SRDC Working Paper Series 06-09

The Effect of the Self-Sufficiency Project on Children

Piotr Wilk
Middlesex-London Health Unit
and
University of Western Ontario

Michael H. Boyle
McMaster University

Martin D. Dooley
McMaster University

Ellen L. Lipman
McMaster University

March 2006

SOCIAL RESEARCH AND DEMONSTRATION CORPORATION

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

Table of Contents

Tables and Figures

Executive Summary

The Self-Sufficiency Project (SSP) was a random assignment study of welfare and work policies. Its objective was to test whether welfare-to-work programs are able to facilitate a successful transition into full-time employment by allowing single parents to gain access to increased financial resources. Adult-oriented welfare-to-work programs have the potential to alter important aspects of family life. Transition into full-time employment and increased financial resources are thought to influence the quality and reliability of child care, the physical and emotional health of the parents, and the quality of interactions between parents and children. Hence, from the policy point of view, it is important to determine the extent to which SSP affected children.

The objective of the current study is to assess whether SSP affected the health, behaviour, and academic achievement of children and whether these effects were sustained beyond the period of the intervention. The impact of welfare-to-work programs such as SSP is a function of two processes: program offer and program take-up. Only about one third of the eligible families in SSP took advantage of the earnings supplement offer. If the observed program effects are disproportionately confined to those children whose parents accepted the offer, then shifts in program take-up could affect the net impact of the program in the general population. Because of uncertainty about supplement use in real world circumstances — when (and if) SSP is introduced to the general population — this study assesses both the overall effect of the program intervention on all children in the program group and the program impact on those children whose parents decided to take advantage of the program. The results obtained through each of these two analytical approaches might be of potential interest, and when combined together, the combined results can better inform policy-makers about the effect of SSP on children.

To estimate the effect of the program intervention on all children from the program group, we relied on the standard intent-to-treat (ITT) analysis, in which children's outcomes are compared by their parents' random assignment into the program group and the control group. To estimate the effect of the receipt of the earnings supplement, we employed the Complier Average Causal Effect (CACE) estimation approach. The CACE methodology compares the outcomes of program group children whose parents took the earnings supplement offer with the outcomes of children in the control group whose parents would have taken the offer, had they been assigned to the program group.

The results from both analyses indicate that children's health, behaviour, and academic achievement were not affected by the SSP program intervention, whether their parents took advantage of the earnings supplement offer or not. In other words, children were not harmed, nor did they benefit from the program intervention. Taking into account how young some of these children were at the time of random assignment, it is reassuring that the increases in full-time employment did not result in negative effects for these children. From the policy point of view, the results of this evaluation suggest that when (and if) the SSP-like welfare-to-work program is introduced to the general population, it is unlikely to affect children's health, behaviour, and academic achievement, regardless of the level of program take-up.

Introduction

One in five Canadian children lives in a single-parent family (Canadian Institute of Child Health, 2000). Over 80 per cent of these families are led by single mothers, and at any given time about 40 per cent of them are receiving social assistance or welfare (Statistics Canada, 1996). Children growing up in these families constitute a large proportion of those at risk for diminished life quality attributable to socio-economic disadvantage (Lipman, Boyle, Dooley, & Offord, 2002; Lipman, Secord, & Boyle, 2001). A large body of literature suggests that low family income, particularly chronic poverty, is related to children's health, intellectual functioning, academic achievement, social behaviour, and psychological well-being (Brooks-Gunn & Duncan, 1997; Duncan & Brooks-Gunn, 1997; Duncan, Brooks-Gunn, & Klebanov, 1994; Ford, Gyarmati, Foley, Tattrie, & Jimenez, 2003; Hill & Sandfort, 1995; Korbin, 1992; McLoyd, 1998; Michalopoulos et al., 2002; Morris & Michalopoulos, 2000).

Welfare-to-work programs (earnings supplement programs) have been identified as an innovative approach to the dual problem of welfare dependency and poverty. These programs offer a temporary earnings supplement to long-term income assistance (IA) recipients who are willing to leave welfare for full-time work. Introduced in 1992, the Self-Sufficiency Project (SSP) was a random assignment study of welfare and work policies. The project was conceived and funded by Human Resources Development Canada and managed by the Social Research and Demonstration Corporation. Its objective was to test whether welfare-to-work programs are able to facilitate a successful transition into full-time employment by allowing families to gain access to increased financial resources. An equally important goal of SSP was to assess whether earnings supplement programs affect children and whether these effects are sustained beyond the period of the intervention. To this end, SSP collected detailed information on children of program participants during the period when their parents were eligible for the earnings supplement and after their parents' eligibility for the supplement had ended.

The overall objective of the current study is to determine whether SSP affected the health, behaviour, and academic functioning of children and whether these effects were sustained beyond the period of the intervention. SSP employed a random assignment evaluation design, which is considered one of the most reliable techniques to measure program impacts. However, only about one third of eligible families in SSP took advantage of the earnings supplement offer. The focus of this study is to assess the effect of the receipt of SSP's earnings supplement. Using the phrase of Heckman, Smith, and Taber (1998), to assess the impact of the earnings supplement use, we need to estimate "the effect of treatment on the treated" (in Dunn et al., 2003).

Although analyses of subgroups participating in experiments pose difficult selection problems, it is important from a policy perspective to estimate not only the net program impacts attributable to experimental manipulation, but also the program effects among those who would take the program intervention if offered. In the present context, this means an examination of the extent to which variability in child outcomes is associated with the program take-up. Uncertainty about supplement use in real world circumstances — when (and if) SSP is introduced to the general population — lends urgency to this study. The

impact of welfare-to-work programs such as SSP is a function of two processes: program offer and program take-up. Since only the minority of the eligible families in SSP took the earnings supplement, any observed program effects on child outcomes are likely to be confined to these particular families. If SSP had little or no effect on the children whose parents did not take the supplement, "then changes in children's outcomes for those families who did take the supplement must have been much larger than the effects of SSP overall" (Morris & Michalopoulos, 2000). If the observed program effects are disproportionately confined to those children whose parents accepted the program offer, then shifts in program take-up could markedly affect the net impact of the program in the general population. It is arguable that program take-up in natural circumstances might be different, conceivably higher from that observed in the SSP experiment. This might be due to differences in conditions of local labour markets, cultural settings, or the levels of public awareness about the program. The level of program take-up might also be related to the artificial conditions associated with running an experiment, which no longer apply when programs are made available to the general population.

From the policy point of view, it is essential to determine the extent to which adult-oriented welfare-to-work programs are affecting children. Furthermore, we strongly believe that it is important to assess both the overall effect of the program intervention on all children in the program group and the program impact on those children whose parents decided to take the program offer. The results obtained through each of these two analytical approaches might be of potential interest, and when combined together, they can better inform policy-makers about the effect of SSP on children. Consequently, we conducted both types of analyses.

The Self-Sufficiency Project

OVERVIEW OF SSP

The Self-Sufficiency Project (SSP) was a three-study research and demonstration project that offered a temporary earnings supplement to long-term, single parent welfare recipients who were willing to leave income assistance (IA) for full-time work. In the Recipient study, the main SSP study, randomly selected long-term welfare recipients in British Columbia and New Brunswick were offered a temporary earnings supplement that consisted of a monthly cash payment. The supplement was available to single parents who had been on welfare for at least one year. To receive the supplement, participants had to find a full-time job within one year of random allocation. The supplement was paid on top of earnings from employment for up to three years in individual months when the participants were employed full-time and remained off IA. The Applicant study, one of the two sub-studies within SSP, targeted single parents in British Columbia who had just begun to receive IA. To establish eligibility for the earnings supplement, single parents who entered the study had to stay on IA for a year; if they left IA before one year, they could never receive the supplement. Then, in order to receive the supplement, parents had to find a full-time job within one year. Finally, the SSP Plus study, the other sub-study within SSP, examined the effect of combining the earnings supplement with other services. In that study, participants were randomly assigned to three groups: the program group, the control group, and the SSP Plus group. Single parents allocated into the SSP Plus group received extensive job-search assistance in addition to the standard earnings supplement received by the members of the program group.[1]

Recruitment into the Recipient study began in November 1992 and was completed in March 1995, while the participants in the Applicant study were recruited between February 1994 and February 1995. SSP Plus took place between November 1994 and March 1995 in New Brunswick. More detailed descriptions of the design of each of the three studies are provided in Ford et al. (2003) and Michalopoulos et al. (2002).

SSP AND CHILDREN

Measured against the direct objectives, SSP was successful: it encouraged full-time employment, increased earnings, reduced dependency on welfare, and reduced poverty among single parent families (Ford et al., 2003; Michalopoulos et al., 2002). In general, improvements in economic outcomes experienced by parents are expected to produce some changes in key aspects of family life that are important to children (Ford et al., 2003; Hoffman, 1989; Michalopoulos et al., 2002; Morris & Michalopoulos, 2000; Parcel & Menaghan, 1997; Zaslow & Emig, 1997; Zaslow, Rabinovich, & Suwalsky, 1991). For instance, transition into full-time employment and increased financial resources are thought to influence the quality and reliability of child care, the physical and emotional health of the

[1]Participants from the SSP Plus study, as well as their children, were excluded for the current analysis.

parents, and the quality of interactions between parents and children. It is primarily through these intermediate outcomes that the SSP earnings supplement program was expected to produce some significant and positive changes in the lives of children whose parents were affected by the program intervention.

The results from the previous evaluations indicate that overall, SSP had less effect on children than expected (Ford et al., 2003; Michalopoulos et al., 2002; Morris & Michalopoulos, 2000). Morris and Michalopoulos's (2000) evaluation of the Recipient study indicates that three years into the program, young children of parents assigned to the program and control groups performed equally well on such outcome measures as a standardized test of vocabulary skills, levels of cognitive and academic achievement, grade repetition, social behaviour, emotional well-being, and health. Among older children, SSP was found to produce small positive effects on cognitive and school outcomes; however, there were no differences on measures of social behaviour and health. Finally, young adolescents from the program group exhibited higher levels of minor delinquency and were reported to have lower levels of academic functioning.

Michalopoulos et al. (2002) evaluated the effect of program intervention on children from the Recipient study at the end of the follow-up period (54 months after random assignment). The results of that study indicate that SSP had no significant effects on very young children — those two years of age and younger at the time of random assignment. Parents in the two research groups reported similar levels of cognitive and academic achievement, grade repetition, behavioural problems, and health. Similarly, SSP was found to have no significant effect on older children and on young adolescents who at the time of random assignment were 13 to 15 and 16 to 18 years of age, respectively. However, children from the program group who were between the ages of three and five at the time of allocation were reported to perform better in school than similar children from the control group.

Finally, the study by Ford et al. (2003) assessed the effects of the Applicant study on the families and their children at the end of the follow-up (72 months after random assignment). The results of that study suggest that SSP had no significant effects on the very young children. Parents from the program group whose children were four to nine years of age at the time of random assignment reported that their children exhibited more pro-social behaviour. However, program group children from that age cohort were also reported to be more likely to have behavioural problems.

Evaluation of Field Experiments

INTENT-TO-TREAT ANALYSIS

Randomized studies are powerful in terms of statistical analysis and inference, and in general, are relatively straightforward to evaluate. In a randomized experiment, analyzing results is simply a matter of comparing the outcomes of subjects by their random assignment into the treatment and control groups. This analysis is commonly referred to as an intent-to-treat analysis (ITT). This type of analysis represents the purist approach to establishing causal evidence of program effectiveness: the results are unequivocal and easy to explain to policy analysts, government officials, and the public. Since the Self-Sufficiency Project (SSP) employed a rigorous experimental design based on randomization, the program and the control groups were similar in all respects except for the program intervention. Hence, to measure the net impact of SSP on children, we estimated differences in outcomes between children whose parents were assigned to the program group and children whose parents were allocated into the control group.

RECEIPT OF TREATMENT

Randomized studies, particularly field experiments, are affected by a number of complications, including treatment non-compliance. Treatment non-compliance occurs when subjects in the treatment group do not comply fully with their assigned intervention; that is, when there is a variation in the level to which a treatment is taken by the study participants. In SSP, a substantial proportion of individuals in the program group did not take the earnings supplement altogether. In total, only about one third of eligible parents (33.9 per cent in the Recipient study and 27.6 per cent in the Applicant study) took the earnings supplement offer.

The primary objective of this study is to assess the effect of receipt of treatment. When the focus of the analysis is on the effect of receipt of treatment, ITT and any other estimation techniques that do not take into account compliance status (as-treated analysis or per-protocol analysis) may produce biased results (Little & Rubin, 2000). If subjects in the treatment group do not comply fully with their assigned intervention, the ITT analysis confounds two distinct phenomena: program efficacy (program impact among those who actively participate) and program take-up (compliance with program activities). Little and Yau (1998) argue that in the presence of non-compliance, a conventional ITT analysis estimates the causal effect of treatment assignment (intention to treat or encouragement) rather than the effect of the treatment actually received.[2]

Recent developments in statistical methods provide the opportunity to estimate intervention effects with non-compliance. To estimate the effect of receipt of treatment, this study employs the Complier Average Causal Effect (CACE) estimation approach (Angrist,

[2]Frangakis and Rubin (1999, 2002) refer to randomized studies with variation in levels of compliance as *encouragement studies,* since they randomize encouragement to participation, not the actual participation.

Imbens, & Rubin, 1996; Imbens & Rubin, 1997; Little & Yau, 1998; Yau & Little, 2001), which is sometimes referred to as the "Local Average Treatment Effect" (Imbens & Angrist, 1994). This refined form of the instrumental variable approach (Bloom, 1984) with clear underlying assumptions was proposed by Angrist et al. (1996). Imbens and Rubin (1997) demonstrated CACE estimation through maximum-likelihood estimation using the expectation-maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977; McLachlan & Krishnan, 1997) and a Bayesian approach using the data augmentation algorithm. Frangakis and Rubin (1999) further extended the CACE approach by providing a flexible framework for designs with both treatment non-compliance and missing outcomes. Little and Rubin (2000) argue that "the CACE is a valid causal effect because it is a summary measure of individual-level effects in a subpopulation of interest, namely compliers." To estimate the CACE effects, we employed the latent variable mixture modeling approach and Mplus software (Little & Yau, 1998; Muthén & Muthén, 2004). Since many readers may be unfamiliar with the CACE approach, we will discuss this method in a more detailed way.

Complier Average Causal Effect

NOTATION

In order to better understand the issue of treatment non-compliance and the logic of the Complier Average Causal Effect (CACE) estimation approach, we will first introduce some notation. Assume a simple experimental setting where there is only one outcome measure (Y) and where the treatment assignment (Z) for each individual i is binary: $Z_i = 1$ for subjects assigned to the treatment group and $Z_i = 0$ for those assigned to the control group. In addition, assume that an indicator of the treatment received (D) for individual i has only two levels: $D_i = 1$ for participants who took up the treatment and $D_i = 0$ for those who did not take up the treatment.

Based on treatment receipt status (D) and given treatment assignment status (Z), the potential behaviour of subjects in randomized studies can be classified into four categories: complier, never-taker, defier, and always-taker (Angrist et al., 1996). Let $D_i(1)$ denote the potential intervention receipt status for individual i who is allocated to the treatment group and $D_i(0)$ denote the potential treatment receipt status for individual i who is assigned to the control group. Then compliers are participants who do what they are assigned to do: $D_i(1) = 1$ and $D_i(0) = 0$. Never-takers are subjects who do not receive the intervention even if they are assigned to the treatment group: ($D_i(1) = 0$ and $D_i(0) = 0$). Defiers are the participants who do the opposite of what they are assigned to do: ($D_i(1) = 0$ and $D_i(0) = 1$). Finally, always-takers are the subjects who always take the treatment, no matter which group they are allocated to: ($D_i(1) = 1$ and $D_i(0) = 1$). Among the four categories of behaviour identified by Angrist et al. (1996), the CACE technique estimates causal effect of treatment for compliers.

OVERVIEW OF CACE

The CACE estimation approach was developed to evaluate the effect of treatment in randomized controlled trials that are affected by treatment non-compliance. Participants in the experimental study are assumed to belong to one of two classes: compliers or non-compliers (i.e. never-takers). In the treatment group, compliers are those individuals who took the treatment; in the control group, compliers are those subjects who would have received the treatment had they been offered it. Non-compliers in the treatment group are those individuals who failed to take the treatment when it was offered to them. In the control group, non-compliers are those participants who would have failed to take the treatment had they been offered it. The CACE approach compares the outcomes observed among the actual compliers from the treatment group with the outcomes of subjects in the control group who would have complied with the treatment, had they been assigned to the treatment condition.

The primary technical difficulty related to the CACE approach is that the compliance status is only partially observed. The compliance behaviour is known in the program group and subjects can be easily classified into either compliers or non-compliers. In the control

group, however, the compliance behaviour cannot be observed, because the intervention was not offered to members of that group. In other words, we do not know which subjects in the control group would have taken up the intervention, if they had been in the program group. As a consequence, each participant in the control group can potentially belong to either of the two classes of compliance. Compliance status in the control group is considered to be a latent variable.

ASSUMPTIONS

Since compliance information for subjects in the control group is latent (missing), the following five assumptions have to be imposed to ensure that the CACE analysis provides an unbiased treatment effect estimate for compliers (Jo, 2002b):

- *Assumption 1 (Randomization).* Treatment assignment is random. Randomization ensures that the proportion of compliers in the control group is the same as that in the treatment group (Bloom, 1994; Sommer & Zeger, 1991). As a consequence, the proportion of unobserved compliers in the control group can be estimated based on the information on the proportion of compliers observed in the treatment group.
- *Assumption 2 (Stable unit treatment value).* Potential outcomes for each subject are unrelated to the treatment status of other individuals (Rubin, 1978).
- *Assumption 3 (Exclusion restriction).* For never-takers and always-takers, the distributions of the potential outcomes are independent of the treatment allocation (Bloom, 1994; Sommer & Zeger, 1991). In other words, the offer of treatment in itself does not influence the outcomes of the never-takers and always-takers.
- *Assumption 4 (Monotonicity).* There are no defiers (Imbens & Angrist, 1994). In the Self-Sufficiency Project (SSP), parents were not allowed to receive a different intervention condition than the one that they were assigned to. Hence, neither always-taker (see *Assumption 3*) nor defier was a likely compliance behaviour.
- *Assumption 5 (Non-zero average causal effect).* The average causal effect of treatment assignment on treatment received is not equal to zero (Angrist et al., 1996).

EXCLUSION RESTRICTION

The primary challenge in the CACE estimation approach is the missing compliance information among subjects in the control group. Since the compliance status is not observed completely, the exclusion restriction assumption (*Assumption 3*) is critical to ensure identifiability in CACE models (Jo, 2002a). The exclusion restriction assumption posits that the offer of treatment in itself does not have any effect on never-takers. Under this assumption, the CACE approach disallows the effect of treatment assignment for non-compliers (Jo, 2002a, 2002b). That is, it is assumed that the (latent) mean value of any outcome variable for the non-compliers (never-takers) in the control group is, on average, equal to the (observed) mean value of that variable for the non-compliers (never-takers) in the treatment group.

Although the exclusion restriction assumption plays a critical role in resolving difficulties related to the identifiability of CACE models, it can often be unrealistic in practice (Hirano,

Imbens, Rubin, & Zhou, 2000; Jo, 2002a, 2002b). In SSP, it is possible that some subjects in the program group looked for full-time employment with the intention of receiving the earnings supplement but were not successful within the required one-year time frame. However, as their supplement-initiated job search eventually paid off in later months, the assignment to the treatment group affected these participants, even though they never received a supplement payment. It is also plausible that the treatment offer on its own had some negative psychological effect on non-compliers in the program group who could, for instance, become demoralized. Non-compliers in the control group would not be exposed to such a negative effect, because the supplement was never offered to them.

Violation of the exclusion restriction assumption may introduce substantial bias into the CACE estimates. It can either underestimate or overestimate the size of the causal effects of treatment, depending on how the assignment of treatment affects non-compliers (Jo, 2002b). The magnitude of this bias can be reduced substantially by including into the model baseline covariates that predict compliance status (Jo, 2002b). Baseline covariates contribute to the identification of CACE by providing information on the pattern of missing data for the compliance status. In theory, the exclusion restriction assumption can be completely relaxed without losing identifiability if the selected covariates are perfect predictors of compliance status (Jo, 2002a). In practice, it is unlikely to have access to such predictors; hence, the potential for bias cannot be eliminated completely.

Data and Measurement Instruments

DATA

To assess the Self-Sufficiency Project (SSP) effect on children, the current study employed the following data sources:

- *The Recipient study: The 36-month follow-up.* This data set contains information on children from the Recipient study that was collected 36 months after random allocation. At that time, supplement takers who went to work shortly after random assignment were nearing the end of their eligibility to receive supplement payments; those who found full-time employment at the end of the first year after random assignment could still receive the supplement for a full year after the survey. Parents who participated in the 36-month survey were asked to complete a questionnaire about each of their children. In addition, children aged 10 to 18 were given a self-administered survey.[3]
- *The Recipient study: The 54-month follow-up.* This data file contains information on the impacts of the program after it ended, that is four and one half years after random assignment. Unlike the child data collected at the 36-month follow-up, the 54-month follow-up data were collected only from parents. Detailed information was collected for up to two children in the age group 5.5 to 9.5.
- *The Applicant study: The 72-month follow-up.* This data file was used to examine the effects of the Applicant study on children. The information contained in this data set allows for the assessment of the impacts of the program intervention six years after random assignment. The 72-month child outcome data were collected only from parents who were asked a set of detailed questions for children aged 18 years and under.

SAMPLE

The literature indicates that children at different developmental stages may respond differently to the transitions in parental employment status and increases in family income. For instance, preschool age children might be more sensitive and vulnerable to parental engagement in full-time employment than older children, especially if they are placed in poor-quality child care (Duncan & Brooks-Gunn, 1997; Brooks-Gunn & Duncan, 1997). To examine the contextual role of the age of the child to account for the possibility that the magnitude of the program impact might differ by children's age, two age cohorts were identified:

- *Younger children* (children who were between the ages of 0 and 5 at the time of random assignment). In the Recipient study, these children were 3 to 8 years old at

[3]For a complete overview of the SSP data sets and for a more detailed description of the variables used in this study, see Ford et al. (2003), Michalopoulos et al. (2002), and Morris and Michalopoulos (2000).

time of the 36-month follow-up and 4.5 to 9.5 years old at the time of the 54-month follow-up (children under the age of 1 at baseline were not assessed in the 54-month follow-up). In the Applicant study, children who were 0 to 5 years old at the time of random assignment were between the ages of 6 and 11 when the 72-month follow-up took place.

- *Older children* (children who were between the ages of 6 and 11 at the time of random assignment). In the Recipient study, these children were 9 to 14 years old at the time of the 36-month follow-up. SSP collected no data on older children at the 54-month follow-up. In the Applicant study, children who were between the ages of 6 and 11 at the random allocation were 12 to 17 years old at the time of the 72-month follow-up (the 72-month follow-up sample also includes 18-year-old adolescents who were 12 at the time of random assignment).

OUTCOME VARIABLES

The following three child outcomes are the focus of the current study:

- *Subjective Health Status* (latent variable). Parents were asked to provide responses to the following statements about the health status of their children: (1) "His/Her health is excellent," (2) "He/She doesn't get sick often," (3) "He/She seems to be less healthy than other children you know," and (4) "When there is something going around he/she usually catches it." The items were rated on a five-point Likert scale ranging from 1 (false) to 5 (true). A confirmatory factor analysis suggested that the four indicators provide a good fit to the hypothesized factor structure. The latent variable *Subjective Health Status* is measured on a five-point scale, with higher scores indicating higher health status.
- *Behavioural Problems* (latent variable). Parents provided responses to a series of statements about the behaviour of their children. Based on the underlying theory, a latent variable *Behavioural Problems* was constructed to measure the overall level of negative behaviour. It was assessed by the following three subscales: (1) *Hyperactivity*: items such as "My child can't sit still," "My child is restless," "My child is hyperactive," or "My child is distractible;" (2) *Conduct problems*: items such as "My child gets into many fights," "My child destroys things belonging to the family or other children;" and (3) *Internalizing problems*: items such as "My child seems to be unhappy," "My child seems to be depressed," or "My child cries a lot." All items were coded on a three-point scale ranging from 1 (Never/not true) to 3 (Often/very true). The three sub-scales had acceptable internal consistency (Morris & Michalopoulos, 2000) and a confirmatory factor analysis indicated that they provide a good fit to the hypothesized factor structure. The latent variable *Behavioural Problems* is measured on a scale from one to three, with higher scores indicating more negative behaviour.
- *Academic Achievement* (latent variable). Parents of children who were attending school at the time of a given follow-up were asked to assess their children's academic functioning in three subjects: reading, writing, and math (parents of older children were asked about writing, math, and science). Answers to these questions were rated on a five-point Likert scale ranging from 1 (Not well at all) to 5 (Very well). A

confirmatory factor analysis confirmed that the three indicators provide a good fit to the hypothesized factor structure. The latent variable *Academic Achievement* is measured on a five-point scale, with higher scores indicating higher levels of academic functioning.

BASELINE COVARIATES

The current study incorporates baseline covariate values ascertained prior to randomization. Their presence in the model reduces the bias due to the potential violation of the exclusion restriction assumption (Jo, 2002b), increases the precision in the estimation of compliance status, and improves the power to detect treatment effects. In SSP, detailed information on parents was collected before they were randomly assigned to their respective groups. The *Recipient Microdata File* and the *Applicant Microdata File* contain administrative records and responses to baseline surveys for the study participants. We employed the following covariates representing parental characteristics at the time prior to the random assignment:

- *Age*: a continuous variable measuring the parent's age at baseline; this variable was centred at the population mean.
- *Male*: a dichotomous measure of the parent's gender.
- *High school*: a dichotomous measure of parental educational achievement indicating whether the parent had a high school diploma or not.
- *Marital status*: three categories of parents' marital status at baseline were identified: (1) *Married* — parents who were either married or lived in common-law unions; (2) *Separated* — parents who were separated, divorced, or widowed; (3) *Single* — parents who were never married. The last category served as a reference group.
- *Immigrant*: a dichotomous variable indicating that a respondent immigrated to Canada.
- *Physical problems*: a dichotomous measure indicating that a parent reported physical problems that limited him or her in the kind or amount of activity he or she could do at home, school, work, or leisure.
- *Illness*: a dichotomous measure indicating that a parent could not take a job within the period of four weeks prior to the baseline survey because of his or her illness or disability.
- *Work intensity*: a dichotomous variable indicating that the respondent worked within the period of four months before the baseline interview.
- *Need full-time job*: a dichotomous variable indicating that a parent's greatest need at the baseline was getting full-time employment.
- *Low self-efficacy*: a dichotomous variable indicating that a parent has a low level of self-efficacy; based on the survey question: "Sometimes I feel that I'm being pushed around in life."

SAMPLE CHARACTERISTICS

The summary statistics for the baseline covariates used in the current study are shown in Table 1. The table presents the mean (*Age*) and proportions (all other variables) for a sample of 5,677 parents from the Recipient study who took part in the 36-month follow-up survey and for a sample of 2,371 parents from the Applicant study who participated in the 72-month follow-up survey. Respondents who did not participate in the two follow-up surveys were excluded from this analysis. The respective statistics are presented by random assignment (Z) and, for subjects from the program group, by treatment received (D). For the purpose of this study, compliers in the program group were defined as parents who initiated the earnings supplement at any time within the allowed three-year time frame, regardless of how long they were receiving the supplement. In other words, compliers are those individuals who received at least one payment of the supplement. Parents who did not receive the supplement altogether were categorized as non-compliers.[4]

The analysis of the baseline data by random assignment indicates that at the start of the study, parents allocated to the program group ($Z = 1$) and the control group ($Z = 0$) were very similar to each other. Except for the proportion of single parents in the Applicant study ($p = 0.045$), there were no statistically significant differences between the two groups of parents.

The second part of Table 1 shows the values of baseline covariates for compliers in the treatment group ($D(1) = 1$) and for subjects allocated to the program group who, according to our definition of compliance status, did not take the treatment offer ($D(1) = 0$). The reported differences indicate that means and proportions for a number of baseline covariates for parents who took the intervention are significantly different from statistics for parents who did not comply.[5] In both SSP studies, parents who complied with the program intervention were, on average, younger, more likely to have a high school diploma, and less likely to have physical problems or illness and had a higher level of self-efficacy than parents who did not take the earnings supplement offer. In addition, compliers from the Recipient study (but not from the Applicant study*)* were, on average, more likely to work within the period of four months before the baseline interview and to express a greater need for full-time jobs. In both studies, compliance behaviour did not differ by gender, immigration status, and marital status.

Although we found that the baseline covariates were well balanced in the two groups as randomized, randomization does not ensure that the groups of parents defined by the compliance status were balanced. In fact, the compliers in the program group, especially those from the Recipient study, were found to have significantly different characteristics than the non-compliers. This implies that comparing subjects by treatment allocation (intent-to-treat [ITT] analysis) may produce different results than comparing subjects by treatment received (Complier Average Causal Effect [CACE] analysis). A statistically significant relationship between supplement take-up and some of the baseline covariates indicates that, if the analysis focuses on assessing of the effect of the earnings supplement receipt, the standard ITT analysis of the SSP data could provide misleading results.

[4]In the Social Research and Demonstration Corporation (SRDC) reports, compliers and non-compliers are referred to as takers and non-takers, respectively.

[5]It is also possible that parents in the program group and parents in the control group who did not complete the 36-month follow-up (Recipient study) and the 72-month follow-up (Applicant study) were systematically different from parents who remained in the analytical samples (Ford et al., 2003; Michalopoulos et al., 2002). However, as no adjustment for the response behaviour was made, any significant differences between the two groups of parents defined by the response pattern may affect the external validity of this study.

Table 1: Baseline Covariates

	Randomization			Compliance (Program Group)		
	Control Group	Program Group		Compliers	Non-Compliers	
Baseline Covariates	***Z* = 0**	***Z* = 1**	***p***	***D*(1) = 1**	***D*(1) = 0**	***p***
Recipient study						
Age	31.949	31.927	0.920	31.040	32.383	0.000*
Male	0.046	0.052	0.284	0.045	0.055	0.267
High school	0.449	0.457	0.557	0.572	0.398	0.000*
Married	0.209	0.200	0.802	0.024	0.018	0.302
Separated	0.499	0.497	0.906	0.496	0.498	0.940
Single	0.480	0.483	0.850	0.480	0.484	0.830
Immigrant	0.138	0.135	0.713	0.124	0.140	0.231
Physical problems	0.261	0.250	0.354	0.191	0.281	0.000*
Illness	0.140	0.147	0.489	0.072	0.185	0.000*
Work intensity	0.231	0.227	0.692	0.387	0.144	0.000*
Need full-time job	0.295	0.291	0.712	0.409	0.230	0.000*
Not looking for job	0.570	0.588	0.155	0.406	0.682	0.000*
Low self-efficacy	0.383	0.398	0.263	0.372	0.411	0.039*
Sample size	2,822	2,855		969	1,886	
Applicant study						
Age	32.460	32.803	0.271	31.932	33.133	0.020*
Male	0.068	0.083	0.162	0.086	0.081	0.802
High school	0.376	0.409	0.096	0.466	0.387	0.013*
Married	0.048	0.064	0.091	0.043	0.072	0.067
Separated	0.701	0.720	0.313	0.706	0.726	0.495
Single	0.251	0.216	0.045*	0.252	0.202	0.066
Immigrant	0.290	0.293	0.867	0.264	0.305	0.168
Physical problems	0.190	0.198	0.611	0.132	0.223	0.000*
Illness	0.099	0.089	0.394	0.049	0.103	0.003*
Work intensity	0.419	0.429	0.601	0.457	0.419	0.232
Need full-time job	0.305	0.296	0.644	0.298	0.295	0.941
Not looking for job	0.448	0.443	0.790	0.445	0.442	0.928
Low self-efficacy	0.354	0.365	0.559	0.304	0.388	0.007*
Sample size	1,185	1,186		326	860	

Source: The Recipient Microdata File and the Applicant Microdata File.

Note: *Estimate significant at $p = 0.05$.

Analysis

EFFECT DEFINITION

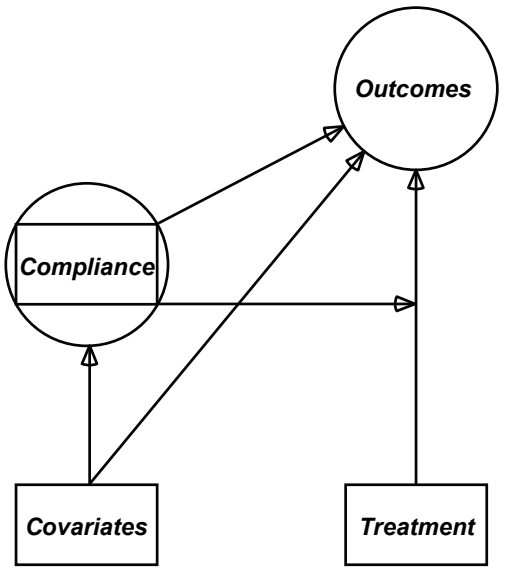
ITT. We defined the intent-to-treat (ITT) effect for a given child outcome as the difference between the average score for children in the program group and the average score for children in the control group at the time of the 36-month (Recipient study), the 54-month (Recipient study), and the 72-month (Applicant study) follow-up, regardless of their parents' compliance with the program offer. This analysis assesses the difference in outcomes between the two groups as randomized.

CACE. We defined the Complier Average Causal Effect (CACE) for a given child outcome as the difference between the average score for children whose parents were compliers in the program group and the average score for children whose parents were compliers in the control group at the time of the 36-month (Recipient study), the 54-month (Recipient study), and the 72-month (Applicant study) follow-ups, regardless of whether the compliance behaviour was actually observed or not. In other words, we compared children of parents from the program group who took up the supplement with children of parents in the control group who would be expected to take up the supplement, if it were available to them.

MODELS

Figure 1 shows a generic CACE model used for the estimation of the effect of earnings-supplement receipt on child outcomes. In this diagram, the observed binary variable *Treatment* indicates random assignment (1 = program group, 0 = control group). The partially observed variable *Compliance* represents compliance status (0 = compliance, 1 = non-compliance). The compliance status is observed among the program group subjects and it is missing (latent) in the control group. The continuous latent variable *Outcomes* denotes three child outcome latent variables: *Subjective Health Status*, *Behavioural Problems*, and *Academic Achievement*. Finally, the variable *Covariates* denotes the baseline covariates that were used to improve the prediction of compliance status (the construct path from the baseline covariates to the partially observed variable *Compliance*). Some of these covariates, namely *Age*, *Male*, *High school*, *Married*, *Separated*, *Single*, *Immigrant*, *Physical problems*, and *Illness*, were also used to improve the power to detect treatment effects, as it is indicated by the construct path from the variable *Covariates* to the variable *Outcomes*. The construct path from the variable *Treatment* to the variable *Outcomes* corresponds to the effect of the treatment receipt, the essence of this analysis. The arrow from the partially observed variable *Compliance* to this path suggests that the program effects were allowed to vary depending on the compliance status. The arrow from the partially observed variable *Compliance* to the dependent variable *Outcomes* indicates that the means of the three outcome latent variables are allowed to differ for the compliers and for the non-compliers.

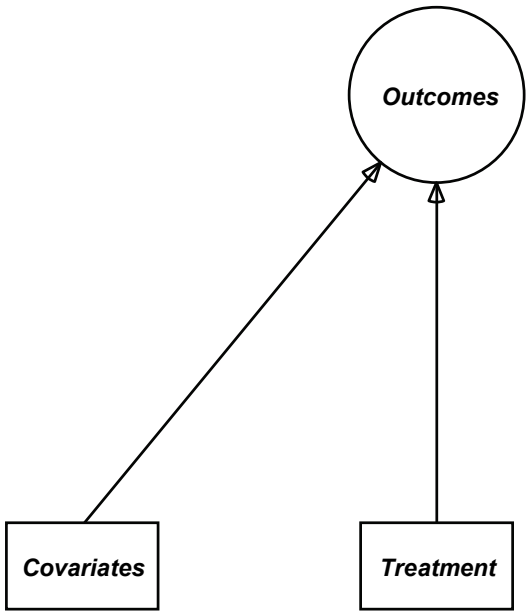
Figure 1: CACE Model



Source: Muthén and Muthén (2004).

Figure 2 shows a generic ITT model used to estimate the effect of the program intervention on child outcomes. As in the previous diagram, the observed binary variable *Treatment* indicates random assignment (1 = program group, 0 = control group). The continuous latent variable *Outcomes* denotes the three child outcome variables: *Subjective Health Status*, *Behavioural Problems*, and *Academic Achievement*. Finally, the variable *Covariates* denotes the selected baseline covariates *Age*, *Male*, *High school*, *Married*, *Separated*, *Single*, *Immigrant*, *Physical problems*, and *Illness,* which were used to improve the power to detect treatment effects (the construct path from the variable *Covariates* to the variable *Outcomes*). The construct path from the variable *Treatment* to the variable *Outcomes* corresponds to the ITT effect of the program intervention.

Figure 2: ITT Model



ESTIMATION

The ITT models and the CACE models were analyzed through the structural equation modelling (SEM) techniques (Bollen, 1989). SEM is a technique for simultaneously estimating the relationships between observed and latent variables (the measurement model) and among latent variables themselves (the construct model). SEM allows for flexible CACE modeling (Jo & Muthén, 2001) where subjects from the two compliance classes can be seen as finite mixtures (Titterington, Smith, & Makov, 1985) of subpopulations that are allowed to have separate distributions and different model parameters (Jo & Muthén, 2001). Following our operationalization of the compliance status, the likelihood of compliance with the allocated treatment is operationalized as a dichotomized variable and it is represented in the CACE model as a discrete latent variable: *Compliance* (see Figure 1). To estimate the unknown compliance status of each subject in the control condition and to estimate average treatment effects for compliers, we employed the maximum-likelihood estimation method using the expectation-maximization algorithm, a two-step iterative algorithm. Parametric standard errors are computed from the information matrix of the maximum-likelihood estimator using both the first- and the second-order derivatives, assuming conditional normality. The ITT's parameter estimates were assessed through the standard maximum-likelihood estimation method. The CACE and the ITT analyses were carried out using the Mplus program (Muthén & Muthén, 2004). The theoretical and practical aspects of this method are presented in detail by Jo and Muthén (2001) and by Muthén and Muthén (2004).

OTHER METHODOLOGICAL ISSUES

Missing data. A number of variables, including indicators of the three latent child outcome variables, have missing data points. CACE techniques can be easily applied to randomized controlled trials in which there are missing data points. CACE estimation deals with the issue of non-ignorable missing data by taking into account the missing data mechanism (Frangakis & Rubin, 1999). Specifically, to derive unbiased CACE estimates from data with non-compliance and missing data, we used an extended general location (EGL) model for CACE (Peng, Little, & Raghuanthan, 2003). The EGL model allows observations with missing data points in the outcome or the baseline covariates to be included in the analysis.

Clustering. Data in the Self-Sufficiency Project (SSP) study are multilevel because children are observed within families. Since siblings within a given family tend to be more similar to one another than to other randomly drawn children, the assumption of independent observations is violated. Applying conventional methods to multilevel data tends to underestimate the size of the standard errors. Analysis of such data requires methods that take into account the clustering effect. To adjust for the dependencies caused by the presence of multiple siblings per household, we used a sandwich estimator (Muthén & Satorra, 1995).

Results

OVERVIEW

In this section, we present the results from the intent-to-treat (ITT) and Complier Average Causal Effect (CACE) models to assess whether the Self-Sufficiency Project (SSP) extended its effects to children and whether these effects were sustained beyond the period of the intervention. The results are presented in five tables:

- *Table 2:* the results for younger children, under 5 years of age at random assignment, from the Recipient study at the 36-month follow-up; these children were 3 to 8 years of age at the time of the 36-month follow-up.
- *Table 3:* the results for older children, 6 to 11 years old at random assignment, from the Recipient study at the 36-month follow-up; these children were 9 to 14 years of age at the time of the 36-month follow-up.
- *Table 4:* the results for younger children, 1 to 5 years old at random assignment, from the Recipient study at the 54-month follow-up; these children were 5.5 to 9.5 years of age at the time of the 54-month follow-up.
- *Table 5:* the results for younger children, 5 years old and under at random assignment, from the Applicant study at the 72-month follow-up; these children were 6 to 11 years of age at the time of the 72-month follow-up.
- *Table 6:* the results for older children, 6 to 12 years old at random assignment, from the Applicant study at the 72-month follow-up; these children were 12 to 18 years of age at the time of the 72-month follow-up.

The results displayed in tables 2 to 6 follow the same format. The left side of each table presents the results from the ITT analysis. The right side of each table presents the results from the CACE analysis. As stated earlier, the ITT effect for a given child outcome variable is defined as the difference between children from the program group and children from the control group, regardless of their parents' compliance behaviour. The CACE effect for a given child outcome is defined as the difference between children of compliers from the program group and children of compliers from the control group, regardless of whether the compliance behaviour was actually observed. The CACE estimates are based on the exclusion restriction assumption. The reported estimates for the three latent variables, *Subjective Health Status* (five-point scale), *Behavioural Problems* (three-point scale), and *Academic Achievement* (five-point scale), are expressed in the units of those variables. Standard errors for the parameter estimates and the ratio of the estimates to the respective standard errors are also reported. Tables 2 to 6 also display the effects of the selected baseline covariates, *Age*, *Male*, *High school*, *Separated*, *Married*, *Immigrant*, *Physical problems*, and *Illness*, on each of the child outcome variables, both for the ITT analysis and for the CACE analysis. However, these results are not discussed in this study. Finally,

tables 2 to 6 show the estimated differences between the compliers and the non-compliers in the child outcome variables (CACE analysis only).[6,7]

Table 2: Younger Children, Aged 0–5 at Random Assignment, Recipient Study — 36-Month Follow-Up

	ITT			CACE		
Child Outcome (a)	**Estimate**	**S.E.**	**Estimate/ S.E.**	**Estimate**	**S.E.**	**Estimate/ S.E.**
Subjective Health Status	0.043	0.034	1.255	0.057	0.084	0.680
on *Age*	-0.007	0.003	-2.033*	-0.007	0.003	-1.948
on *Male*	0.101	0.130	0.779	0.097	0.130	0.750
on *High school*	0.017	0.034	0.501	0.014	0.036	0.408
on *Separated*	0.064	0.040	1.621	0.064	0.040	1.626
on *Married*	0.177	0.103	1.729	0.176	0.103	1.707
on *Immigrant*	0.084	0.061	1.388	0.085	0.061	1.406
on *Physical problems*	-0.213	0.053	-4.064*	-0.212	0.053	-3.975*
on *Illness*	-0.052	0.070	-0.736	-0.050	0.070	-0.715
Compliers vs. non-compliers	n/a			-0.004	0.100	-0.043
Behavioural Problems	0.017	0.018	0.985	0.001	0.039	0.022
on *Age*	0.002	0.002	1.267	0.002	0.002	1.232
on *Male*	0.045	0.074	0.603	0.046	0.074	0.614
on *High school*	-0.053	0.018	-2.983*	-0.052	0.018	-2.804*
on *Separated*	-0.008	0.022	-0.372	-0.008	0.021	-0.363
on *Married*	-0.121	0.041	-2.916*	-0.122	0.042	-2.913*
on *Immigrant*	-0.126	0.032	-3.886*	-0.126	0.033	-3.875*
on *Physical problems*	0.046	0.026	1.809	0.046	0.026	1.769
on *Illness*	0.010	0.037	0.279	0.010	0.037	0.280
Compliers vs. non-compliers	n/a			-0.002	0.045	-0.044
Sample size (missing)	2,356	(131)		2,356	(131)	

Source: The Recipient study: The 36-month follow-up.

Note: Data of *Academic Achievement* were not collected for younger children.

*Estimate significant at $p = 0.05$.

[6]The means of the outcome variables for the non-compliance were fixed to zero to allow for such a comparison.

[7]Following the example of other studies (Dunn et al., 2003; Jo, 2002a; Jo & Muthén, 2001; Muthén & Muthén, 2004), we assumed that the variance of each of the latent outcome variables is equal across the two compliance classes (compliers and non-compliers). Mixture modeling procedure allows subjects from the two compliance classes to have separate distributions and different model parameters. The results of the log-likelihood test indicated that some of the variances in the outcome variables were different in each class of compliance. However, when we relaxed the across-class equity constraints, some of the CACE estimates changed sign from positive to negative and some became statistically significant.

Table 3: Older Children, Aged 6–11 at Random Assignment, Recipient Study — 36-Month Follow-Up

	ITT			CACE		
Child Outcome	**Estimate**	**S.E.**	**Estimate/ S.E.**	**Estimate**	**S.E.**	**Estimate/ S.E.**
Subjective Health Status	-0.008	0.042	-0.180	-0.011	0.078	-0.147
on *Age*	-0.004	0.004	-0.989	-0.004	0.004	-0.826
on *Male*	0.179	0.111	1.609	0.153	0.114	1.342
on *High school*	0.073	0.040	1.806	0.056	0.041	1.382
on *Separated*	-0.055	0.043	-1.296	-0.061	0.042	-1.439
on *Married*	-0.180	0.136	-1.326	-0.151	0.136	-1.108
on *Immigrant*	-0.070	0.073	-0.958	-0.078	0.072	-1.085
on *Physical problems*	-0.107	0.054	-2.005*	-0.112	0.053	-2.102*
on *Illness*	-0.077	0.071	-1.083	-0.056	0.072	-0.774
Compliers vs. non-compliers	n/a			0.136	0.081	1.686
Behavioural Problems	-0.024	0.024	-1.015	-0.054	0.066	-0.818
on *Age*	-0.006	0.002	-2.475*	-0.006	0.002	-2.412*
on *Male*	-0.016	0.069	-0.227	-0.016	0.070	-0.226
on *High school*	-0.058	0.025	-2.330*	-0.059	0.026	-2.315*
on *Separated*	0.019	0.026	0.738	0.018	0.026	0.690
on *Married*	0.099	0.084	1.175	0.101	0.084	1.198
on *Immigrant*	-0.113	0.037	-3.027*	-0.113	0.037	-3.021*
on *Physical problems*	0.014	0.030	0.482	0.014	0.030	0.484
on *Illness*	0.008	0.039	0.201	0.009	0.041	0.211
Compliers vs. non-compliers	n/a			0.036	0.079	0.456
Academic Achievement	-0.017	0.039	-0.427	-0.028	0.087	-0.320
on *Age*	-0.003	0.004	-0.782	-0.002	0.004	-0.568
on *Male*	-0.019	0.095	-0.196	-0.013	0.097	-0.130
on *High school*	0.234	0.043	5.501*	0.222	0.043	5.122*
on *Separated*	-0.032	0.043	-0.747	-0.036	0.043	-0.853
on *Married*	-0.186	0.138	-1.345	-0.170	0.139	-1.223
on *Immigrant*	0.199	0.059	3.357*	0.197	0.060	3.285*
on *Physical problems*	0.012	0.048	0.246	0.009	0.048	0.189
on *Illness*	0.016	0.065	0.245	0.033	0.066	0.501
Compliers vs. non-compliers	n/a			0.105	0.096	1.091
Sample size (missing)	1,749	(126)		1,749	(126)	

Source: The Recipient study: The 36-month follow-up.

Notes: *Estimate significant at $p = 0.05$.

Table 4: Younger Children, Aged 1–5 at Random Assignment, Recipient Study — 54-Month Follow-Up

	ITT			CACE		
Child Outcome	**Estimate**	**S.E.**	**Estimate/ S.E.**	**Estimate**	**S.E.**	**Estimate/ S.E.**
Subjective Health Status	0.015	0.038	0.390	0.018	0.088	0.202
on *Age*	-0.003	0.004	-0.621	-0.002	0.004	-0.552
on *Male*	0.356	0.133	2.668*	0.349	0.135	2.587*
on *High school*	0.058	0.038	1.547	0.051	0.039	1.301
on *Separated*	0.013	0.045	0.282	0.011	0.045	0.252
on *Married*	0.020	0.143	0.141	0.016	0.143	0.110
on *Immigrant*	-0.395	0.067	-5.898*	-0.392	0.067	-5.848*
on *Physical problems*	-0.184	0.061	-3.016*	-0.179	0.062	-2.868*
on *Illness*	-0.199	0.090	-2.211*	-0.195	0.090	-2.164*
Compliers vs. non-compliers	n/a			0.044	0.099	0.449
Behavioural Problems	-0.004	0.019	-0.231	-0.019	0.054	-0.349
on *Age*	-0.002	0.002	-1.265	-0.002	0.002	-1.185
on *Male*	0.052	0.065	0.797	0.048	0.064	0.743
on *High school*	-0.017	0.019	-0.910	-0.020	0.020	-1.030
on *Separated*	0.015	0.022	0.689	0.014	0.022	0.659
on *Married*	-0.051	0.048	-1.053	-0.053	0.049	-1.084
on *Immigrant*	-0.206	0.025	-8.307*	-0.205	0.025	-8.240*
on *Physical problems*	0.054	0.027	2.006*	0.056	0.027	2.074*
on *Illness*	0.003	0.034	0.093	0.005	0.034	0.143
Compliers vs. non-compliers	n/a			0.032	0.063	0.504
Academic Achievement	0.060	0.034	1.773	0.167	0.095	1.754
on *Age*	-0.014	0.004	-3.814*	-0.014	0.004	-3.815*
on *Male*	0.125	0.133	0.940	0.130	0.134	0.974
on *High school*	0.131	0.034	3.839*	0.132	0.035	3.709*
on *Separated*	0.001	0.040	0.019	0.001	0.040	0.015
on *Married*	-0.175	0.110	-1.591	-0.170	0.111	-1.539
on *Immigrant*	0.115	0.046	2.473*	0.116	0.047	2.491*
on *Physical problems*	-0.093	0.050	-1.868	-0.095	0.051	-1.878
on *Illness*	-0.009	0.072	-0.131	-0.010	0.072	-0.144
Compliers vs. non-compliers	n/a			-0.081	0.107	-0.757
Sample size (missing)	2,291	(3)		2,291	(3)	

Source: The Recipient study: The 54-month follow-up.

Note: *Estimate significant at $p = 0.05$.

Table 5: Younger Children, Aged 0–5 at Random Assignment, Applicant Study — 72-Month Follow-Up

	ITT			CACE		
Child Outcome	**Estimate**	**S.E.**	**Estimate/ S.E.**	**Estimate**	**S.E.**	**Estimate/ S.E.**
Subjective Health Status	0.037	0.052	0.707	0.074	0.240	0.307
on *Age*	-0.001	0.005	-0.265	-0.001	0.005	-0.277
on *Male*	0.116	0.144	0.807	0.133	0.147	0.900
on *High school*	-0.023	0.050	-0.461	-0.027	0.051	-0.536
on *Separated*	0.015	0.065	0.225	0.028	0.071	0.398
on *Married*	-0.089	0.110	-0.806	-0.074	0.115	-0.643
on *Immigrant*	-0.186	0.063	-2.935*	-0.181	0.063	-2.873
on *Physical problems*	-0.112	0.087	-1.289	-0.104	0.091	-1.142
on *Illness*	-0.383	0.178	-2.154*	-0.391	0.179	-2.190*
Compliers vs. non-compliers	n/a			0.040	0.284	0.142
Behavioural Problems	0.014	0.034	0.419	0.081	0.067	1.219
on *Age*	-0.007	0.003	-2.146*	-0.007	0.003	-2.122*
on *Male*	-0.081	0.090	-0.896	-0.106	0.094	-1.132
on *High school*	0.057	0.034	1.662	0.062	0.034	1.813
on *Separated*	0.037	0.038	0.952	0.019	0.039	0.481
on *Married*	0.013	0.061	0.220	-0.006	0.064	-0.095
on *Immigrant*	0.001	0.042	0.017	-0.006	0.041	-0.143
on *Physical problems*	0.072	0.049	1.453	0.058	0.051	1.145
on *Illness*	0.095	0.080	1.197	0.110	0.086	1.274
Compliers vs. non-compliers	n/a			-0.151	0.070	-2.147*
Academic Achievement	0.040	0.071	0.568	-0.066	0.213	-0.308
on *Age*	0.003	0.007	0.488	0.003	0.007	0.466
on *Male*	0.060	0.229	0.262	0.083	0.236	0.351
on *High school*	-0.029	0.070	-0.421	-0.032	0.070	-0.453
on *Separated*	0.097	0.086	1.129	0.122	0.091	1.332
on *Married*	0.233	0.128	1.823	0.257	0.138	1.871
on *Immigrant*	0.093	0.083	1.119	0.100	0.083	1.195
on *Physical problems*	0.001	0.100	0.007	0.020	0.101	0.201
on *Illness*	-0.243	0.166	-1.467	-0.265	0.172	-1.539
Compliers vs. non-compliers	n/a			0.159	0.229	0.693
Sample size (missing)	849	(32)		849	(32)	

Source: The Applicant study: The 72-month follow-up.

Note: *Estimate significant at $p = 0.05$.

Table 6: Older Children, Aged 6–12 at Random Assignment, Applicant Study — 72-Month Follow-Up

	ITT			CACE		
Child Outcome (a)	**Estimate**	**S.E.**	**Estimate/ S.E.**	**Estimate**	**S.E.**	**Estimate/ S.E.**
Subjective Health Status	0.034	0.060	0.578	0.096	0.590	0.163
on *Age*	-0.018	0.008	-2.356*	-0.018	0.008	-2.219*
on *Male*	0.239	0.138	1.740	0.239	0.142	1.686
on *High school*	0.073	0.063	1.162	0.074	0.079	0.929
on *Separated*	-0.225	0.076	-2.960*	-0.226	0.076	-2.966*
on *Married*	-0.025	0.162	-0.157	-0.026	0.191	-0.139
on *Immigrant*	-0.056	0.070	-0.801	-0.055	0.072	-0.769
on *Physical problems*	-0.150	0.086	-1.738	-0.151	0.091	-1.670
on *Illness*	-0.111	0.129	-0.857	-0.114	0.138	-0.827
Compliers vs. non-compliers	n/a			-0.054	0.735	-0.073
Academic Achievement	-0.003	0.084	-0.038	-0.025	0.581	-0.043
on *Age*	-0.017	0.009	-1.930	-0.017	0.009	-1.793
on *Male*	-0.373	0.187	-1.991	-0.374	0.191	-1.954
on *High school*	0.044	0.086	0.517	0.045	0.099	0.456
on *Separated*	0.341	0.127	2.675	0.341	0.127	2.676*
on *Married*	0.566	0.238	2.378	0.564	0.261	2.164*
on *Immigrant*	0.085	0.100	0.851	0.085	0.103	0.827
on *Physical problems*	-0.083	0.102	-0.815	-0.084	0.105	-0.798
on *Illness*	-0.071	0.134	-0.535	-0.073	0.140	-0.519
Compliers vs. non-compliers	n/a			0.004	0.705	0.005
Sample size (missing)	658	(30)		658	(30)	

Source: The Applicant study: The 72-month follow-up.

Notes: Data of *Behavioural Problems* were not collected for younger children.

*Estimate significant at $p = 0.05$.

PARAMETER ESTIMATES

ITT Analysis: The results from the ITT models indicate that none of the parameter estimates for the ITT effects of the program intervention on the three latent variables, *Subjective Health Status*, *Behavioural Problems*, and *Academic Achievement*, was statistically significant. In other words, based on parental assessment, the health, behaviour, and academic achievement of younger and older children, across the two SSP studies and at all the follow-ups, were not affected by the program intervention.

CACE Analysis: The results presented in tables 2 to 6 show that none of the parameter estimates for the CACE effects of program intervention on the three latent variables, *Subjective Health Status*, *Behavioural Problems*, and *Academic Achievement*, was statistically significant. In other words, compared with the compliers from the control group, the compliers from the program group from the two SSP studies reported that their children were equally healthy, had the same frequency of behavioural problems, and had the same level of academic achievement.

Similarly, tables 2 to 6 (the compliers versus non-compliers part of the analysis) indicate that the compliers, regardless of their random allocation, reported that the health, behaviour,

and academic achievement of their children was the same as the health, behaviour, and academic achievement of non-compliers. The only exception to this generalization was the compliers from the Applicant study who, at the time of the 72-month follow-up, indicated that their younger children (under 5 years old at random assignment; 5 to 11 years old at the follow-up) had fewer behavioural problems than the corresponding children of non-compliers. The size of this difference was estimated at -0.151, measured on a three-point scale. Negative value of this coefficient indicates that the children of the compliers were less likely to have behavioural problems than children of the non-compliers.

Compliers vs. non-compliers. The logit coefficients for the effect of baseline covariates on the indicator of the compliance status (not shown in tables 2 to 6) show that, in general, parents from the Recipient study who were assigned to the control condition were more likely to comply with the program offer if they were younger, had high school diplomas, worked within the period of four months before the baseline interview, or indicated that their greatest need at the baseline was full-time employment. Parents from the control group would also be less likely to comply with the offer when, at baseline, they reported having some physical problems that limited the kind and amount of activity they could do. The results from the predictive model for the compliance status for parents from the Applicant study were less pronounced. It appears that only the indicator of low levels of self-efficacy was statistically significant. That is, parents from the control group who had high levels of self-efficacy at baseline would be more likely to comply with the assigned treatment, had they been assigned to the program group.

SENSITIVITY ANALYSIS

The results from the predictive model for the compliance status in the control group suggest that the baseline covariates are not perfect predictors of the compliance status, particularly in the case of the Applicant study. Thus, the identification of the CACE models is based primarily on the exclusion restriction assumption, which posits that the offer of the earnings supplement in itself did not have any effect on the non-compliers in the program group. Any significant violation of this assumption would introduce the possibility that the results presented in this study's CACE estimates are biased.

The restriction exclusion assumption might be questionable in the SSP study. As indicated before, the never-takers in the program group (non-compliers) could be demoralized by failing to take the earnings supplement offer. It is also possible that their supplement-initiated job search paid off after the allocated one-year time frame but before the follow-up survey. In this study we explore the sensitivity of the CACE estimates to the violation of the exclusion restriction assumption, employing the methodology proposed by Jo (2002a; Jo & Muthén, 2001). To demonstrate how the violation of the exclusion restriction assumption is affecting the estimated program effects, we relaxed this assumption by specifying different values for the effect of offering the earnings supplement to the non-compliers in the program group. In other words, we examined the validity of the exclusion restriction assumption by proposing a series of alternative assumptions on the effect of treatment allocation for the non-compliers. Following Dunn et al. (2003), we set up the maximum absolute values of this effect at approximately the same magnitude as the highest observed size of the CACE effect, that is, at 0.15 for the effects on all three latent variables: *Subjective Health Status*, *Behavioural Problems*, and *Academic Achievement.*

Table 7 displays the results of this assessment. These results indicate that the modified CACE estimates move gradually between the two extreme values that we set up in the expected directions. For instance, under the exclusion restriction assumption, the effect of the program receipt on the latent variable *Subjective Health Status* for the cohort of the younger children from the Recipient study at the 36-month follow-up was estimated at 0.057. We expected that setting the effect of program allocation for the non-compliers to a negative value would increase the CACE estimate. As expected, when we fixed the program effect for the non-compliers to the value of -0.15, the modified CACE estimate increased to the value of 0.252. On the other hand, when we set up the program effect for the non-compliers to the maximum positive value of 0.15, the modified CACE estimate decreased, as expected, to the value of -0.028. Some of the modified CACE estimates are statistically significant at $p = 0.05$. For instance, if the true effect of the program offer on the health (*Subjective Health Status*) and behaviour (*Behavioural Problems*) of children of the non-compliers from the Recipient study was -0.15, then the adjusted CACE effects for the compliers would be 0.252 and 0.093, respectively. Hence, we concluded that although most of the CACE estimates of the earnings supplement are not substantially affected by the violation to the exclusion restriction assumption, some of these effects might be significant if the exclusion assumption is violated.

Table 7: Sensitivity Analysis — Exclusion Restriction Assumption

		Assumed Program Effect for Non-compliers						
		-0.15	-0.10	-0.05	0.00	0.05	0.10	0.15
Age Cohort — SSP Study — Follow-Up		**CACE Estimates**						
Younger children (aged 0–5 at random assignment) from the Recipient study at 36-month follow-up								
Subjective Health Status	Line A	**0.252***	0.208	0.131	0.057	0.011	-0.014	-0.028
Behavioural Problems	Line B	**0.093***	**0.070***	0.040	0.001	-0.045	-0.091	-0.130
Older children (aged 6–11 at random assignment) from the Recipient study at 36-month follow-up								
Subjective Health Status	Line C	0.229	0.165	0.064	-0.011	-0.047	-0.063	-0.073
Behavioural Problems	Line D	0.079	0.057	0.011	-0.054	-0.117	**-0.156***	**-0.171***
Academic Achievement	Line E	0.203	0.147	0.050	-0.028	-0.074	-0.098	-0.114
Younger children (aged 1–5 at random assignment) from the Recipient study at 54-month follow-up								
Subjective Health Status	Line F	0.131	0.094	0.048	0.018	0.011	0.010	0.004
Behavioural Problems	Line G	**0.080***	0.063	0.033	-0.019	-0.090	**-0.142***	**-0.170***
Academic Achievement	Line H	**0.385***	0.332	0.252	0.167	0.095	0.042	0.001
Younger children (aged 0–5 at random assignment) from the Applicant study at 72-month follow-up								
Subjective Health Status	Line I	0.378	0.355	0.270	0.074	-0.028	-0.075	-0.094
Behavioural Problems	Line J	**	**	**	0.081	0.032	-0.041	-0.076
Academic Achievement	Line K	0.196	0.150	0.049	-0.066	-0.132	-0.149	-0.165
Older children (aged 6–12 at random assignment) from the Applicant study at 72-month follow-up								
Subjective Health Status	Line L	**	**	**	0.096	-0.019	-0.075	-0.114
Academic Achievement	Line M	**	**	**	-0.025	-0.083	-0.169	-0.246

Notes: *Estimate significant at $p = 0.05$.
**The estimates were not reliable.

Conclusion

The objective of this study was to assess whether the Self-Sufficiency Project (SSP) affected the health, behaviour, and academic achievement of children and whether these effects were sustained beyond the period of the intervention. The impact of earnings supplement programs such as SSP is a function of two processes: actual program effects and program coverage or take-up. Since only the minority of the eligible parents in SSP took advantage of the program offer, the observed program effects could be disproportionately confined to those children whose parents complied with the program intervention. Because of the uncertainty about supplement use in real world circumstances, we examined both the net program impacts on all children whose parents were randomly assigned to the program group (the intent-to-treat [ITT] analysis) and the program effects on those children whose parents took advantage of the earnings supplement (Complier Average Causal Effect [CACE] analysis).

The results from both analyses suggest that children were not harmed, nor did they benefit from the program intervention. In particular, children's health, behaviour, and academic achievement were not affected by the program intervention, whether their parents took advantage of the earnings supplement offer or not. Taking into account how young some of these children were at the time of random assignment, it is reassuring that the increases in full-time employment did not result in negative effects for these children. From the policy point of view, the results of this evaluation suggest that when (and if) a welfare-to-work program similar to SSP is introduced to the general population, it is unlikely to affect children's health, behaviour, and academic achievement, regardless of the level of program take-up.

The results of other studies using ITT analysis indicate that SSP had a small positive effect on children between the ages of two and five (Michalopoulos et al., 2002; Morris & Michalopoulos, 2000) and a small negative effect on adolescents (Morris & Michalopoulos, 2000). The discrepancy in the ITT estimates between the results from the previous studies and those presented in this report might be due to differences in methodological approach. In particular, (1) the estimates of program effects were adjusted for the values of baseline covariates, (2) all the outcome variables were measured as latent variables, and (3) some measures were taken to address the problems of missing data and the presence of multiple siblings per household.

There is some empirical evidence, including results from SSP, to suggest that adult-oriented welfare-to-work programs have the potential to alter various aspects of family life (Ford et al., 2003; Hoffman, 1989; Morris & Michalopoulos, 2000; Parcel & Menaghan, 1997; Zaslow & Emig, 1997; Zaslow et al., 1991). These effects might be either favourable or unfavourable for children. The results from Atlanta's labour force attachment program, for instance, indicate that the transition into full-time employment has the potential to increase parents' sense of time stress. At the same time, increased financial resources can improve positive parenting skills (Zaslow, McGroder, & Moore, 2000). Counterbalancing positive and negative effects of this kind might help to explain the absence of significant effects reported in the current SSP study.

At this time, we did not separate positive and negative effects. As a result, it remains unknown whether the negative effects of the transition into full-time employment were counterbalanced by the positive effects of increased financial resources or whether neither of these two effects had any influence on the children. We would like to argue that from the policy point of view, it is important to understand the nature and direction of multiple program influences on family life. It is plausible that the positive effects of the SSP-like welfare-to-work programs on children can be significantly increased by strengthening the pathways that produce positive effects and weakening the pathways that produce negative effects for children. Further research into the causal mechanism (pathways) of the SSP program intervention will be critical to provide a definite answer to this question.

References

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Society, 91*, 444–455.

Bloom, H. S. (1994). Accounting for non-compliers in experimental evaluation designs. *Evaluation Review*, *8*, 225–246.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: John Wiley and Sons.

Brooks-Gunn, J., & Duncan, G. (1997). The effects of poverty on children. *The Future of Children, 7*(2)*,* 55–71.

Canadian Institute of Child Health. (2000). *The health of Canada's children*. Ottawa: Author.

Dempster, A., Laird, N., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, *Series B*, *39*, 1–38.

Duncan, G., & Brooks-Gunn, J. (Eds.). (1997). *Consequences of growing up poor*. New York: Russell Sage Foundation.

Duncan, G., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development, 65*, 296–318.

Dunn, G., Maracy, M., Dowrick, C., Ayuso-Mateos, J. L., Dalgard, O. S., Page, H., et al. (2003). Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. *British Journal of Psychiatry*, *183*, 323–331.

Ford, R., Gyarmati, D., Foley, K., & Tattrie, D. (with Jimenez, L.). (2003). *Can work incentives pay for themselves? Final report on the Self-Sufficiency Project for welfare applicants*. Ottawa: Social Research and Demonstration Corporation.

Frangakis, C. E., & Rubin, D. B. (1999). Addressing complications of intent-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*, *86*: 366–379.

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, *58*, 20–29.

Heckman, J., Smith, J., & Taber, C. (1998). Accounting for dropouts in evaluations of social programs. *Review of Economics and Statistics*, *80*, 1–14.

Hill, M. S., & Sandfort, J. R. (1995). Effects of childhood poverty on productivity later in life: Implications for public policy. *Children and Youth Services Review*, *17*, 91–126.

Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. (2000). Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics*, *1*, 69–88.

Hoffman, L. W. (1989). Effects of maternal employment in the two-parent family. *American Psychologist*, *44*, 283–292.

Imbens, G. W., & Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*, 467–476.

Imbens, G. W., & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics, 25*, 305–327.

Jo, B. (2002a). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioural Statistics*, *27*, 385–409.

Jo, B. (2002b). Model misspecification sensitivity analysis in estimation causal effects of interventions with non-compliance. *Statistics in Medicine*, *21*, 3161–3181.

Jo, B., & Muthén, B. O. (2001). Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 57–87). Mahwah, NJ: Lawrence Erlbaum Associates.

Korbin, J. E. (1992). Introduction: Child poverty in America. *Behavioral Scientist*, *35*, 213–219.

Lipman, E. L., Boyle, M. H., Dooley, M. D., & Offord, D. R. (2002). Child well-being in single-mother families. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 75–82.

Lipman E. L., Secord, M., & Boyle, M. H. (2001). Moving from the clinic to the community: The Alone Mothers Together Program [Letter to the editor]. *Canadian Journal of Psychiatry*, *46*, 657.

Little, R. J., & Rubin, D. B. (2000). Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. *Annual Review of Public Health*, *21*, 121–145.

Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods*, *3*, 147–159.

McLachlan, G. J., & Krishnan, T. (1997). *The EM algorithm and extensions*. New York: John Wiley and Sons.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist*, *53*, 185–204.

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

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

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. V. Marsden (Ed.), *Sociological Methodology* (pp. 267–316). Washington: The American Sociological Association.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus User's Guide*. Los Angeles: Muthén and Muthén.

Parcel, T. L., & Menaghan, E. G. (1997). Effects of low-wage employment on family well-being. *The Future of Children*, *7*(1), 116–121.

Peng, Y., Little, R., & Raghuanthan, T. E. (2003). *An extended general location model for causal inference from data subject to noncompliance and missing values* (Working Paper 7). Ann Arbor, Michigan: University of Michigan School of Public Health, Department of Biostatistics.

Rubin, D. B. (1978). Bayesian inference for causal effects. *The Annals of Statistics*, *6*, 34–58.

Sommer, A., & Zeger, S. (1991). On estimating efficacy from clinical trials. *Statistics in Medicine*, *10*, 45–52.

Statistics Canada. (1996). *Census families by presence of children, 1996 Census*. Ottawa: Author.

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*. New York: John Wiley and Sons.

Yau, L. H. Y., & Little, R. J. A. (2001). Inference for complier-causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for unemployed. *Journal of the American Statistical Association*, *96*, 1232–1244.

Zaslow, M. J., & Emig, C. A. (1997). When low-income mothers go to work: Implications for children. *The Future of Children*, *7*(1), 110–115.

Zaslow, M. J., McGroder, S. M., & Moore, K. A. (2000). *Summary report on the two-year follow-up of the Child Outcomes Study: National Evaluation of Welfare-to-Work Strategies.* Washington, DC: U.S. Department of Health and Human Services and U.S. Department of Education.

Zaslow, M. J., Rabinovich, B. A., & Suwalsky, J. T. (1991). From maternal employment to child outcomes: Preexisting group differences and moderating variables. In J. V. Lerner & N. L. Galambos (Eds.), *Employed mothers and their children* (pp. 237–282). New York: Garland.