full-time. For values of \( c(h) < 1 \) the cost function is concave for feasible labor market time, creating a tendency to prefer part-time work. On the other hand, costs are convex when \( c > 1 \), which creates a tendency either to stay at home or work full time.

○ Equation in Text: ( 10)

○ Transition: \( h' = h^*(h, \pi_h, H) \)

S4. Outside Support Opportunities: \( s \in S = \{0, 1/3, 2/3, 1\} \)
Role: determines the cash-equivalent amount of support available to the parent that, if accepted, disqualifies the parent from IA.

Equations in Text: (9)

Transition: \( s' = s^*(s, \pi_s, S) \)

**S5. Upper bound on working hours:** \( b \in \{0, PT, 1\} \)
- Role: constraint on work hours in current job
- Auxiliary Equations: See feasible actions below
- Transition:

\[
\begin{array}{c|c}
 b'(\alpha, \theta) & P\{b'|(\alpha, \theta)\} \\
0 & B [m > 0] \pi_l \\
1 & (1 - p_j(\alpha, \theta))B [b < 2] + B [m = 0] \\
2 & p_j(\alpha, \theta)\pi_j \\
3 & p_j(\alpha, \theta)\pi_j \\
b & B [m > 0] (1 - \pi_l). \\
\end{array}
\]  

**A27**

**S6. Accumulated Skill:** \( x \in \{1/4, 1/2, 3/4, 1\} \)
- Role: level of earnings and future growth potential
- Equations in Text: (11), (12), and (13)
- Transition:

\[
x' = x^* \left( x, [m\pi_a + B [m = 0] \pi_a], [\min\{\max\{1/4, x + B [m > 0] /4 - B [m = 0] /4\}, 1\}] \right)
\]  

**A28**

**S7. Wage Offer:** \( n \in \{0\} \cup N = \{1, 2, 3, 4, 5\} \)
- Role: search-sensitive component of wages
- Equations in Text: (11)-(13)
- Transition: with MW= 0,

\[
n'(\alpha, \theta) = n^*(n, \{ap_j\pi_m, ap_j(1 - \pi_m)\}, \{0\} \cup N).
\]

With MW > 0

\[
n'(\alpha, \theta) = n^* \left( n, [ap_j\pi_m, ap_j(1 - \pi_m)\phi_x, ap_j(1 - \pi_m)(1 - \phi_x)] \right), \quad \{0\} \cup \{1/6, \ldots, (5 - \tilde{n}(x'))/6\} \cup \{(6 - \tilde{n}(x'))/6, \ldots, 5/6\} \\
\]

(A29)

Note that the distribution of \( n' \) depends on the contemporaneous state through the value of \( x' \). So between periods \( x' \) must be determined before \( n' \).

**S8. Job Loss:** \( l \in \{0, 1\} \)
- Role: exogenous loss of job.
- Transition: \( l = l^*(0, B [m > 0] \pi_l, \{1\}) \)

**S9. Employed Previously:** \( p \in \{0, 1\} \)
- Role: tracks whether the person worked last period (with \( l \) can infer the parent quit).
- Transition: \( p' = l^*(B [m > 0], 0, \emptyset) \)
Actions.

A1. Labor market hours: \( m \in M = \{0, 1/4, 1/2, 3/4, 1\} \)

\( \diamond \) Equations in Text: (10)

A2. Active Job Search: \( a \in \{0, 1\} \)

\( \diamond \) Equations in Text: (15), (10)

A3. Accept Income Assistance: \( i \in \{0, 1\} \)

\( \diamond \) Equations in Text: (9), (25), (25)

Feasible Actions

\[
A(\theta) = \{(m \quad a \quad i) \in \{M \times \{0, 1\} \times \{0, 1\}\} : m < b & ma = 0\}. \quad (A30)
\]

IX.B The SSP Experimental Design

A subject’s status in the treatment program is defined by the sub-vector \( \theta_{\text{clock}} = (t \quad r \quad f) \), where \( f \) is the current phase of treatment, \( r \) is the number of periods the subject has resided in that phase, and \( t \) is experimental time, which is defined below. The SSP program of treatment is defined by a vector of parameters,

\[
\Psi_t[g] = (R[1] \quad \cdots \quad R[5] \quad f_n(y) \quad PT \quad TB \quad UL). \quad (A31)
\]

In the SSP experiment there are seven phases, numbered from 0 to \( F = 6 \). Both \( f = 0 \) and \( f = 6 \) correspond to the real, non-experimental world, before random assignment (0) and after treatment has ended (6). By definition, control groups \( (g = 3) \) transit immediately from phase 0 to phase 6. The treatment groups transit from phase 0 to the initial phase for their treatment group (listed in Table A.1). Ultimately they reach phase 6 as well. Phase 1 is the entry phase, where a parent must remain on IA for twelve months to get a chance to qualify for the SSP treatment. Phase 2 is the qualification period in which the parent becomes eligible for the SSP supplement if and when they begin a full-time job. They remain eligible for the supplement during phases 3 to 5. \( R[f] \) is the maximum duration of treatment phase \( f \). Since each phase of the SSP lasts at most 12 months, \( R[f] = 12 \) for \( f = 1, 2, \ldots, 5 \). The parameter \( f_n(y) \) is shorthand for a set of deterministic transition rules for next period’s phase. In other words, it describes how the SSP treatment progresses. Table A summarizes the selection, assignment, and transition rules in the SSP.

The remaining elements of \( \Psi_t \) are parameters that determine the value of the SSP supplement, \( \text{SUP}(\alpha, \theta) \), which enters utility defined in (A7) through income defined in (A8). The full equation for \( \text{SUP}(\alpha, \theta) \) appears in (A25). The red line in Figure 1 that passes through OS and 2.9MW+OS illustrates the effect of the supplement on the household budget.

The treatment variables \( r \) and \( f \) are not useful for coordinating observations across groups. For example, one parent may take 8 months to leave phase 2 while another may take only 4 months. After seven months the first parent’s clock would read \((7 \quad 2)\), the second \((3 \quad 3)\), and for all parents assigned to the control group it would read \((1 \quad 6)\). And the values of \( r \) and \( f \) are meaningless for parents assigned to control groups. To make results generated by the model compatible across groups a separate data clock, \( t \), tracks the experimental month at which a measurement is taken.
IX.C The SSP samples

A subject’s treatment group in the SSP is indexed by the sub-vector $\theta_{\text{exp}} = (e \ g)$, where $e$ is the experimental sample and $g$ is the randomly assigned treatment status within samples. The Recipient Study ($e = 3$) includes parents that had been on IA for at least one year. The Applicant Study ($e = 2$) includes parents initiating (applying for) a new spell of receiving IA after a period of at least six months without IA. The treatment variable $g$ takes on three values. Besides a control group ($g = 3$) and a treatment group ($g = 2$), the separate SSP Plus group ($g = 1$) was offered job-search and employment services in addition to the SSP supplement. Each treatment group has associated with it an initial post-assignment clock setting, a pre-assignment selection period and a sequence of feasible histories.

$$\Psi_{\tau}[e] = \left( \tilde{\theta}_{\text{clock}} \ T \ H[y;\theta_{\text{cond}}] \right).$$ (A32)

The elements of $\Psi_{\tau}$ are listed in Table 2. To make measurements consistent across groups the experimental clock $t$ must be coordinated. The time $t_0$ corresponds to the point of random assignment in the group and is normalized to 0 in the group that enters the program of treatment last. Thus $t = 0$ at the beginning of the qualification phase ($f = 2$) which is when the Recipient Study ($e = 2$) is randomly assigned.

Prior to $t_0$ is the period of sample selection. For the Recipient Study this period is of length $T = 12$ and stretches back to $t_{\min} = -11$. It requires the parent receive IA each period, so only outcomes with $i = 1$ are feasible during this time. The Applicant Study ($e = 1$) is randomly assigned at $t_0 = -11$ and the selection period is $T = 7$ periods long, extending back to period $t_{\min} = -17$. In the first six periods the feasible condition is $i = 0$, and the last period is the condition $i = 1$, the start of a new spell of receiving IA. One fine point is that after random assignment the Applicant sample has already spent one month on IA and requires only eleven more months to enter phase 2. Therefore the initial clock setting has $r = 2$. Formally the selection criteria can be represented several different ways. Table 2 represents them as a 0/1 indicator for a measurement vector.

C. Ferrall. Explaining The SSP IX. Technical Appendix Page 42
C. Ferrall. Explaining The SSP

Table A.2. Policy Vectors Contained in $\Psi_{pol}$.

<table>
<thead>
<tr>
<th>d</th>
<th>IAB</th>
<th>MW</th>
<th>SA</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>712</td>
<td>650</td>
<td>200</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>755</td>
<td>755</td>
<td>712</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>982</td>
<td>780</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1175</td>
<td>780</td>
<td>712</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>g</th>
<th>PT</th>
<th>TB</th>
<th>UL</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>75%</td>
<td>50%</td>
<td>3.90</td>
</tr>
<tr>
<td>1</td>
<td>75%</td>
<td>50%</td>
<td>3.90</td>
</tr>
</tbody>
</table>

$\Psi_s[e]$ $H[y; \theta_{cond}]$

$\Psi_{pol} = (\Psi_p \ \Psi_x \ \Psi_t)$

(33)

IX.D Conditional Distributions

To compute the selection into sample $e$ by group $d$, begin by setting $t = t_0 - T + 1$ and $g = 3$, which determines the value of the conditioning vector $\theta_{cond}$. Choose an unobserved

Provincial minimum wages changed during the experiment. These changes are not accounted for in the model but they are accounted for when classifying parents as working at a minimum wage job or not in computing mwgm. i means $i=1$ (on IA); $-i$ means $I=0$ (off IA).
type \(k\) and use the corresponding ergodic distribution as the starting value:

\[
\Omega \{ \theta' \mid k, \theta_{\text{cond}} \} = P_\infty(\theta).
\]

Initialize the selected weight of type \(k\) to one:

\[
\omega(k; \theta_{\text{cond}}) = 1.
\]

During selection the feasible choices are imposed on the choice probabilities.

\[
P^* \{ \theta' \mid \theta \} = \sum_\alpha P \left\{ \theta' \mid [B[t = t_0]] \theta_{\text{clock}} \right\} \mathcal{H} \{Y(\alpha, \theta); \theta_{\text{cond}}\} P \{\alpha \mid \theta\}. \tag{A35}
\]

The notation \(\mid_x\) means to set elements of the state vector to \(x\) holding other elements constant. The condition \(B[t = t_0]\) means this only happens at time \(t_0\). In other words, subjects make their last choice before random assignment ignorant of the experiment. Then during the transition to the next month’s state, those in a treatment group have their clocks reset to the initial clock for that experimental sample. They ‘wake up’ in the program treatment with all other states determined by choices before the experiment.

Working recursively forward in time first compute the fraction of type \(k\) households that make it to the next period:

\[
\omega(k; \theta_{\text{cond}}|_{t+1}) = \omega(k; \theta_{\text{cond}}) \sum_{\theta'} \sum_\theta P^* \{ \theta' \mid \theta \} \Omega \{ \theta \mid k, \theta_{\text{cond}} \}. \tag{A36}
\]

The proportion of the unselected population that is eligible for assignment may become very small. Thus, the distribution across states is updated and re-normalized to sum to one:

\[
\Omega \{ \theta' \mid k, \theta_{\text{cond}}|_{t+1} \} = \frac{\omega(k; \theta_{\text{cond}})}{\omega(k; \theta_{\text{cond}}|_{t+1})} \sum_\theta P^* \{ \theta' \mid \theta \} \Omega \{ \theta \mid k, \theta_{\text{cond}}|_{t-1} \}. \tag{A37}
\]

Once \(t + 1 = t_0\) we have the distribution eligible for random assignment. All of these calculations can be done independently (in parallel) across both \(d\) and \(k\). But once generating the predictions after random assignment the type-specific distributions must be adjusted:

\[
\lambda^*(k; \theta_{\text{cond}}) \equiv \lambda[k, d] \frac{\omega(k; \theta_{\text{cond}})}{\sum_{k'} \omega(k'; \theta_{\text{cond}})}. \tag{A38}
\]

Since the clock was set properly at \(t_0\), the updating rules (35)-(37) apply for \(t > t_0\) as well. Since all actions are feasible after random assignment in the SSP, \(\omega(k; \theta_{\text{cond}}|_{t+1})\) becomes constant and correction factor on \(\Omega \{ \theta' \mid k, \theta_{\text{cond}}|_{t+1} \}\) becomes one. This assumes that attrition is uncorrelated with unobserved types (and unobserved states).
IX.E Solving the Model and Computing Predictions

The size of the model and some technical details of the solution are listed in Table 1 and Table 6. The size of the system is notable. Even though each endogenous state variable is restricted to a small set of values, an individual subject can be in one of 2,304 states outside the experiment. The post-treatment infinite horizon problem requires convergence of the value function at these points, although some points in the state space are, from the subject’s point of view, redundant and do not require re-solving the maximization problem (A17). For example, the household is not affected by the values of \( t \) and \( p \), and a currently unemployed worker \( (b = 0) \) does not care about values of \( n \).

Since a stationary distribution \( P_{\infty} \) over states is computed, 16 different linear systems of size 2,304 must be solved on each iteration of the model. The SSP program of treatment adds 60 additional values of \( f \) and \( r \). With the separate SSP Plus treatment and Applicant sample over 4 phases leads to 51,840 total states for an individual. In keeping track of all states while tracking experimental results in a total of 6,672,384 different combinations are possible. Up to 12 actions are available at each state. When aggregating over all states (including demographic, unobserved, and equivalent variation) the result is an outcome space of size 80,068,608.

The value function (A18) is solved to a level of precision under the infinite horizon. Evaluating the model ‘from scratch’ takes a bit more than an hour using a single processor of a high-end server. The required time is sensitive to the size of the discount factor \( \delta \). This cost can be cut by roughly \( 1/(16) \) through the use of 16 processors to solve in parallel the separate problems defined by \( d \) and \( k \). Further substantial savings occur when computing numerical gradients by taking account of the limited interactions across parameters implied by a finite mixture model (Ferrall 2005). These savings are essential to making the model feasible to solve. With the computing resources currently available a full iteration of the BFGS algorithm can completed in approximately an hour.

---

Steps in Computation.
A0. Set \( \theta_{\text{exog}} = \theta_{\text{exog}}^0 \) and call an optimizer to minimize \( W(\theta_{\text{exog}}) \).
A1. To evaluate \( W(\theta_{\text{exog}}) \); Set \( d = D \).
A2. Solve completely for one group \( d \). Set \( k = K \).
B0. Solve for behavior. Set \( f = F, r = 1, g = G, e = E \).
B1. Iterate on \( V(\theta) \) in 18 to convergence.
B2. Once converged, loop one more time over \( \theta_{\text{end}} \) to compute choice probabilities \( (P \{ a | \theta \}) \) in 19 and \( E \{ Y | \theta \} \).
B3. Solve the linear system that defines \( P_{-\infty} \) for \( k \) and \( d \).
C0. Solve for the endogenous sample in entry group \( e \). Set \( t = t_{\text{min}} \).
C1. From \( P_{-\infty} \), compute the first value of \( \omega (k; \theta_{\text{cond}}) \) and \( \Omega(\theta | k, \theta_{\text{cond}}) \).
C2. Increase \( t \) by 1. Update \( \Omega \) and \( \omega \) by looping through all transitions.
C3. Repeat previous step until \( t = t_0 \).
C4. Store \( \Omega \) to be used for all \( g \) given \( e, k, d \).
C4. **Solve for behavior under treatment.** If \( g = G \) set \( f = 0 \) and skip this part.

- **E0.** Decrease \( f \) and set \( r = R[f] \).
- **E1.** Solve for \( V() \), choice probabilities, and \( E[Y | \theta] \).
- **E2.** Decrease \( r \) by 1. Return to E1 until \( r = 0 \).
- **E3.** Repeat the previous two steps until \( f = 0 \).

C5. **Compute expected outcomes given \( k \).** Set \( t = t_0 \) and restore \( \Omega \).

- **F0.** Loop through \( \theta_{\text{end}} \) and setting the clock to \( \hat{\theta}_{\text{clock}} \). Compute \( E[Y | k, \theta_{\text{cond}}] \) and update \( \Omega \) for the next period.
- **F1.** Increase \( t \) by 1. Repeat previous step until \( t > t_{\text{max}} \).

B1. Decrease \( g \) by 1. If \( g > 0 \) set \( f = F \) and return to section E.

B2. Decrease \( e \) by 1. If \( e > 0 \) then reset \( g = 2 \) and return to section D.

B3. Decrease \( k \). If \( k > 0 \) return to B0.

B4. **Compute empirical predictions.** Set \( e = E \), \( g = G \), \( t = t_0 \), and \( k = K \).

- **G0.** Loop over \( k \) to compute the sample-selected mixture for values of \( t, e, \) and \( g \) that apply for \( d \).

- **G1.** Compute the contribution \( \hat{\Delta}(\theta_{\text{cond}}) \) to the econometric objective as defined in (24).

- **G2.** Iterate on \( t \) through \( t_{\text{max}} \) then decrease \( g \) and \( e \) until 0.

A3. Accumulate \( W(\theta_{\text{exog}}) \). Decrease \( d \). If \( d > 0 \) return to step A2.

A4. Use the optimizer to minimize the objective with respect to \( \theta_{\text{exog}} \).

A5. Iterate on the weighting matrix \( \Sigma \), return to previous steps to compute \( \hat{\theta}_{\text{exog}} \).

---

**IX.F GMM Estimation Procedure**

**IX.F.1 First Stage**

The weighted discrepancy between the data and the model used in the first stage is

\[
Z^1(\theta_{\text{exog}}) = \sum_{d=1}^{4} \sum_{e=1}^{2} \sum_{g=1}^{3} \sum_{t=0}^{t_{\text{max}}(e)} \frac{n(\theta_{\text{cond}})}{265159} \hat{\Delta}(\theta_{\text{cond}})' \Sigma_0 \hat{\Delta}(\theta_{\text{cond}}),
\]

where \( \Sigma_0 \) is a 12 \( \times \) 12 diagonal matrix with elements listed in Table 4. For the monetary values the weights are the inverse of the grand mean of the moment over conditioning states. For the binary variables a weight of \( 1/ \ 5 = 2.0 \) was chosen to avoid putting excessive weight on turnover values which are near 0 and noisy across months. The cell sizes \( n_{\text{cond}} \) (in Table B.11) sum to 265,159 in (A39). The Appendix discusses how variation across samples, treatments, provinces, experimental time and elements of the measurement vector \( Y(\alpha, \theta) \) contribute to the identification of parameters of the model.
Let $\hat{\theta}_{\text{exog}}^1$ denote the parameters chosen to minimize $Z^1$. From these estimates the covariance matrix of the moments is computed. Given the random assignment to groups and the assumption that demographic groups are different (exogenous) mixtures across types, the moments are uncorrelated across groups defined by $e$, $g$ and $d$. That is, the sequence of observed vectors $Y(\alpha, \theta)$ for an individual is correlated, but across entry, treatment and demographic groups the sequence of individual shocks are independent. The population covariance matrix of moments is block diagonal with non-zero entries only across $t$. There are 14 blocks varying in size between 326 and 417. To compute the covariance matrix, individual paths of $Y$ are simulated from $\hat{\theta}_{\text{exog}}^1$ following the model. First, initial states are drawn from the ergodic distribution. Then an action vector is drawn from the choice probabilities and finally a new state is drawn from the primitive transition. Let $Y^r_i(\theta_{\text{cond}})$ denote the $r$th simulated path with vectors concatenated across experimental time $t$. Then the deviation of the path from the mean $E_i[Y | \theta_{\text{cond}}]$ is computed and weighted by the endogenous type proportion for the type within the sample. The outer product of the vector of deviations is computed and averaged across simulations. The resulting matrix is a consistent estimate of the covariance of the block of moments for the group $\theta_{\text{cond}}$. The inverse of the matrix is computed for each block:

$$
\Sigma(\theta_{\text{cond}}) = \left( \frac{1}{R} \sum_{r=1}^{R} \sum_{k=1}^{4} \lambda^r(\theta_{\text{cond}}) (Y_i^r(\theta_{\text{cond}}) - E_i[Y | \theta_{\text{cond}}]) (Y_i^r(\theta_{\text{cond}}) - E_i[Y | \theta_{\text{cond}}])' \right)^{-1}.
$$

Based on this new weighting of the moments the parameter fit changed a great deal. Therefore, the correlation matrices is computed once more from the new parameter values. The final stage objective is:

$$
Z^2(\theta_{\text{exog}}) = \sum_{d=1}^{4} \sum_{e=1}^{2} \sum_{g=1}^{3} \Delta_i(\theta_{\text{cond}})' \Sigma(\theta_{\text{cond}}) \Delta_i(\theta_{\text{cond}}).
$$

The GMM estimates are then

$$
\hat{\theta}_{\text{exog}} = \arg \min_{\theta_{\text{exog}}} Z^2(\theta_{\text{exog}}). \tag{A40}
$$

Let $D(\theta_{\text{cond}})$ denote the matrix of gradients for the vector $\Delta_i$ with respect to the estimated parameters. The estimated variance matrix and standard errors were computed using the standard formula

$$
\text{Var} \left[ \hat{\theta}_{\text{exog}} \right] = \left\{ \frac{1}{4} \sum_{d=1}^{4} \sum_{e=1}^{2} \sum_{g=1}^{3} D(\theta_{\text{cond}}) \Sigma(\theta_{\text{cond}}) D(\theta_{\text{cond}})' \right\}^{-1}. \tag{A41}
$$

### IX.G Identification

The estimated parameters are identified from three sources of variation:

- Controlled and time-varying (path of treatment and assignment to experimental group)
- Uncontrolled and time-invariant (variation in policy and demographic groups).
Uncontrolled and time-varying (unobserved endogenous states and treatment status)

The first two sources are captured in the vector of conditioning variables $\theta_{\text{cond}} = (t, g, e, d)$. Different loadings on these three factors will produce different patterns within months (across contemporaneous moments), across months (progress of treatment and initial selection), across studies (differing selection and information), across treatment groups (impact), and across demographic groups (variation in the mixture across exogenous types). It is not possible to prove analytically that the estimated parameters are identified from data generated by the experiment. Instead, a heuristic argument is given. The sources of variation are appealed to roughly in the order given above.

Begin with the case of no unobserved heterogeneity ($K = 1$) and a simple parameter to identify, the job-loss probability $\pi_t$. In the model job loss occurs exogenously and the SSP survey records reasons why a parent stop working. These were grouped into losses and quits as reported in Table 4. Thus the proportion of working parents losing a job each month is available in the data and is directly determined by the value of $\pi_t$. Since the observed proportions differ across demographic groups it is feasible to consider unobserved heterogeneity in $\pi_t$ with different mixtures across groups. Of course, the estimates of $\pi_t$ enters into all other aspects of the model.

Parents in the control group receiving IA do not quit jobs unless the convexity parameter $c(h)$ changes value. And some parents go on and off IA with no change in labor market status, which occurs in the model only when the level of outside support changes. The measurement vector includes quits and IA status but not these conditional switch rates. However, the joint movement over time (within control groups) of IA, labor market status, and quits help identify the jump probability for $h$ the jump probability for outside support, $\pi_s$. How the quit rate correlates with labor market earnings helps identify the distribution of $c(h)$ and thus $\zeta$. Mean earnings and the square of mean earnings are included in $Y(\alpha, \theta)$ so that two moments of the accepted distribution are available to match the mean and variance of the offer distribution. Wage growth and duration dependency in accepted starting wages identify the skill accumulation and depreciation parameters. The correlation between income and welfare benefits helps identify the income reporting rate.

In a stationary model estimated on non-experimental data, the job search parameters (cost of search, offer probability, proportion of full-time jobs) would have to be identified through the reservation wage and the proportion of households working part-time (along with parametric assumptions on the offer distribution already made). It is not guaranteed that they would be identified in such data. The SSP experiment, however, includes exogenous variation in the value of job search and the value of keeping a full-time job. For example, the change in the proportion of people working part-time in the first month of the SSP (relative to the controls) picks up the proportion of accepted jobs that are potentially full-time.

Now consider more subtle variation across the Applicant ($e = 1$) and Recipient ($e = 2$) samples. An impact study focuses on differences between a treatment group and their matched control group. For the Applicant Study, this consists of those who know the SSP subsidy exists and can anticipate becoming eligible for it (i.e., they are in phase $f = 1$), and those in the control group who cannot become eligible ($f = 0$). The model makes
clear predictions between the behavior of these two groups. The value of taking a job and/or leaving IA changes with the time spent in phase 1. As \( r \), the months residing in the phase, approaches \( R(1) \) the higher the value of continued receipt of IA becomes among the treated. The rate at which outcomes diverge across the two groups as \( r \) increases reflects this approach to the change in phase. The change in the value of IA across groups as \( R(1) \) approaches is sensitive to the transition probabilities. For example, high offer probabilities imply the treatment group can afford to reject offers received earlier and/or cease active job search. The pattern of impacts helps identify these probabilities, although there is no one observable difference that can be matched to each parameter.

Treated households in the Applicant and Recipient Studies are in identical situations if and when they reach the qualifying phase of the experiment (\( f = 2 \)). From that point on, any difference between the behavior of the eligible households within the two groups is, within the model, forced to come from the difference in household states conditional upon reaching phase 2. In the Recipient Study reaching phase 2 is exogenous to the SSP and unexpected, whereas for the Applicant Study it is completely endogenous and can be expected and partially controlled up to one year in advance. Thus, the two samples provide experimental variation in unobserved household states caused by lagged decisions made while anticipating different future opportunities. Many model parameters affect this cross-sample variation. For example, if job offers are rare then parents in the Applicant Study may not respond strongly to the information they have relative to the Recipient Study before assignment. As argued above, other variation in the data contribute to identifying parameters like job offer rates. For purposes of this discussion, if we treat the other parameters as identified without comparing the entry and applicant treatment groups, then their comparison reveals the discount factor \( \delta \).

The final parameter to discuss is the smoothing factor \( \rho \). When \( \rho = 0 \) each feasible action has equal probability independent of the household’s state. This allows for a conclusion of completely ‘irrational’ behavior to be drawn from the data. The estimated model avoids this result because it is required to match the overarching patterns across groups and across experimental states that indicate systematic variation in choice probabilities across states. For example, under complete irrationality, the proportion of households receiving IA each month would be the same no matter the assigned treatment group or how long ago random assignment occurred. Since statistically significant differences in choice probabilities exist across groups and experimental time, the estimated parameters will choose \( \rho > 0 \).

The point of the discussion so far is that each of the 19 exogenous parameters interacts with the design of the SSP experiment to affect specific aspects of the 12 matched results. The arguments account for the presence of many unobserved endogenous states, but they do not as yet account for unobserved exogenous parameters. Identification of unobserved heterogeneity in the parameters would be strengthened by applying the model to individual outcomes, because the likelihood or the predicted moments for a single individual would be conditioned on a single type. The computational cost of imposing these additional requirements is, however, prohibitive.

Recall that demographic variation plays a restrictive role in the model. It determines the value of the policy parameters, such as the level of IA benefits, which are pre-determined and not free to explain variation in the data. The behavior of the unobserved types will respond to the differences in the policy parameters but there are no free
parameters that directly control the influence of the demographic variables on predictions. That is, there is nothing like a ‘provincial coefficient’ in the wage offer distribution or a ‘number of children’ coefficient in the cost of time. Therefore, the model greatly restricts the freedom to calibrate responses in order to match the wide variation in experimental results across demographic groups. The only way for the estimates to gain more leverage in explaining the wide variation across demographic groups is to allow variation in the within-group proportions of each type. Thus it is likely (but not obvious how to demonstrate ahead of time) that the mixture parameters \( \Lambda \) will be identified from the data along with differences in the underlying parameter vectors \( \Gamma[\kappa] \).

**IX.H Supplementary Material**

As of the writing of this draft, computer programs, data and output are available from [www.econ.queensu.ca/~ferrall/papers/SSPinformation](http://www.econ.queensu.ca/~ferrall/papers/SSPinformation).
X. References


### Table 1. Endogenous Variables and Actions

<table>
<thead>
<tr>
<th>Item</th>
<th>Variable</th>
<th>Description</th>
<th>Num.</th>
<th>Values / Calculation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L</td>
<td>Lost job entering this month</td>
<td>2</td>
<td>{0,1}</td>
<td>does not affect utility or transitions</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>worked Previous month</td>
<td>2</td>
<td>{0,1}</td>
<td>does not affect utility or transitions</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>current earnings offer</td>
<td>6</td>
<td>{0,1/5,...,4/5}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>x</td>
<td>experience level</td>
<td>4</td>
<td>{1/4,1/2,3/4,1}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>θ_end</td>
<td>upper bound on hours in job</td>
<td>3</td>
<td>{0,1,2}</td>
<td>Figure 1</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>Outside Support</td>
<td>4</td>
<td>{0,1/3,2/3,1}</td>
<td>Figure 1</td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>Opp. cost of time outside Household</td>
<td>3</td>
<td>{1/4,2/4,3/4}</td>
<td>Figure 1</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>Demographic group</td>
<td>4</td>
<td>{1,2,3,4}</td>
<td>Table 2</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>unobserved type</td>
<td>4</td>
<td>{1,2,3,4}</td>
<td>Table 7</td>
</tr>
<tr>
<td></td>
<td>Θ_end</td>
<td>Real states for individual (S₁)</td>
<td>2,304</td>
<td>= 2<em>4</em>6<em>4</em>4*3</td>
<td>l &amp; b stored as 1 var. w/ 4 values</td>
</tr>
<tr>
<td></td>
<td>xΘ_clock</td>
<td>All states given assignment (S₂)</td>
<td>138,240</td>
<td>= S₁ * 12 * 5</td>
<td>Cond. on ne &amp; ng (group assignment)</td>
</tr>
<tr>
<td></td>
<td>xΘ_exp</td>
<td>All individual states (S₃)</td>
<td>417,024</td>
<td>= S₁ + S₂ * 3</td>
<td>Control+SSP+SSP_Plus+Applicant</td>
</tr>
<tr>
<td></td>
<td>xKxD</td>
<td>Complete State Space (S)</td>
<td>6,672,384</td>
<td>= S₃ * 4 * 4</td>
<td>K * D</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>labor Market work hours</td>
<td>5</td>
<td>{0,1,4,1/2,3/4,1}</td>
<td>constrained by u(b); see Figure 1</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>engage in Active job search</td>
<td>2</td>
<td>{0,1}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>accept IA</td>
<td>2</td>
<td>{0,1}</td>
<td>see Figure 1.</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>Feasible Action Space (A)</td>
<td>12</td>
<td>= 6*2</td>
<td>m &amp; a stored as one var. with 6 values</td>
</tr>
<tr>
<td></td>
<td>(A,Θ)</td>
<td>Outcome Space</td>
<td>80,068,608</td>
<td>= S * A</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Demographic, Treatment, and Experimental Groups

<table>
<thead>
<tr>
<th>Vector</th>
<th>Index</th>
<th>Description</th>
<th>Subjects</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>1</td>
<td>New Brunswick, 1 Child</td>
<td>1728</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>New Brunswick, 2+ Children</td>
<td>1217</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>British Columbia, 1 Child</td>
<td>3058</td>
<td>34%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>British Columbia, 2+ Children</td>
<td>2895</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>8898</td>
<td>100%</td>
</tr>
</tbody>
</table>

| g      | 3     | Control                               | 4305     | 48%        |
|        | 2     | SSP Treatment                         | 4300     | 48%        |
|        | 1     | SSP+ Treatment (NB only)              | 293      | 3%         |
|        |       | Total                                 | 8898     | 100%       |

| e      | 2     | Recipient Study                       | 5682     | 63%        |
|        | 1     | Applicant Study (BC only)             | 3316     | 37%        |
|        |       | Total                                 | 8998     | 100%       |

Observations dropped: invalid or missing age, high school attendance or number of children (from IA records).
Table 3. Experimental Results (Moments) Selected for Matching

<table>
<thead>
<tr>
<th>Var.</th>
<th>Description</th>
<th>Model</th>
<th>Unit</th>
<th>Count</th>
<th>Mean</th>
<th>St.Dev</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>earn</td>
<td>Rep. Earnings</td>
<td>$(1-\beta_i)mW(\alpha,\theta)$</td>
<td>$100$</td>
<td>450</td>
<td>3.439</td>
<td>1.564</td>
<td></td>
</tr>
<tr>
<td>earnsq</td>
<td>Earnings Sq.</td>
<td>$\text{earn}^2$</td>
<td>$100^2$</td>
<td>470</td>
<td>62.590</td>
<td>43.470</td>
<td></td>
</tr>
<tr>
<td>ia</td>
<td>IA Received</td>
<td>$\text{IA}(\alpha,\theta)$</td>
<td>$100$</td>
<td>466</td>
<td>5.966</td>
<td>1.869</td>
<td>Fwd.2 mth</td>
</tr>
<tr>
<td>iasq</td>
<td>IA Receiv Sq.</td>
<td>$\text{IA}^2$</td>
<td>$100^2$</td>
<td>466</td>
<td>57.782</td>
<td>27.593</td>
<td>Fwd.2 mth</td>
</tr>
<tr>
<td>gsu</td>
<td>SSP Suppl</td>
<td>$\text{SUP}(\alpha,\theta)$</td>
<td>$100$</td>
<td>240</td>
<td>1.530</td>
<td>0.600</td>
<td>Fwd.2 mth</td>
</tr>
<tr>
<td>onia</td>
<td>Received IA</td>
<td>i</td>
<td>0/1</td>
<td>470</td>
<td>0.708</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>mwg</td>
<td>Worked at MW</td>
<td>$(n*&lt;6-#n)(m&gt;0)$</td>
<td>0/1</td>
<td>470</td>
<td>0.777</td>
<td>0.059</td>
<td>hrly w &lt;= MW+$0.10</td>
</tr>
<tr>
<td>leftjb</td>
<td>Left/quit a job</td>
<td>p(l=0)(m=0)</td>
<td>0/1</td>
<td>456</td>
<td>0.003</td>
<td>0.004</td>
<td>Excl. job-to-job</td>
</tr>
<tr>
<td>lossjb</td>
<td>Loss a job</td>
<td>l</td>
<td>0/1</td>
<td>456</td>
<td>0.004</td>
<td>0.004</td>
<td>Excl. job-to-job</td>
</tr>
<tr>
<td>emft</td>
<td>Full Time</td>
<td>m&gt;PT</td>
<td>0/1</td>
<td>470</td>
<td>0.223</td>
<td>0.088</td>
<td></td>
</tr>
<tr>
<td>empt</td>
<td>Part-time</td>
<td>0 &lt; m &lt;= PT</td>
<td>0/1</td>
<td>470</td>
<td>0.130</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>onXem</td>
<td>IA &amp; Working</td>
<td>$\text{i} \times (m&gt;0)$</td>
<td>0/1</td>
<td>470</td>
<td>0.161</td>
<td>0.045</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>4884</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A summary of the complete data listed in Table B.1-B.10. Count is the number of cells in Table B panel. Mean and standard deviation are across cells not individuals. #n denotes the order of n in the feasible set. For example, #0 = 1, #1/6 = 2, etc.
Table 4. Relative Impacts on Selected Moments in Months -11,1,13,25

<table>
<thead>
<tr>
<th>Var.</th>
<th>t</th>
<th>NB / 1 Child Recipients</th>
<th>NB / 2+ Recipients</th>
<th>BC / 1 Child Appl. Recipients</th>
<th>BC / 2+ Appl. Recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SSP+</td>
<td>SSP</td>
<td>SSP+</td>
<td>SSP</td>
</tr>
<tr>
<td>Earn</td>
<td>-10</td>
<td>1</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>0.51</td>
<td>0.39</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>0.58</td>
<td>0.30</td>
<td>0.74</td>
</tr>
<tr>
<td>OnIA</td>
<td>-11</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>(0.24)</td>
<td>(0.19)</td>
<td>(0.29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>(0.32)</td>
<td>(0.18)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Mwgm</td>
<td>-11</td>
<td>1</td>
<td>0.00</td>
<td>(0.01)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>(0.13)</td>
<td>(0.11)</td>
<td>(0.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Emft</td>
<td>-11</td>
<td>1</td>
<td>0.11</td>
<td>(0.07)</td>
<td>(0.09)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>0.97</td>
<td>0.76</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>0.90</td>
<td>0.64</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Difference between SSP and Ctrl columns in Table A divided by Ctrl column. Negative impacts in () and in red. Largest absolute impact within the table shaded for each moment.
Table 5. Contemporaneous Correlations Across Results

<table>
<thead>
<tr>
<th>Applicants (e=1)</th>
<th>earn</th>
<th>ia</th>
<th>onia</th>
<th>mwg</th>
<th>left</th>
<th>emft</th>
<th>Group (g); Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>earn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ia</td>
<td>-0.360</td>
<td>-0.314</td>
<td>-0.678</td>
<td>-0.029</td>
<td>0.655</td>
<td></td>
<td>(3):42,056</td>
</tr>
<tr>
<td>onia</td>
<td>-0.356</td>
<td>0.733</td>
<td>0.360</td>
<td>-0.002</td>
<td>-0.375</td>
<td></td>
<td>(2); 40,875</td>
</tr>
<tr>
<td>mwg</td>
<td>-0.303</td>
<td>0.720</td>
<td>0.302</td>
<td>-0.011</td>
<td>-0.338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>left</td>
<td>-0.668</td>
<td>0.379</td>
<td>0.303</td>
<td>0.047</td>
<td>-0.677</td>
<td></td>
<td></td>
</tr>
<tr>
<td>emft</td>
<td>-0.029</td>
<td>0.008</td>
<td>0.004</td>
<td>0.050</td>
<td>-0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.640</td>
<td>-0.400</td>
<td>-0.353</td>
<td>-0.672</td>
<td>-0.044</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Recipients (e=2)</th>
<th>earn</th>
<th>ia</th>
<th>onia</th>
<th>mwg</th>
<th>left</th>
<th>emft</th>
<th>Group (g); Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>earn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ia</td>
<td>-0.317</td>
<td>-0.294</td>
<td>-0.550</td>
<td>-0.011</td>
<td>0.564</td>
<td></td>
<td>(3); 95,302</td>
</tr>
<tr>
<td>onia</td>
<td>-0.409</td>
<td>0.692</td>
<td>0.330</td>
<td>-0.018</td>
<td>-0.369</td>
<td></td>
<td>(2); 96,220</td>
</tr>
<tr>
<td>mwg</td>
<td>-0.396</td>
<td>0.733</td>
<td>0.267</td>
<td>-0.021</td>
<td>-0.324</td>
<td></td>
<td></td>
</tr>
<tr>
<td>left</td>
<td>-0.608</td>
<td>0.402</td>
<td>0.376</td>
<td>0.024</td>
<td>-0.576</td>
<td></td>
<td></td>
</tr>
<tr>
<td>emft</td>
<td>-0.018</td>
<td>-0.009</td>
<td>-0.027</td>
<td>0.031</td>
<td>-0.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.638</td>
<td>-0.503</td>
<td>-0.511</td>
<td>-0.611</td>
<td>-0.030</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 6. Summary of the Estimation

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of linear system to compute ergodic distribution</td>
<td>2,304</td>
<td>See Table 1</td>
</tr>
<tr>
<td>Number of Type-Specific Parameters (N)</td>
<td>16</td>
<td>Table 7.1-8.4</td>
</tr>
<tr>
<td>Number of Common Parameters (C)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Number of free exogenous parameters</td>
<td>79</td>
<td>(D^* (K-1) + K\times N + C)</td>
</tr>
<tr>
<td>CPU Time to Evaluate Objective (min.)</td>
<td>14</td>
<td>16 X UltraSPARC-III</td>
</tr>
<tr>
<td>Value of Objective ((Z^2))</td>
<td>19.426</td>
<td></td>
</tr>
</tbody>
</table>
Table 7.1. $\hat{\theta}_{\text{exog}}$ : Estimated Type Proportions ($\Lambda[d]$)

<table>
<thead>
<tr>
<th>d</th>
<th>Description</th>
<th>Type Index (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>NB, One Child</td>
<td>0.0063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
</tr>
<tr>
<td>2</td>
<td>NB, Two+ Children</td>
<td>0.000003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.860)</td>
</tr>
<tr>
<td>3</td>
<td>BC, One Child</td>
<td>0.5228</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.438)</td>
</tr>
<tr>
<td>4</td>
<td>BC, Two+ Children</td>
<td>0.5736</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

GMM estimates based on (40). See (5) and (16) for roles of the parameters. Estimated standard errors in parentheses is the square root of diagonal elements of (41).
Table 7.2. $\hat{\theta}_{\text{exog}}$ : Estimated Dynamic Programming Parameters ($\delta_k$ and $\rho_k$)

<table>
<thead>
<tr>
<th>Var</th>
<th>Description</th>
<th>Type Index (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Discount Factor</td>
<td>0.9999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Income Reporting</td>
<td>0.399</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Smoothing</td>
<td>36.220</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.081)</td>
</tr>
</tbody>
</table>

GMM estimates based on (40). See (5) and (16) for roles of the parameters. Estimated standard errors in parentheses is the square root of diagonal elements of (41).
Table 7.3. $\hat{\theta}_{\text{exog}}$ : Estimated Utility Shifters ($\tau$)

<table>
<thead>
<tr>
<th>Var.</th>
<th>Description</th>
<th>Type Index (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$\mu$ Job Offer Mean</td>
<td>-1.560</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.045)</td>
</tr>
<tr>
<td>$\sigma$ Job Offer St. Dev.</td>
<td>1.999</td>
<td>1.632</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.014)</td>
</tr>
<tr>
<td>$\chi$ Outside Support</td>
<td>1.427</td>
<td>1.080</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.037)</td>
</tr>
<tr>
<td>$\nu$ Cost of FT Work</td>
<td>0.346</td>
<td>0.409</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>$\kappa$ Cost of Job Search</td>
<td>0.461</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.036)</td>
</tr>
<tr>
<td>$\eta$ Return to Skill</td>
<td>1.355</td>
<td>2.964</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.155)</td>
</tr>
<tr>
<td>$\zeta$ 1 / Mean Convexity</td>
<td>2.997</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GMM estimates based on (40). See (5) and (16) for roles of the parameters. Estimated standard errors in parentheses is the square root of diagonal elements of (41).
Table 7.4. $\hat{\theta}_{\text{exog}}$: Estimated Transition Shifters (II)

<table>
<thead>
<tr>
<th>Sub.</th>
<th>Description</th>
<th>Type Index (k)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>Job Offer (b&gt;0)</td>
<td></td>
<td>0.069</td>
<td>0.730</td>
<td>0.99994</td>
<td>0.831</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.510)</td>
<td>(0.060)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>f</td>
<td>Prop. Full Time</td>
<td></td>
<td>0.999996</td>
<td>0.889</td>
<td>0.99955</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.000)</td>
<td>(0.080)</td>
<td>(0.206)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>m</td>
<td>Prop. MW job (n=0)</td>
<td></td>
<td>0.131</td>
<td>0.439</td>
<td>0.531</td>
<td>0.249</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.016)</td>
<td>(0.321)</td>
<td>(0.512)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>l</td>
<td>Job Loss</td>
<td></td>
<td>0.021</td>
<td>0.0011</td>
<td>0.018</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Home</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>Support Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>h</td>
<td>Prob. Costs Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.999989</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>SSP Plus Effect</td>
<td></td>
<td>0.823</td>
<td>0.009</td>
<td>0.641</td>
<td>0.910</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
<td>(0.255)</td>
<td>(1.060)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>a</td>
<td>Accumulation</td>
<td></td>
<td>0.995</td>
<td>0.4787</td>
<td>0.1245</td>
<td>0.2601</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.038)</td>
<td>(0.898)</td>
<td>(0.071)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>d</td>
<td>Depreciation</td>
<td></td>
<td>0.0006</td>
<td>0.0037</td>
<td>0.7654</td>
<td>0.1007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.021)</td>
<td>(1.066)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

GMM estimates based on (40). See (5) and (16) for roles of the parameters. Estimated standard errors in parentheses is the square root of diagonal elements of (41).
Table 8. Estimated Standard Errors by Group

<table>
<thead>
<tr>
<th>Par./Sub.</th>
<th>k=1 / common</th>
<th>k=2</th>
<th>k=3</th>
<th>k=4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all</td>
<td>Ctrl</td>
<td>Treat</td>
<td>all</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.0005</td>
<td>1.5261</td>
<td>0.0003</td>
<td>0.0811</td>
</tr>
<tr>
<td>k=1</td>
<td>0.0363</td>
<td>41.9370</td>
<td>0.0269</td>
<td>0.3497</td>
</tr>
<tr>
<td>k=2</td>
<td>0.0449</td>
<td>11.4133</td>
<td>0.0329</td>
<td>0.1543</td>
</tr>
<tr>
<td>k=3</td>
<td>0.0135</td>
<td>2.2822</td>
<td>0.0098</td>
<td>0.0688</td>
</tr>
<tr>
<td>k=4</td>
<td>0.0163</td>
<td>32.5299</td>
<td>0.0106</td>
<td>0.3211</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.0150</td>
<td>------</td>
<td>------</td>
<td>0.2552</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.0379</td>
<td>39.1994</td>
<td>0.0273</td>
<td>0.8979</td>
</tr>
<tr>
<td>( \mu )</td>
<td>0.0008</td>
<td>0.0192</td>
<td>0.0006</td>
<td>0.0210</td>
</tr>
<tr>
<td>( \nu )</td>
<td>0.0027</td>
<td>2.5649</td>
<td>0.0019</td>
<td>0.5096</td>
</tr>
<tr>
<td>( \xi )</td>
<td>0.0100</td>
<td>0.0060</td>
<td>0.0005</td>
<td>0.0065</td>
</tr>
<tr>
<td>( \chi )</td>
<td>0.0371</td>
<td>1.3163</td>
<td>0.0251</td>
<td>3.6653</td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.0001</td>
<td>0.0192</td>
<td>0.0001</td>
<td>0.0800</td>
</tr>
<tr>
<td>( \eta )</td>
<td>0.0043</td>
<td>0.5478</td>
<td>0.0032</td>
<td>0.7668</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.1551</td>
<td>371.7602</td>
<td>0.1108</td>
<td>30.4060</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.8912</td>
<td>1114.7384</td>
<td>0.7502</td>
<td>143.9372</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.0040</td>
<td>0.0389</td>
<td>0.0029</td>
<td>0.0395</td>
</tr>
<tr>
<td>( \zeta )</td>
<td>0.0912</td>
<td>3.3796</td>
<td>0.0498</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5a. Result: Earnings, New Brunswick

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 6a. Result: On IA, New Brunswick

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 7a. Result: Total Government Transfers (IA+SSP)

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 7b. Result: Total Government Transfers (IA+SSP)

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 8. Variation from Policy, Selection and Heterogeneity

earn for $k=1$

New Brunswick 1 Child

New Brunswick 2+ Children

British Columbia 1 Child

British Columbia 2+ Children

Type 1 under each policy; triangle represents the ergodic (unselected) mean

earn for $k=2$

New Brunswick 1 Child

New Brunswick 2+ Children

British Columbia 1 Child

British Columbia 2+ Children

Type 2 under each policy; triangle represents the ergodic (unselected) mean
Figure 9a. Forecast: Earnings by Province

New Brunswick Pooled

British Columbia Pooled

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 9b. Forecast: On IA by Province

New Brunswick Pooled

British Columia Pooled

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
Figure 12. Experiment 3: Total Transfers under Short-Lived, Large, Flat Bonus

C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment
C: Recipient control; CA: Applicant Control;
T: Recipient treatment; A: Applicant treatment; +: Plus Treatment