Estimating family labour supply:
A review of prior NCB work and charting a future course

Greg Mason
PRA Inc. and the University of Manitoba
June 12, 13, 2006

1.0 Introduction

This paper sketches issues in the following four themes:

- key issues in family labour supply including what models must try to capture
- revisiting the data landscape in Canada to support econometric estimates of labour supply by National Child Benefit (NCB) recipients
- surveying the prior NCB estimates of labour supply using matching techniques and the Survey of Labour and Income Dynamics (SLID)
- new directions for estimating the labour supply of NCB recipients using discrete choice and simulation as alternatives to econometrics.

Two basic points are made in this paper. First, quasi-experimental estimation of labour supply, using statistical matching on the SLID data, represents the state of the art within the neo-classical framework. This approach requires the acceptance of a major assumption, which many cannot support, about the creation of a program and comparison group for a universal program. Second, micro-simulation using discrete choice methods presents an alternative method to model family labour supply for NCB clients. This approach steps outside the somewhat rigid traditional paradigm typically used in econometric evaluations. It, too, has important assumptions that can qualify results.

2.0 Key issues in estimating labour supply

Neo-classical labour supply theory typically presents a constrained maximization problem where the actor attempts to maximize utility (U), subject to a time constraint (16 waking hours). Individuals are assumed to find an equilibrium between the wage rate and work-leisure. The typical set-up appears as:

maximize U = f(Income, Leisure)
s.t.
labour hours + leisure hours $\leq$ 16 waking hours

or in terms of labour hours (L) as follows:

maximize U = f(w*L, 16-L),
where w = wage rate.

This leads to the standard conclusion that workers will equate the marginal rate of substitution between income and leisure to the wage that represents the *opportunity cost* of leisure time in terms of foregone income.

As with all theoretical structures, application to actual decision making requires many adjustments. Typically, these are reflected in the model constraints, and developments in this area try to account for institutional barriers, such as the fixed workday, and the fact that most people do not divide their days into periods of work and non-work. More fundamental adjustments require the reformulation of the basic assumptions of the utility function. For example, equating non-work to leisure is a strange concept for most people, especially low-income parents with young children.

Specific issues that must be built into a family labour supply model relevant for NCB households include the following:

- Time allocation decisions are joint decisions among household members, especially those who contribute to the operation of the household.

- A contribution to the household may be cash or in kind. This is not simply a matter of pricing housework at its market equivalent. Rather, the work decision-making problem within the household may need to include factors reflected in the utility function and/or the constraints. For example, the presence of a child with a disability may be modelled as a constraint on the labour force participation of one of the adults or as a preference by the household to have one adult care for the child as opposed to increasing income and paying for care.

- Other institutional factors include the rules of social assistance. Much has been written about the welfare wall and how the financial and non-financial benefits of being on welfare provide powerful incentives not to participate in paid work.

- As elucidated by recent qualitative research, the concept of the welfare wall setting a relatively high reservation wage is quite general. Employment insurance also encourages EI clients to have high reservation wages. The decision to work or accept training is the outcome of a calculation about the probability of finding work at a specific target wage, perceived costs and benefits of moving to higher wage areas (e.g., Alberta), and the lifestyle possible with the social safety net.

Models of family labour supply therefore need to incorporate these factors. Much of the development based on the neo-classical framework has introduced increasing complexity into the model constraints, but there is an understandable reluctance (and conceptual impediment) to creating more complex utility functions.

One of the significant impediments to estimating family labour supply is the availability of suitable data. The next section updates the data assessment conducted as part of the evaluability assessment completed for the first round of the NCB. It examines the databases that could conceivably support the neo-classical perspective on estimation of labour supply.

3.0 Data sources for estimating family labour supply of NCB participants

The NCB evaluability assessment in 2000 reviewed all existing databases that could potentially support the analysis of family labour supply. Appendix A presents this analysis, updated to the present.

The main conclusion from this analysis is that family labour supply can rest on the following databases:

- *Longitudinal Administrative Database (LAD)* can create family tax files, which will allow the extraction of those receiving the NCB. Within the limits of the tax forms, this data will support analysis of labour supply, as inferred by earnings, earnings to income, and earnings to social assistance ratios. The income of other earners is included and number/age of children can be determined, as well as disabilities (to the extent that costs associated with disabilities are claimed). A serious limitation is that certain important variables, such as education, do not appear.

- The *Survey of Labour and Income Dynamics (SLID),* identified in the NCB evaluability assessment, remains the most useful non-experimental database available for the NCB. It provides longitudinal links over time, family and individual measures, tax-based variables, and a range of family and socio-economic variables. The NCB eligible sub-sample is about 5,000, and so provides a useful basis for looking at various subsets such as those receiving social assistance. The SLID cannot support analysis on specific important subsets such as immigrants, Aboriginal populations, or low-income parents of children.

- *Provincial reinvestments.* Downplayed in 2000 because many of these programs had just started, data from some provincial reinvestment programs could potentially support labour supply modelling. The Saskatchewan Earnings Supplement is a case in point. This requires separate data development with the relevant provinces, and the resulting files may need to be supplemented with other information from training programs and client surveys.

- *Social assistance files.* Most provinces maintain social assistance caseload data, part of which they share with the federal government. It is feasible to develop a standardized extract of these data to create a file to support various forms of analysis. For Phase 3 of the NCB evaluation, the social assistance data from several provinces formed the basis for a survival analysis to determine the impact of the NCB on participation in welfare. These data do not include the "working poor" who never take social assistance, but they do include many who participate partially and who have earnings from work, so this information has potential.

The obvious requirement to complete a net impact assessment is the creation of a valid counterfactual. Since the NCB is a universal program, open to those who submit a tax return, no database will support the creation of a valid program and comparison group. While the SLID offers the most complete range of variables, and longitudinal linkages, the trick is creating this counterfactual.

4.0 The NCB evaluation – using statistical matching methods and the SLID

The NCB evaluation used standard statistical matching procedures to create a program and comparison group. Since all eligible NCB recipients have children, the comparison group must consist of households for which all variables are the same as program recipients, save that they will not have children. The matching equation essentially becomes a fertility model. This follows the work by Eissa and Hoynes (1999) on the Earned Income Tax Credit.

Many disengage from the analysis at this point, arguing that the labour market decision making for those with children differs markedly from the decision making for those without. Even with statistical matching, unmeasured differences between the matched program and comparison groups, which may account for deciding to have a child, may also affect labour market participation.

If the matching model produces a close match between program and comparison groups, based on both general statistical measures of quality, plus balancing tests, it is probably valid to express less concern about the approach. However, the reality is that most practical applications of statistical matching produce mediocre matches at best, largely because of omitted variable bias. Accordingly, a fertile ground exists in which to breed criticism of the method.

That being said, within the standard econometric paradigm, statistical matching remains the state of the art. Table 1 presents some key results from the NCB evaluation completed in 2004.

Extracting from that report, main findings are as follows:

Labour force attachment. *The effect of the program to increase family weeks worked is concentrated in the group without new children, as would be expected. In families without new children, the program effect is to increase work by 7.5 weeks, whereas the program effect on weeks worked for families with new children is insignificant.*

Income. *The estimated effects of the program on total family earnings and total family income are similar for the two groups. The program has a negative impact on family earnings of $6,007 for families with new children and $6,550 for families without new children. Total family income is $7,881 lower in the program group with new children and $6,524 lower in the program group without new children. Families with new children saw no difference in SA income relative to the comparison group, whereas families without new children saw a decline of $182.*

Source: NCB Evaluation Net impact assessment, Technical Document 2, 2004

Table 1: Difference-in-differences matching estimates of NCB program impact for the full sample (with 95% confidence interval estimates)

SLID respondent indicators:	New child – 1996 (n=561)		No new child – 1996 (n=6,309)	
	Difference-in-differences (change relative to comparison group)	95% confidence interval	Difference-in-differences (change relative to comparison group)	95% confidence interval
Weeks worked	-1.77	[-4.27, 0.73]	-0.18	[-0.92, 0.57]
Annual paid hours	-50.9	[-169.8, 68.1]	-34.4	[-72.5, 3.7]
Annual earnings	-$1,665	[-3410, 81]	-$4,326*	[-4924, -3728]
SA	-$187	[-407, 33]	-$262*	[-335, -188]
Total income	-$2,405*	[-4164, -647]	-$4,124*	[-4716, -3532]
Family indicators:				
Weeks worked	-3.54	[-8.47, 1.40]	7.51*	[5.80, 9.22]
Earnings	-$6,007*	[-9125, -2889]	-$6,550*	[-7651, -5448]
SA	-$69	[-428, 289]	-$182*	[-280, -84]
Total income	-$7,881*	[-11038, -4725]	-$6,524*	[-7631, -5416]
% Below LICO (pre tax)	-2.67	[-6.97, 1.63]	-0.74	[-2.09, 0.60]
% Below LIM (pre tax)	-1.07	[-5.69, 3.55]	1.2	[-0.16, 2.57]
% Below LICO (post tax)	-3.03	[-7.21, 1.15]	-0.6	[-1.85, 0.65]
% Below LIM (post tax)	-0.89	[-4.87, 3.09]	0.97	[-0.32, 2.26]
Depth of poverty (LICO pre tax)	-$204	[-688, 280]	$92	[-51, 235]
Depth of poverty (LIM pre tax)	$155	[-309, 619]	$301*	[162, 441]
Depth of poverty (LICO post tax)	-$82	[-437, 272]	$100	[-6,206]
Depth of poverty (LIM post tax)	$45	[-313, 403]	$202*	[93, 312]

Notes: (1) Based on 0.5% caliper. (2) Sample size is reduced for individual weeks worked and paid hours and family weeks worked because of missing values (responded "don't know" or refused to answer; not applicable coded to zero where appropriate). For the full sample, sample size is 5,326 for individual weeks worked, 4,593 for individual paid hours, and 5,023 for family weeks worked. (3) Approximate statistical significance based on the z-statistic for the difference of means; * denotes 5% level of significance.

One of the results that caused some concern is that the NCB appears to have had "perverse" results. For individuals, *there is no program impact on weeks worked. The program impact on SLID respondents is to decrease hours worked in the under $32,000 group (-94.6) and the over $32,000 group (-46.0).*

Again, families under $32,000 show the "best" program impact. While the program is estimated to reduce total family earnings (-$3,083) and total family income (-$1,359), these reductions are less severe for families under $32,000. Families with incomes above $32,000 react to the program in such a way as to have the greatest impact on total family earnings (-$10,501) and total family income (-$10,889).

These large reductions, especially in earnings, imply that some families are participating in the labour force less than the comparison group, despite the fact that as a group, the program families are working more weeks on average. However, we can conclude that the relative earnings and income declines for NCB recipients observed for the full sample in Table 6 are concentrated in families with incomes over $32,000. These results, as one might expect, suggest that NCB recipients with higher incomes are adjusting their work arrangements to take care of children and that this has some consequences for earnings growth relative to families without children. These results also imply that the NCBS[1] and its SA[2] replacement, which is directed to those under $32,000, improve work incentives.

Source: Source: NCB Evaluation Net impact assessment, Technical Document 2, 2004

[1] National Child Benefit Supplement
[2] Social Assistance

The analysis of the SLID data could not confirm that higher income parents used the income security of the NCB to "buy" time to parent. That insight came from the focus groups, especially from parents with disabled children.[3] This illustrates the limitation of evaluations that attempt to use statistical methodology to "parse" the causal relationships. Any experimental or quasi-experimental method must create a strict regime to isolate the causal factors that account for the observed outcomes. The standard experimental design that operates as the putative gold standard in research renders all changes in outcomes to a single factor. Statistical matching mimics this process, but admits to allowing variations in the treatment to have a range of outcomes (changes in income, incidence of poverty, earnings and hours worked in the case of the NCB).

It is not possible to identify the potential causes for higher income parents withdrawing from the labour market within any given matched sample. One needs to develop a hypothesis about increased preference for parenting from qualitative analysis such as focus groups, dyads and case studies. Confirming "hunches" discovered then requires matching on sub-samples of parents with and without disabled children and different levels of income. Each hypothesis needs a new matched sample. New software makes this process easy; the problem is the lack of sample. In general, matching requires a sample of at least 1,000 participants and non-participants. The SLID cannot support such specific analysis.

In summary, statistical matching using difference-in-differences remains the state of the art in terms of econometric modelling based on the neo-classical framework and represents a valuable line of evidence. It should be replicated in any future NCB evaluation. As well, any future evaluation should expand the qualitative component early in the analysis to generate the important hunches that would be confirmed quantitatively. It is unfortunate that qualitative analyses are included in outcome evaluations, where their insights can not be integrated into quantitative analysis. This represents a defect in the current evaluation paradigm in Canada, which tries to measure and confirm outcomes without connecting results to any deeper understanding of how the program actually works.

If statistical matching, albeit a powerful method, has defects, what alternatives exist? The next section sketches recent developments in micro-behaviour modelling using discrete choice theory. Closely related to experimental economics, these methods represent promising approaches to evaluation and policy design.

[3] The focus groups were conducted well before the finalization of the SLID analysis for the NCB.

5.0 New directions for estimating labour supply

5.1 Micro simulation and discrete choice

Experimental economics is slowly emerging as a policy tool. For most of the two decades of its development, this branch of economics has mined the depths of rational choice theory, auction markets and micro-decision-making under uncertainty.[4] Recently, it has started to emerge as a policy tool.[5]

An important aspect of experimental economics is the use of stated choice (discrete choice) methods that use randomized subsets of a sample to try policy options. Well developed as a pricing methodology in market research, this method is very simple in principle. A sample is randomly divided into a number of sub-sets (three or more) and presented with different options such as a price for a product. The pattern of "yes/no" responses to a willingness to buy question presents an empirical measure of demand. This model has seen wide application in environmental economics to price intangible assets such as the preservation of wilderness areas.[6] Figure 1 presents a synopsis of how experimental economics aligns with observational studies.

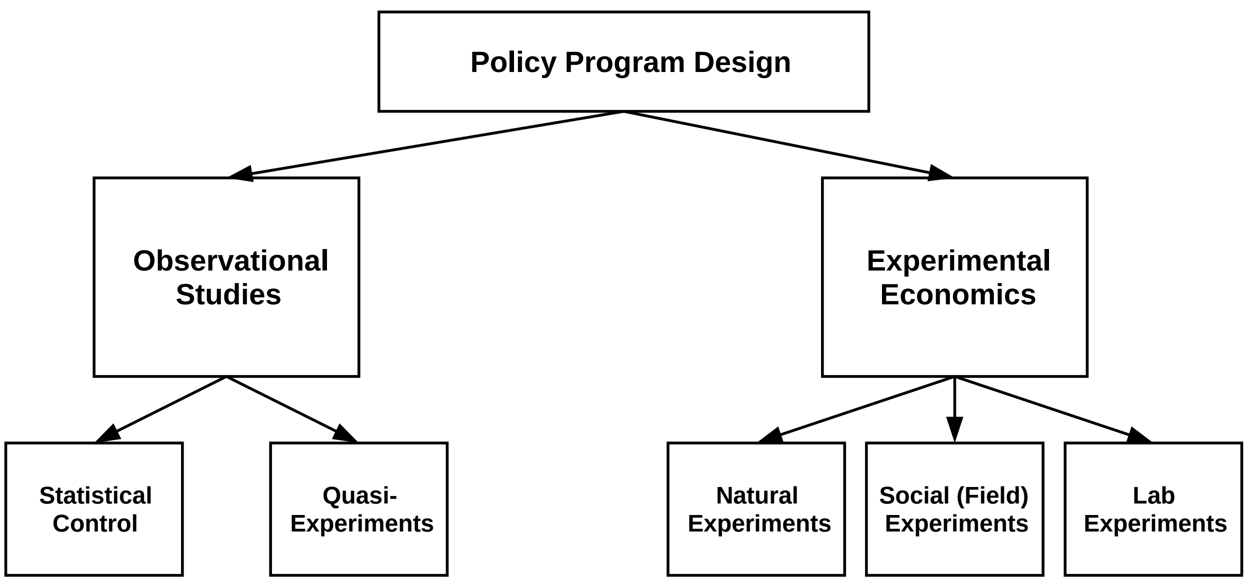


Figure 1: Policy Program Design

Blundell et al. (2000) apply discrete choice models to estimating the labour supply of families in receipt of an earnings tax credit. Bourguignon and Spadaro (2006) include simulation on discrete choice results to model a range of redistribution policies. A preliminary effort to apply discrete choice in the context of the NCB was attempted for the first evaluation. The following sections outline this preliminary attempt.

4 Guala (2005) provides an excellent account of this literature.
5 See Cumming et al. (2004).
6 The standard reference is in Louviere, Hensher and Swait (2000).

5.2 The welfare wall as a measure of labour supply[7]

The second goal of the NCB is to encourage labour force participation by reducing disincentives to work. Many NCBS clients currently receive or have received SA in the past,[8] and as we discovered in the focus group study in the NCB evaluation, most are keenly aware of the barriers to work – or the elements of the "welfare wall." The concepts of the welfare wall and labour supply are clearly connected on a fundamental level.

The welfare wall can be expressed in many ways. Conceptually, it represents the barrier to accepting work because the social assistance recipient cannot command a sufficient wage to replace the financial and non-financial benefits of welfare. Further, since SA allows recipients to have all their time as "leisure" (i.e., non-work in the sense of the model described in section 2 above), there is an extra incentive not to work.

Much of the discussion of the welfare wall is cast in the form of the marginal tax rate (MTR). The traditional social assistance model taxed earnings at 100%; every extra dollar earned resulted in a dollar reduction in benefits. Welfare reform has created earnings set asides and other inducements to work, but most analysts still believe that the effective marginal tax rate (EMTR) is higher than the highest rate on income. Critics of the MTR as an operational concept argue that it is less important than *the choice between full-time and part-time work — (which) depends more on the difference in average tax rates between low and middle income levels than on the MTRs levied at specific points in the distribution. A policy that clawed back more benefits from low and middle-income families might have undesirable effects on participation, particularly of secondary earners"* Bird and Smart (2001).[9]

Several commentators cast doubt on the usefulness of empirical estimates of the EMTR as a measure of the welfare wall and the assumption that SA recipients fit the "rational decision maker model." The critique is not directed at the rationality of SA clients, but rather at the narrow precepts of economic analysis and its reliance on the price/tax system as a predictor of actual behaviour.

Battle and Torjman (2000) state:

> *Labour market behavior is affected by a variety of factors, amongst which marginal tax rates are but one consideration. For example, the availability of affordable child care and the fierce desire of many Canadians – especially the thousands of working poor and modest income employed parents – to work for pay and not rely on welfare are two such factors. Some people work because they like the social aspects of the workplace, or the challenge of work, or for any number of non-monetary reasons. It seems implausible to us that higher marginal tax rates on working poor families resulting from the Canada Child Tax Benefit – even if such families understood the implications, which is doubtful – would result in parents deciding to quit their job or to move from full-time to part-time work.*

---

7 This material is extracted from "An empirical test of the welfare wall: a discrete choice approach to the reservation wage." Unpublished NCB Technical Study.

8 Using provincially weighted NCB client survey data, about 6 in 10 clients have received SA in the past, and about 1 in 4 has received SA in the past year.

9 http://www.csls.ca/events/slt01/bird.pdf

Another perspective is offered by Caniglia (1996), who writes:

> *Conventional analysts seem to view all individuals as fundamentally the same and as relatively strict adherents to the maximizing framework…. Not all individuals, of course, are rational economic agents in an even general sense. Some are "idealists" who are self-sacrificing in their desire to do the "right thing" to such a degree that they betray a fundamental lack of consistency with the rational behavior assumption. This could be motivated by religion, morality, or social ideals.*[10]

The consensus is that the welfare wall, while a potentially useful policy tool, if it includes the long-term costs and benefits of improved parenting, is neither empirically tractable nor universally applicable. An alternative approach is to measure the willingness to work of NCBS clients using a reservation wage approach. This captures the personal and cultural as well as the financial factors in the decision to work.

5.3 Measuring the welfare wall through the reservation wage

The reservation wage of an individual is a function of the wages he/she is offered, the number of jobs offered to him/her, and the costs of searching for work. Search costs are a function of personal circumstances and the technology that firms use to publicize the availability of work. The reservation wage represents the wage a worker would require to give up an additional hour of leisure to work that hour (Hammermesh and Rees, 1984). It is the lowest wage at which a person would accept a formal paid job. Although the work-leisure trade-off is common to labour supply research, for many NCBS clients, the alternative to formal, paid work is increased time parenting, which few regard as "leisure."

The reservation wage is a subjective value influenced by many factors such as:

- the nature of the work sought (with dangerous employment and "off-hour" work requiring a higher wage)
- distance to the job
- other income earned by the household
- the hours presently spent in formal paid work
- the duration of present unemployment
- the flow of job offers and perceived availability of employment
- the length of prior employment, the wage earned, and the level of savings
- other family responsibilities, etc.

---

[10] Caniglia, Alan S. (1996). "How large are welfare's work disincentives?" *International Journal of Social Economics*, Vol. 23 No. 9, pp. 61-68.

The survey-based studies on the reservation wage all use a direct question such as, *"How much would the wage need to be for you to accept a position now?"* A direct parallel exists with market research analyses, where a consumer might be asked, *"How much are you willing to pay for product 'X'?"* Market researchers have recognized that directly posing a pricing question risks strategic response bias, where the respondent purposely under or overstates his/her willingness to pay in an attempt to sway results of the research.

A discrete choice questionnaire structure counters this bias. Respondents simply agree or disagree to a specific question – "would you accept a job at $Y per hour?" Those who respond "YES" can be asked whether they would accept a job at a test wage of $Y-$1, and those who reply "NO" may be tested with $Y+$1. The percentage of respondents who answer YES (i.e., they would accept a job that provides a "YES") may be plotted against the three test wages (Y, Y-$1, Y+$1). A researcher can widen the scope of the test by reducing or increasing the test price by more than $1.

A further refinement of this technique randomly allocates respondents into more than one test price, where the sequence can be repeated. If one uses three test wages $X, $Y, and $Z, with the "NOs" being tested at a higher wage (e.g., $X+$1, $Y+$1 and $Z+$1) and the "YESs" at the prices less $1, one can assess the response across nine points.[11]

Two methods of analysis are typical:

- Since respondents are randomly allocated to the test wages, it is possible to use these prices as "least wage" observations for the dependent variable in a regression. Independent variables will include the various income, household, and personal characteristics. This approach assumes that the test wage is the lowest wage at which the respondent will accept work: which is, strictly speaking, not accurate. Someone agreeing to work at $10 per hour may actually be prepared to accept a job at $9.50 per hour, but without direct testing, there is no way to measure this.

- It is also possible to use a probability model, where the decision to accept/reject a position becomes a 0-1 dependent variable in a logit or probit estimate with the usual independent variables as cited above plus the test wages associated with agreement and disagreement. This model allows one to calculate the probability of accepting a job if the wage increases by "$1" as well as the marginal impact of other independent variables on the probability of working.

[11] Appendix B presents the question flow for implementing the discrete choice tests.

Table 2 shows the mean value of the estimated reservation wage for the unemployed respondents. We computed this estimate by taking the average of the lowest test wages that a respondent stated he/she would accept for a job. It also included the stated reservation wage for those who declined work at either of the two test wages we offered. These estimates may be biased upward somewhat by the nature of the discrete choice model, where some respondents accepted work at 1.5 times and twice the minimum wage when they might have been prepared to commit at a lower wage had it been offered.

Table 2: Reservation wage
(Computed as the test wage accepted or the reservation wage offered if both test wages rejected)

Group	SA*	Non-SA	Total
	(Sample size in parentheses)		
All clients	$9.97 *(711)*	$10.67 *(1,011)*	$10.39 *(1,728)*
Single parents	$10.04 *(509)*	$10.77 *(302)*	$10.33 *(814)*
Dual parents	$9.79 *(202)*	$10.63 *(709)*	$10.44 *(914)*

* Receipt of SA in the last year (prior to interview date)

A logit model provides more useful insight into the effect of the reservation wage on the probability of working. We constructed the dependent variable as described above, and Table 3 presents the model and the diagnostics.

Table 3: Logistic regression results – reservation wage
(Total sample n = 4,392)[12]

Dependent variable = 0 or 1 based on acceptance of a job at the test wage (ACCEPT)

Variable	B	Wald	p
CONSTANT	-.856	10.055	.000
AMOUNT	.216	298.730	.000
CHILDREN $\leq$ 6	-.337	48.916	.000
CHILDREN $\geq$ 7	.054	2.229	.136
EDUCATION	-.066	20.723	.000
AGE	-.023	26.586	.000
STATUS (= 1 if Aboriginal or visible minority, 0 otherwise)	-.385	17.152	.000
COURSE (= 1 if yes in last year, 0 otherwise)	.140	3.325	.068
DISABLED (= 1 if disability in household, 0 otherwise)	.717	89.822	.000
UNEMPLOY (Provincial unemployment rate)	.053	48.298	.000
MARITAL (0 = single, 1 = dual parent)	-.086	1.289	.256
SA Status (0 = not in last year, 1 = in last year)	-.583	55.049	.002
Goodness of fit	4,515.56		
Nagelkerke R^2	.185		
Percent correctly classified	67.26%		
Chi-square *(p value)*	32.633 *(.000)*		

12 This sample was "doubled-up" as each respondent supported two observations, one at the initial test wage and another at the test wage + $1, effectively doubling the number of observations relative to the original number of respondents. This "doubled-up" sample was then reduced due to the automatic deletion of observations with missing data on one or more variables.

Overall, these are robust results. The Nagelkerke R^2 at .185 is typical of cross-section data, as is the percent correctly classified at 67.26%.[13] The Chi-square of 32.633 is highly significant and performs the same role as the F test in a standard regression model.[14] In this estimation, all coefficients are highly significant except for children $\geq$ 6 and marital status (the Wald statistic on the coefficients performs the same role as the t-test). Whether someone reported having taken a course in the last year or not is marginally significant with a p-value of .068

The interpretation of the coefficients produced by a logit model requires care. A logit model (and a probit model) assumes that the change in probability of observing a behaviour (here the respondent reporting that he/she would accept a job at the offered wage) diminishes as one moves away from the mean value of an independent variable. In contrast, a standard linear regression model assumes that the dependent variable changes at a constant rate throughout the range of the independent variable(s). Therefore, it is usual to evaluate the logit model at the mean value of the independent variables.[15] One can approximate the marginal impact of a unit change in the independent variable (at a mean value of the dependent variable, which is often assumed to be .5) by dividing the logit coefficient by 4.[16] This is a general rule that may not suit the sample if the independent variables are not symmetric about their means and if the average value of the dependent variable diverges from .5. Alternatives are to compute the marginal effect at the median or modal values.[17] Here the mean value of the dependent variable is .56 so we can safely use the simple rule of dividing by 4.

Interpreting the marginal changes for a dummy variable is more problematic since they change from 0 to 1 and a mean value makes little sense. A common approach is to simply note whether the sign is in the correct direction and the variable is statistically significant. However, here we follow the example presented in Pampel (2000; p. 27).[18]

---

13 The percent correctly classified is a useful measure of model success because it indicates the difference between the dependent variable observed from the survey (those who actually accepted or rejected a job offer at the stated wage) and the prediction of the model based on the estimated coefficients.

14 That is, it tests the null hypothesis that all coefficients are simultaneously 0 and is often termed the test of overall model significance.

15 In a general way, this is because the logistic model forms an "S" curve that is compressed between 0 and 1 on the vertical axis and the range of the independent variable(s) on the horizontal axis. The point of inflection for the "S" reaches maximum steepness at the mean of the independent variable(s), which is why most interpretation of the coefficients occurs at the mean value of the independent variable(s).

16 See Pampel (2000) for a convenient summary of logistic regression. The marginal effect of a coefficient on the probability of accepting a job, which can often be evaluated at the approximate mid-point of an independent variable, is given by $\partial p/\partial X_k = B_k * p(1 - p)$ where p is usually set to .5. This yields the general rule that the coefficient is multiplied by .25 (divided by 4) to estimate the impact on the probability for a unit change in the independent variable. Note that the marginal effect falls as p diverges from .5. For example, if p = .7, then the coefficient is multiplied by .21.

17 Some analysts dispute the value of calculating marginal effects (Demaris 1992), arguing that these models are useful only for sorting out the impact of the independent variable in a general way.

18 For dummy variables, it is possible to compute the probability of the dependent variable of the group without the dummy variable ("omitted group") and subtract that from the probability of the dependent variable for the dummy variable group.

Table 4 presents the impact of a marginal change in the independent variable on the probability of accepting a job at a specified wage.

Table 4: Estimated marginal impact on probability of accepting a job	
Variable	**Marginal impact (change in probability)***
AMOUNT	.05
CHILDREN ≤ 6	-.08
CHILDREN ≥ 7	Not significant
EDUCATION	-.02
AGE	-.01
STATUS (Aboriginal or visible minority = 1, 0 otherwise)	-.09
COURSE (= 1 if yes in last year, 0 otherwise)**	.03
DISABLED (= 1 if disability in household, 0 otherwise)	.16
UNEMPLOY (Provincial unemployment rate)	.01
MARITAL (0 = single, 1 = dual parent)	Not significant
SA Status (0 = no, 1 = dual parent)	-.14
* Dummy variable contribution calculated using change of probability for omitted and dummy variable group. Other variables calculated at p = .5. ** Marginally significant.	

Table 4 shows that a $1 change (increase) in test wages/wages offered will increase the probability of participating in the workforce by .05, while an additional child under six reduces this likelihood by .08. The addition of children over six makes no difference to the probability of labour force attachment. An additional step in the educational ladder *reduces* the probability of working by .02. This apparently contradictory result may be due to the way in which education data are collected in the survey (as an ordinal scale and not a continuous scale). It could also reflect the assumption that more educated people tend to be more discriminating in their job search. Older NCBS clients have a slightly lower likelihood of working, and older clients are less likely to have children under the age of six.[19] Being Aboriginal or a visible minority reduces the likelihood of working by .09, while having taken a course in the last year increases the likelihood of working by .03. The largest impact on willingness to take a job is whether someone in the house has a disability, and this increases the probability of working by .16. A respondent who is married or living common-law is less likely to work by .02, while a respondent who has collected SA in the last year is .14 less likely to accept a job.

For the most part, these results appear reasonable. It is important to bear the following in mind:

- The survey collected respondents' reported willingness to accept a job at a stated wage. This is not the same as observing whether a respondent accepts a job if it is really offered at that wage. This "artificiality" is an important barrier to wider acceptance of these experimental approaches.

- The high positive impact of a disability in the household on the willingness to work needs to be placed in the questionnaire context, where we offered respondents a full-time or part-time job at a specified wage. It may be that many who are disabled or who have a disabled household member (this variable includes both) are keen to work part-time and may accept lower reservation wages. Unfortunately, we cannot explore this interesting

[19] Note that simple correlations among the variables reveal very low values.

idea within these data. This illustrates an important limitation of these approaches: large samples are still needed to test various propositions.

- Having recently taken a course tends to increase the likelihood of accepting a job and may reflect the job readiness of those who participate in education and training. At the same time, those with higher levels of education appear to be slightly less willing to accept a job. Some degree of selection bias is clearly at work.

5.4 From discrete choice model to policy simulation

Within a survey sample, the administration of discrete choice options across a variety of policies and sub-samples generates behavioural parameters. The parameters form the “data” for policy simulation models that draw on distributions of the population. Using high quality samples such as the SLID or even the census provides the basis for applying behavioural parameters from discrete choice models to predict the total reaction to changes in a range of variables. Thus, changes to policies such as earnings set asides, tax rates, benefit levels etc. are founded in the predictions from discrete choice methods. Clearly the big assumption is that “stated” choice will approximate “revealed” choice. Those arguing for empirical studies using observational studies criticize these behavioural simulation techniques because they disbelieve the responses provided by survey respondents.

6.0 Conclusions

The future evaluation of the NCB has many opportunities for measuring program impact and uncovering insights into the labour market behaviour of families. It goes without saying that the statistical matching analysis completed in round 1 should be replicated.

The following two methodologies also appear worthwhile:

- qualitative studies (focus groups, dyads and case studies), especially on specific family types to generate testable hypotheses of how programs affect behaviour

- policy simulations based on parameters discovered by discrete choice methods applied to NCB clients and other low-income households.

Bibliography

Battle, K., & Torjman, S. (1994). The welfare wall: An analysis of the welfare/tax system in Ontario. In A. Maslove (ed.), *Taxation and the distribution of income.* Toronto: University of Toronto Press.

Blundell, R., Duncan, A., McCrae, J., & Meghir, C. (2000). The labour market impact of the working families tax credit. *Fiscal studies, vol. 21*.

Bourguignon, F., & Spadaro, A. (2006). Microsimulation as a tool for evaluating redistribution policies. *Journal of inequality, vol. 4,* 77-106.

Caniglia, Alan S. (1996). How large are welfare's work disincentives? *International journal of social economics, vol. 23, no. 9,* 61-68.

Cummings, Ronald G., Holt, Charles A., & Laury, Susan K. (2004). Using laboratory experiments for policy making: An example from the Georgia irrigation reduction auction. *Journal of policy analysis and management.*

Eissa, N., & Hoynes, H. (1999). The earned income tax credit and the labor supply of married couples. Institute for Research on Poverty Discussion Paper No. 1194-99, August.

Guala, Francesco. (2005). *The methodology of experimental economics.* Cambridge: Cambridge University Press.

Heath, Alexandra, & Swann, Troy. (January 1999). Reservation wages and the duration of unemployment. *Reserve Bank of Australia*, Research Discussion Paper, 1999-02.

Pampel, F.C. (2000). *Logistic regression.* Sage: Thousand Oaks, CA.

Appendix A

Summary of the potential of secondary databases to support a net impact assessment (Labour market participation)
Updated from NCB Evaluability Assessment, March 8, 2000

	Description	Dependent/outcome variable support	Independent variable coverage	Advantages	Disadvantages	Usefulness for a net impact assessment		
						low	med	high
Surveys								
Census of Population	Each census has detailed information on households. Analysis is possible on small areas, but not for families. Income and work activities are limited and no measurement of participation in NCB programs exists. Data are collected every 5 years.	Measures of poverty cannot be supported since no access exists to individual records. Crude poverty indicators can be constructed from income data. Similarly, because the data are collected every 5 years, pre-post observations are difficult to construct. Some data on labour force participation. After 5 or 6 years, public use samples can offer access to information at the individual record level.	The census has a wide array of data on families and income. It contains no data on programs such as the NCB or any details on provincial policies.	• Widely available • High reliability • Relatively low costs • Useful for general descriptions of low-income families	• Dependent and independent variables are hard to construct • Records at the family level delayed • Program intervention information absent • Periodicity of five years does not capture the NCB	●		
Labour Force Survey (LFS)	This large sample (50,000), monthly survey collects data on labour force activity. It collects limited information on the household and further questions are needed to measure incomes, participation in NCB programs, etc. While it uses a rotating sample, it is complex and requires special techniques. This has been folded into the SLID.	NA see SLID below	NA (Combined with SLID)	NA	NA			
Survey of Consumer Finances (SCF)	This is an annual survey that uses the same design as the LFS. It has detailed income and labour force activity data. The SCF has been folded into SLID since 1993.	NA (see SLID below)	NA (combined the SLID)	NA	NA			

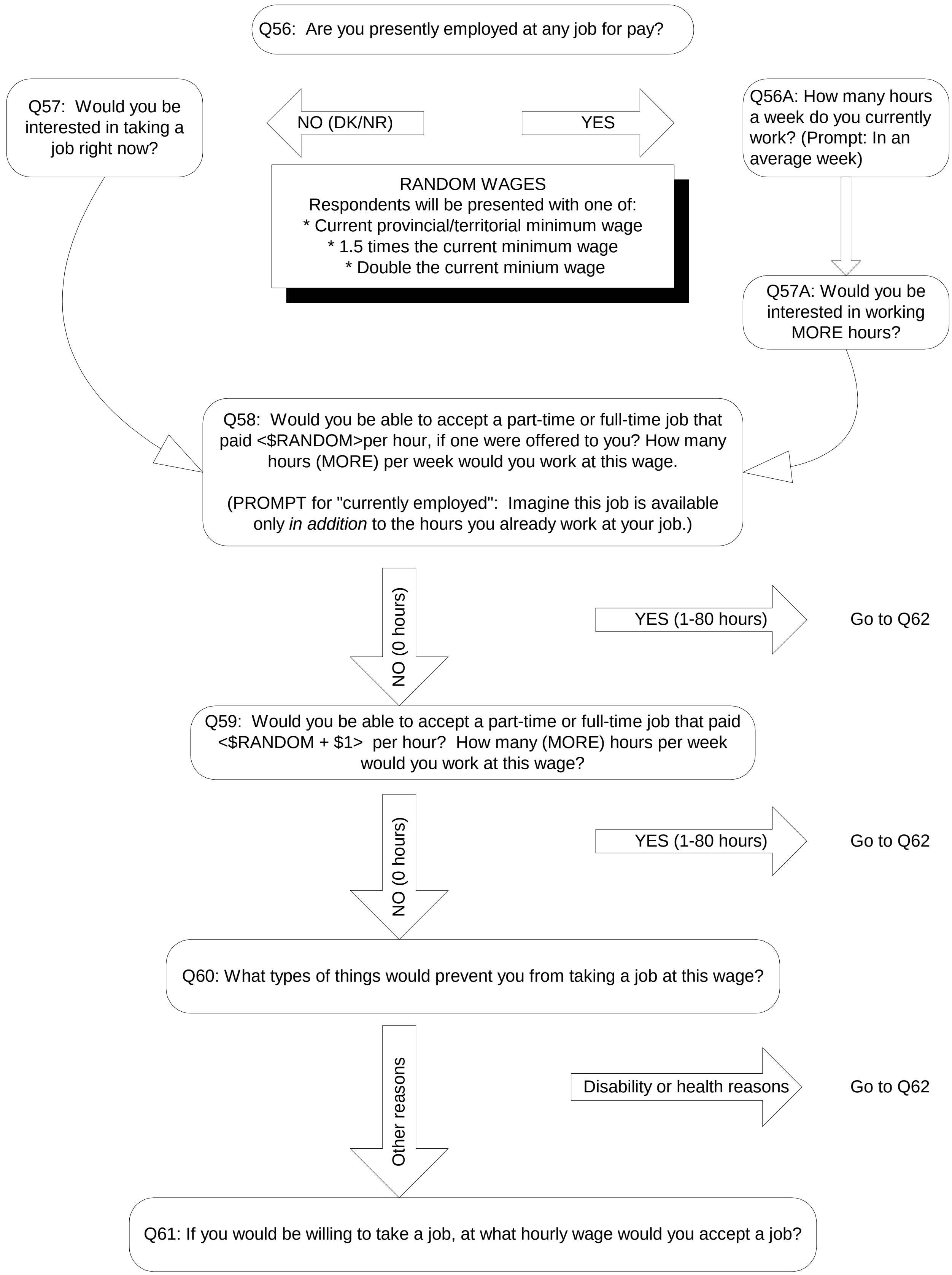
Summary of the potential of secondary databases to support a net impact assessment (Labour market participation)
Updated from NCB Evaluability Assessment, March 8, 2000

	Description	Dependent/outcome variable support	Independent variable coverage	Advantages	Disadvantages	Usefulness for a net impact assessment		
						low	med	high
Survey of Labour and Income Dynamics (SLID)	This replaces the SCF. Data began to be collected in 1993 for both cross-sectional and longitudinal samples. The longitudinal survey runs for 6 years to 2001. It represents the best single data source for the NCB impact assessment. It is also the basis for calculating statistics on poverty measures.	This dataset is being used to support measures of poverty and has the required outcome measures for the net impact assessment.	It contains a rich array of independent variables. It lacks program detail on each province.	A stable, well-developed survey with high reliability (although it has incurred some sampling bias issues from time to time).	• Lacks information on childcare • Lagged by 18 – 24 months • Insufficient program detail			●
General Social Survey	This annual survey collects information on social trends. With a sample of 10,000 across the provinces, it cycles through themes on education and work, family and friends and social support. Little detail on income and labour force activity.	This survey does not generate the outcome variables needed for an impact assessment.	It also does not generate the data needed to construct the needed independent variables.	The GSS can support single studies on the general trends in the family and therefore offers "collateral" monitoring of the state of the family.	• Sample sizes are small in each jurisdiction • Not suitable for testing hypotheses about earnings supplements or child benefits		●	
Administrative data								
Social Assistance Data (HRDC source)	Collected by HRDC from 6 (soon to be 8) jurisdictions. Varying range and type of variable collected. Some offer 100% others a 10% sample. Most series are annual back to 95/96.	In principle, this should be a very high quality source of information on labour force activity of SA clients. Not suitable for poverty indicators. HRDC extract does not maintain uniform data/variables among the jurisdictions.	These data usually have detail on the family and sources of income. Program participation may also be present.	• Detailed information on SA clients • Longitudinal • Linkable to TI data	• HRDC version is uneven in terms of content • Coverage very uneven among reporting provinces • This project could prove a useful source to the NCB in a few years	●		
Social Assistance Data (Jurisdictions source)	Potentially available from each jurisdiction. Cost and confidentiality are important barriers. Most jurisdictions can provide monthly data back to 1995.	Labour force participation of SA clients is kept for most jurisdictions.	Most relevant variables on the family can be collected.	• Detailed information on SA clients • Longitudinal • Linkable to TI data • Reasonable consistency among a majority of jurisdictions • Supports incidence of poverty and labour force supply	• Cost and confidentiality barriers exist with jurisdiction provision • Selected analysis by jurisdiction is feasible		●	

Summary of the potential of secondary databases to support a net impact assessment (Labour market participation)
Updated from NCB Evaluability Assessment, March 8, 2000

	Description	Dependent/outcome variable support	Independent variable coverage	Advantages	Disadvantages	Usefulness for a net impact assessment		
						low	med	high
EI Status Vector	Offers detailed longitudinal information on beneficiaries to EI since 1971.	Not useful for NCB						
NESS Intervention (EI Part II)	Includes income benefit data for EI clients	Not useful for NCB						
Record of Employment (ROE)	This database tracks job leaving of all Canadians contributing to the Employment Insurance.	Not useful for NCB						
T1 Longitudinal File (Longitudinal Administrative Database – LAD)	The most reliable source of information on individual incomes and sources of income. Can be combined to produce data on the family unit.	Contains all the data collected on tax forms. Could generate measures of poverty to infer labour market attachment. Measure of earnings by source are reliable.	Some key variables can be extracted for use in labour supply and property incidence such as: • dependents • disability (not by type) • dependent exemptions	• High reliability in the income data • Can be used to produce variables that describe participation in certain NCB programs • Can be a complement to the SLID • It is feasible for creating a family taxation file	• Does not contain socio-economic data needed to include in the net impact assessment • Requires strong justification for access		●	
Program data from NCB reinvestments	Few NCB reinvestments programs can track recipients. Some such as the earnings supplements in Saskatchewan offer important potential for labour supply analysis.	Earnings would track labour supply directly and could be very useful for assessing family labour supply.	Data would contain many key independent variables that could be supplemented by surveys.	Detailed information on income and earnings for selected provinces.	These databases do not represent Canada or regions other than where these reinvestments are implemented. There may be a lack of generality.		●	

Appendix B


Q56: Are you presently employed at any job for pay?
NO (DK/NR)
YES
Q57: Would you be interested in taking a job right now?
RANDOM WAGES
Respondents will be presented with one of:
* Current provincial/territorial minimum wage
* 1.5 times the current minimum wage
* Double the current minium wage
Q56A: How many hours a week do you currently work? (Prompt: In an average week)
Q57A: Would you be interested in working MORE hours?
Q58: Would you be able to accept a part-time or full-time job that paid <$RANDOM>per hour, if one were offered to you? How many hours (MORE) per week would you work at this wage.
(PROMPT for "currently employed": Imagine this job is available only *in addition* to the hours you already work at your job.)
NO (0 hours)
YES (1-80 hours)
Go to Q62
Q59: Would you be able to accept a part-time or full-time job that paid <$RANDOM + $1> per hour? How many (MORE) hours per week would you work at this wage?
NO (0 hours)
YES (1-80 hours)
Go to Q62
Q60: What types of things would prevent you from taking a job at this wage?
Other reasons
Disability or health reasons
Go to Q62
Q61: If you would be willing to take a job, at what hourly wage would you accept a job?