Estimating Labor Supply Responses and Welfare Participation: Using a Natural Experiment to Validate a Structural Labor Supply Model*

Jorgen Hansen† & Xingfei Liu‡

October 19, 2012

Abstract

We estimate an economic model of labor supply and welfare participation using data on single men from Quebec drawn from the 1986 Canadian Census. Detailed budget sets for each work-welfare combination - accounting for income taxes, tax credits and welfare benefit rules - are derived using a micro-simulation model. Predictions generated by the model replicate actual changes in labor supply and welfare participation due to a welfare reform that took place in 1989. The advantage of having estimated a structural model is illustrated by showing how labor supply and welfare participation change when income taxes and benefit levels change.

Keywords: labor supply; welfare participation; unobserved heterogeneity; natural experiment; regression discontinuity; micro-simulation

*We thank Rolf Aaberge, Olivier Bargain, Christian Belzil, Richard Blundell, Lennart Flood, Peter Gottschalk, Fabian Lange, Bruce Shearer and conference and seminar participants at CSLRN, ESEM in Malaga, Goteborg University, IZA and University of Montreal for helpful comments. We are also grateful to Kevin Milligan for providing detailed information on tax and benefit parameters. This is a revised version of Hansen and Liu (2011).

†Department of Economics, Concordia University, 1455 de Maisonneuve Blvd W, Montreal, Quebec, H3G 1M8, Canada, CIRANO, CIREQ and IZA.

‡Department of Economics, Concordia University, 1455 de Maisonneuve Blvd W, Montreal, Quebec, H3G 1M8, Canada.

1 Introduction

Structural models that describe agents' actions are common in a wide range of topics in empirical economics and it is important to provide evidence on the validity of such models. This is especially true given that such models often rely on parametric or functional form assumptions. One validation strategy is to compare how well predictions from the model compare with actual observations and reject models that do not fit the data. However, a model whose outcomes compares favorably with the (within-sample) data may not be able to correctly forecast the impact of a large policy change. Indeed, as mentioned in Keane and Wolpin (2007), forecasting the impact of large changes in the economic environment is a great challenge for structural models. Thus, a more appropriate approach to validate the model is to examine the behavioral implications of the model, in combination with within-sample model fit tests.

Ideally, validation of structural models requires access to randomized social experiments or large policy changes. Unfortunately, such opportunities are rare. One example is McFadden (1977) who utilized a large regime shift (the introduction of the Bay Area Rapid Transit (BART)) in his work on forecasting the demand for rail transport in the San Francisco Bay area. He estimated a structural model of travel demand before the introduction of the BART system and compared the forecast generated by his model to actual usage after BART's introduction. Other examples include Lise, Seitz and Smith (2005) who used data from the Self-Sufficiency Project, a Canadian large-scale experiment designed to move people off welfare and into work, to validate a calibrated search-matching model of labor market behavior. Similarly, Todd and Wolpin (2006) used data from a large-scale school subsidy experiment in Mexico, where villages were randomly assigned to control and treatment groups. Using only the control group villages, they estimated a behavioral model of parental decisions about child schooling and work, as well as family fertility. The validity of the model was then assessed according to how well it could predict the behavior of households in the treatment villages. Keane and Wolpin (2007) used

a "non-random holdout sample" instead of a randomized social experiment or a large policy change. Their holdout sample differed significantly from the estimation sample along the policy dimension that the model was meant to forecast. Finally, Choi (2012) uses welfare reform experiments in Minnesota and Vermont to assess the capacity of a structural labor supply model to correctly predict changes in labor supply.[1]

In this paper we exploit a large policy change that took place in the Canadian province of Quebec in 1989. Prior to the change, unattached persons younger than 30 years old received substantially less in welfare payments than similar individuals 30 years of age or older.[2] This form of age discrimination was terminated in 1989 and monthly benefits increased from $185 to $507 (an increase of 175%) for those younger than 30 years old. A similarly large policy change is unlikely to happen again and it provides an opportunity for researchers to evaluate the impact of dramatic increases in welfare generosity on welfare use. Indeed, this policy has been utilized to assess how the benefit increase affected duration of welfare spells (Fortin, Lacroix and Drolet, 2004) and the impact on employment and welfare participation rates (Lemieux and Milligan, 2008).

Our approach to validate our structural model is somewhat different from that used by McFadden (1977) who also utilized a large policy change. In particular, we first estimate our model on data from the 1986 Census (this Census was also used by Lemieux and Milligan, 2008). We then use our estimated preference parameters and predict outcomes given changes in the welfare benefit rules corresponding to the policy change. Finally, we compare our estimated impacts on employment and welfare participation with those obtained using a Regression Discontinuity (RD) approach. We also compare our results with those in Lemieux and Milligan (2008).

This paper makes two major contributions to the literature. First, we are able to assess the performance of a structural model of labor supply

[1] In addition to these papers, a number of studies have evaluated non-experimental methods by comparing estimates with experimental results, see LaLonde (1986), Heckman and Hotz (1989) and Friedlander and Robins (1995).

[2] Recipients were required to complete a form each month allowing officials to determine if respondents had reached age 30.

and welfare participation beyond a simple comparison of actual and predicted distributions of hours of work and welfare use. This should be very valuable as the capacity of this type of labor supply models to recover true effects under stringent parametric assumptions is virtually unknown.

Secondly, this paper illustrates the benefits of having access to a set of estimated preference parameters along with a structural model of labor supply behavior. In particular, our model allows us to recover a distribution of welfare responses as we marginally change benefit levels. This distribution is not identified in studies such as Lemieux and Milligan (2008), where only a single effect corresponding to a particular benefit change is identified.

Our results are encouraging as our preferred model specification predicts employment reductions, as a result of the dramatic increase in welfare benefits, similar to those obtained using RD. We also show that the changes in labor supply and welfare use are smaller for high school graduates than for high school drop-outs. Moreover, the effects are largest among those with lowest incomes. For those in the lowest income quartile, the benefit increase is predicted to reduce employment by 25 percent and increase welfare participation with 20 percent. We further illustrate how employment, welfare use and hours of work change as we marginally increase social assistance benefits. The responses to these benefit changes are nonlinear which suggest that results from experimental studies may only be of limited value when considering policy changes that differ from those used in the experiment. Finally, we present results on how labor supply and welfare use change as we change the income tax system.

The remainder of this paper is organized as follows. Section 2 provides a brief description of social assistance and income taxes in the province of Quebec. Section 3 presents the economic model and the empirical specification while Section 4 describes extensions of the basic model. The data used in the analysis are presented in Section 5 and the results are discussed in Section 6. Section 7 concludes the paper.

2 Social Assistance and Income Taxes in Quebec

Prior to 1996, Canada's social assistance (or welfare) system was administered under the Canada Assistance Plan (CAP), an arrangement that allowed the cost of social policies to be shared by the federal government and the provinces and territories. On April 1, 1996, the Canada Health and Social Transfer (CHST) replaced the CAP and this meant that expenditures on health, education and social assistance became provincial jurisdictions. Consequently, there are thirteen distinct social policy systems in Canada as each province and territory designs, administers and delivers one of its own, although some federal assistance programs, most importantly the Canada Child Tax Benefit (CCTB), are also in place. All the provincial systems have complex rules which regulate the eligibility, rates of assistance, amounts of income and assets exempted from the calculation of the needs test, and other system specifics.

Unlike the U.S., social assistance in Canada provides any type of household, including those consisting of single persons without children, with financial support. In general, the only eligibility requirement is a household's need, regardless of cause. Eligibility for social assistance is established using a needs test.[3] That is, the total value of a household's non-exempted assets is calculated and compared with the maximum allowable level. Then, provided that a household's assets do not exceed the allowable level, the household's income from non-exempted sources is compared with its basic needs (food, clothing, shelter and utilities, household necessities, etc.), and if the cost of these basic needs exceeds the income, the household is eligible for social assistance.

In Quebec, social assistance programs are administered by "Ministère de l'Emploi et de la Solidarité sociale" and eligibility requires Quebec residency and that the applicant is 18 years or older, in addition to passing the basic needs test.

[3] Although the obligation to conduct a needs test was abolished when the Canada Health and Social Transfer replaced the Canada Assistance Plan in 1996, no province or territory has yet disposed of it.

Regarding income taxes, the Quebec system was - and still is - composed of two parts, a provincial component and a federal component. In general, similar tax rules regarding exemptions and deductions apply but individuals file separate returns. However, the marginal tax rates as well as the income levels at which they change differ between the two jurisdictions. For example, in 1985 (the year for our analysis) there were 10 federal tax brackets and 21 provincial tax brackets. The marginal tax rates are presented in Figure 1 and illustrate the high degree of progressivity that applied to income taxes in that year. As the figure shows, the marginal tax rates at the federal and provincial levels were similar, especially for incomes exceeding $38,000 per year. The highest tax rate was 34 percent at the federal level and 33 percent at the provincial level. These high tax rates applied to earnings above $62,657 (federal) and $60,715 (provincial). Both the number of tax brackets and the level of the tax rates were substantially reduced during the late 1980s. During the same period, the basic tax exemption at the federal level was increased, further lowering the average federal tax rate.

3 Economic Model and Empirical Specification

The traditional way to model labor supply assumes that the decision variable, hours of work, is continuous. However, this generally implies that restrictive assumptions must be made in order to guarantee statistical coherency (see for instance the discussion in MaCurdy, Green, and Paarsch (1990)). Moreover, an underlying assumption in traditional labor supply models is that the individual (or household) budget set is convex. Hence, to estimate a continuous labor supply model, a number of important simplifications of the income tax and transfer system must be made.

In this paper, we model labor supply as a discrete choice instead, following previous work by van Soest (1995), Hoynes (1996), Keane and Moffitt (1998), Blundell et al. (2000), Flood, Hansen and Wahlberg (2004) and Blundell and Shephard (2012). Unlike the continuous labor supply model, the discrete choice model allows us to include as many details as possi-

ble regarding the budget set. Specifically, we assume that each individual can choose among all the alternatives in the choice set of consumption (C)-leisure (L) combinations $\left(C_j, L_j\right)$, where $j = 1, ..., J$. Further, $L_j = TE - h_j$ where TE denotes total time endowment and is set to 4,000 hours per year.[4] Thus, the choice set for an individual contains J different hours of work combinations. In the empirical part of the paper, we set $J = 7$.[5]

We assume that utility depends not only on consumption and leisure, but also on participation in welfare programs. We further assume that the utility function is increasing in consumption and leisure and decreasing in welfare participation. The disutility from participation in a welfare program is assumed to primarily reflect the non-monetary costs associated with participation in such programs, such as fixed costs or "stigma", and is included to account for nonparticipation among eligible individuals.[6]

Following van Soest (1995), we use a trans-log specification of the direct utility function, and for any specific individual we have

$$\begin{aligned} U\left(C_j, L_j\right) =& \beta_C log\left(C_j\right) + \beta_L log\left(L_j\right) + \beta_{Csq}\left(log\left(C_j\right)^2\right) + \\ & \beta_{Lsq}\left(log\left(L_j\right)^2\right) + 2\beta_{C,L} log\left(C_j\right) log\left(L_j\right) - \phi_{SA} d_{SA} \\ j =& 1, ..., J \end{aligned}$$

where it is assumed that the disutility from receiving social assistance (ϕ_{SA}) is separable from the utility of leisure and consumption (following Moffitt (1983) and Hoynes (1996)).

The individual chooses leisure time (L), social assistance status (d_{SA})

[4] TE can also be regarded as a parameter that can be estimated together with all other parameters.

[5] Specifically, we set $h_1 = 0, h_2 = 500, h_3 = 1,000, h_4 = 1,500, h_5 = 2,000, h_6 = 2,500$ and $h_7 = 3,000$. Previous research has shown that the performance of discrete choice labor supply models similar to the one used in this paper is robust to alternative values of J (e.g. van Soest (1995), Gong and van Soest (2002) and Flood and Islam (2005)).

[6] What may appear as "stigma" or disutility from welfare participation may also result from the inability of the econometrician to measure true welfare eligibility. Moreover, imperfect information regarding benefit eligibility on behalf of the individual is also included in this non-monetary cost.

and consumption (C) by maximizing utility subject to the following budget constraint:

$$C_j = I_j + B_{SA}\left(I_j\right) d_{SA}$$

where

$$I_j = Wh_j + Y_j - t\left(Wh_j + Y_j^T - D_j\right)$$

C_j is the income net of taxes at hours combination j $\left(h_j\right)$, W equals the before-tax hourly wage rate, Y_j denotes annual non-labor income, $t(.)$ is a function that determines income taxes, Y_j^T is taxable non-labor income, and D_j represents deductions.[7] Finally, $B_{SA}(.)$ is a function that determines social assistance benefits.[8]

The addition of the disutility of welfare participation implies that an individual faces $2J$ work-welfare possibilities. However, some welfare states may be infeasible if the individual's income from work is sufficiently high to render him ineligible for SA. Solving the optimization problem requires evaluating the utility function for each possible combination of hours of work and welfare program participation and choosing the state that yields the highest utility.

To make the model operational, random disturbances are added to the utilities of all choice opportunities:

$$U_{j,r} = U\left(C_j, L_j\right) + \varepsilon_{j,r}$$

where j represents the individual's choice of labor supply, r represents the individual's welfare participation state and $U_{j,r}$ denotes the individual utility of choice (j, r). We assume that $\varepsilon_{j,r}$ follows a type I extreme value distribution with cumulative density $Pr\left(\varepsilon_{j,r} < \epsilon\right) = exp\left(-exp\left(-\epsilon\right)\right)$. The

[7] Non-labor income and deductions are indexed by j since they may depend on earnings.

[8] Apart from consumption or income, eligibility for welfare may also depend on wealth. However, there is no wealth information in the data and we therefore ignore this possibility. Given our sample selections (single males with at most a high school diploma) we do not believe that this assumption invalidates our results in any significant way.

error term $\varepsilon_{j,r}$ can be interpreted as an unobserved alternative specific utility component or as an error in an individual's assessment of the utility associated with choosing the work-welfare combination (j, r) (optimization error). Given the distributional assumptions of the stochastic terms in the utility function, the contribution to the likelihood function for a given individual is

$$l = \sum_{r=1}^{2} \sum_{j=1}^{J} p_{j,r} \delta_{j,r}$$

where

$$p_{j,r} = \frac{exp\left(U_{j,r}\right)}{\sum_{s=1}^{2} \sum_{t=1}^{J} exp\left(U_{t,s}\right)}$$

and where $\delta_{j,r}$ is an indicator for the observed state for each individual.

4 Extensions of the Basic Model

4.1 Heterogeneity in Preferences

Previous studies have shown the importance of allowing for heterogeneity in preferences (e.g. Flood et al. (2004)). Here, we introduce heterogeneity in preferences for leisure and welfare as follows

$$\beta_L = \sum_{k=1}^{K} \beta_{L,k} x_k + \theta_L$$
$$\beta_C = \sum_{k=1}^{K} \beta_{C,k} x_k + \theta_C$$
$$\phi_{SA} = \sum_{k=1}^{K} \beta_{SA,k} x_k + \theta_{SA}$$

where the elements of the vector x are observed individual character-

istics (age, education and immigrant status). K denotes the dimension of the vector x, while the $\theta's$ represent unobserved variables that affect preferences for leisure, consumption and welfare.

It is reasonable to assume that an important source for population heterogeneity is unobserved. In order to account for this, we formulate a finite mixture model, which allows for unobserved heterogeneity in a flexible way without imposing a parametric structure. This way of representing unobserved heterogeneity is similar to what Heckman and Singer (1984) suggested for duration data models. We assume that there exist M different sets of $(\theta_L, \theta_C, \theta_{SA})$ that determine an individual's preferences, each observed with probability π_m (where $\pi_m > 0$ and $\sum \pi_m = 1$, $m = 1, ..., M$. This specification allows for an arbitrary correlation between the individual's work effort, preference for consumption and preference for welfare participation.

Given the distributional assumptions of the stochastic terms in the utility function, the contribution to the likelihood function for a given individual is

$$l = \sum_{m=1}^{M} \pi_m \left\{ \sum_{r=1}^{2} \sum_{j=1}^{J} (p|\Theta)_{j,r} \right\} \delta_{j,r}$$

where

$$(p|\Theta)_{j,r} = \frac{exp\left(U_{j,r}|\Theta\right)}{\sum_{s=1}^{2} \sum_{t=1}^{J} exp\left(U_{t,s}|\Theta\right)}$$

and where $\Theta = \{\theta_L, \theta_C, \theta_{SA}\}$. This expression simply denotes the probability that the utility in state (j, r) is the highest among all possible work-welfare combinations, conditional on unobserved preferences.

4.2 Fixed Costs of Work

The model fit obtained from the basic model reveals that such a simple representation of preferences is unable to capture the observed distribution of

hours of work in our sample. In particular, the fraction of part-time workers is overestimated and the peak around full-time is underestimated. A similar problem has been noted in previous work (e.g. Dickens & Lundberg (1993), van Soest (1995), Aaberge, Colombino and Strom (1999) and Flood et al. (2004)).

Aaberge et al. (1999) address this issue by specifying a general labor supply model that incorporates job offers with different contracted hours of work. In this paper we adopt a simpler approach following van Soest and Das (2001). In particular, we replace $log\,(C)$ for employed individuals with $log\,(C) - log\,(FC)$ where FC is a parameter to be estimated that denotes fixed costs of work. Since utility increases with income, positive costs (FC) reduce the utility of working while having no effect on the utility associated with not working.[9]

5 Data

5.1 Description of the Data and Sampling Procedures

The data employed in this study are drawn from the 1986 Canadian Census.[10] The data are extracted from a sample of 500,434 individuals who accounted for approximately two percent of the Canadian population at the time.[11] The data contain information on demographic, social and economic data such as income, age, education and immigration status.

In this paper we attempt to obtain a sample that resembles the one used in Lemieux & Milligan (2008) as much as possible. However, minor differences exist since we do not have access to the restricted Census master files. Like Lemieux & Milligan (2008), we restrict our sample to single

[9] In addition to adding heterogeneity in preferences and fixed costs of work, we have also explored the possibility that wages are measured with errors. Similar to the results in van Soest (1995), we found no evidence suggesting that incorporating this feature into our model improved its performance.

[10] Specifically, we utilized the public use micro-data file on individuals from the 1986 Census of Canada.

[11] The target population for the 1986 Census includes all individuals except for institutional residents and residents of incompletly enumerated Indian reserves or settlements.

males without children residing in the province of Quebec in 1985. Furthermore, we remove males younger than 20 and older than 39 as of June 3rd in 1986. Unlike Lemieux & Milligan (2008), we remove married males from the sample.[12] While Lemieux & Milligan (2008) focus on a sample consisting of those who dropped out of high school, we also include individuals who have graduated from high school but have acquired no further education. Removing high school graduates from our sample yields a sample of only 327 males. Our final sample, including high school graduates, consists of 1,034 males.

5.2 Variable Definitions

In order to estimate our structural labor supply model, we need information on individuals' labor supply decisions and welfare participation. More importantly, our model assumes that individuals optimally select themselves into different work-welfare categories. Thus, we need to calculate net (or disposable) income for each person and each work-welfare alternative in the model. These calculations are based on a small micro-simulation model that we constructed using details from the 1985 Quebec and Federal Income Tax Schedules.[13]

In order to create annual hours of work for each individual, we combine information on weeks worked in 1985 with hours worked in the reference week (which occurred in 1986). This measure includes working for wages, salary, tips or commission, working in own business, farm or professional practice, or working without pay in a family business or farm owned or operated by a relative living in the same household. Males who worked more than 1 week but less than 13 weeks during the year were excluded from our sample. Similarly, males who worked more than 1 hour but less than 4 hours during the reference week were also excluded. We further

[12] Including married individuals would necessitate a different economic model that accounts for joint household decisions.

[13] We utilized information on tax and benefit parameters kindly provided by Kevin Milligan (see Milligan, 2008). The input in the micro-simulation model is earnings at hours option j ($j = 1, ..., 7$). Using earnings, we calculate deductions, tax credits and income taxes at both the federal and provincial level.

removed males who reported more than 3,640 hours of work in 1985.

The 1986 Census collects yearly wage and salary income, which refers to gross wages and salaries before deductions for items such as income tax. Military pay and allowances, tips, commissions, bonuses and piece-rate payments as well as occasional earnings in 1985 are also included in this variable. Hourly wages are calculated by dividing the yearly wage and salary income measure by our constructed measure of annual hours of work. We exclude males who received self-employment income and those whose hourly wage fell below the minimum wage level in Quebec in 1985. Moreover, we excluded males whose hourly wage exceeded $68.[14]

Unfortunately, the Census data contain no direct information on social assistance (or welfare). However, there is information on whether the respondent received any government transfer payments apart from child related transfers (family allowances and federal child tax credits), pensions (Old Age Security, Guaranteed Income Supplements and Canada/Quebec Pension) and unemployment insurance benefits. For our sample of single, young males, we believe that this measure provides a reasonable measure of social assistance receipt. As mentioned in Lemiuex & Milligan (2008), "... social assistance benefits are by far the largest component of the "other transfers" variable".[15]

5.3 Descriptive Statistics

In Table 1 we present average employment and welfare participation rates for a sample that includes both high school drop-outs and high school graduates as well as separately for the two groups. The employment rates are substantially lower for males without a high school diploma. Defining employment using information on weeks worked in 1985, the employment rate is 0.607 for high school drop-outs while it is 0.867 for those with high school as their highest degree. An alternative employment definition,

[14] As is standard in this literature, we used predictions from a Mincer-type of wage regression for all non-workers.

[15] Lemieux and Milligan (2004) show that 85 percent of income included in "other transfers" in the Census is social assistance payments.

based on work status during the Census reference week, is available in the data. Again, we find that there is a large employment gap between the two groups, 0.638 for drop-outs compared to 0.897 for high school graduates. Regardless of definition, the employment rates for high school drop-outs are comparable to those presented in Lemieux and Milligan (2008).

Table 1 also shows the proportion of respondents that received any social assistance benefits. As for employment rates, there are large differences in welfare participation across educational attainment and the proportion of recipients among high school graduates is 0.107 while it is 0.252 for high school drop-outs. The figure for the latter group again compares favorably to the proportion in Lemieux and Milligan (2008). Overall, labor market outcomes of our sample of high school drop-outs is quite similar to those presented in Lemieux and Milligan (2008), although the sample sizes differ.

In Table 2, we present sample averages of selected individual characteristics in our sample, separate for those with and without social assistance. As shown in Table 1, Table 2 reveals that welfare participants generally have lower educational attainment (only 48.1 percent are high school graduates). Furthermore, welfare recipients are slightly older while the fraction of the sample that was born in a foreign country is higher among those who did not receive welfare. Finally, only a small fraction of those who received welfare were employed.[16] Also, for the welfare recipients who worked, they worked fewer hours than those who did not claim welfare benefits (250 hours per year versus 1,680 hours).

6 Results

6.1 Model Comparison and Validation

We have estimated a number of different specifications of the structural labor supply model described above and used the estimated preference parameters to predict changes in employment and welfare use as a result

[16]This is true regardless of how employment is defined (employment rate last year or during Census week). The employment rate in Table 2 is based on employment rate last year.

of the introduction of higher welfare benefits for those under the age of 30. The results are presented in Table 3.[17] We assess each specification using three criteria. First, we compare predictions of annual hours of work with those observed in the data. Using the seven classes of hours of work, we calculate a goodness-of-fit statistic that is chi-square distributed. These statistics are reported in column one. Secondly, we compare the predicted changes in employment and welfare use with those obtained using RD methods.[18] Finally, we use the Akaike Information Criteria (AIC) to further discriminate among the model specifications.

The first set of rows show results from model specifications where the preference parameters associated with leisure, consumption and welfare are parameterized as functions of observed characteristics (age, education and immigrant status) but where there is no unobserved heterogeneity embedded in the preferences. Within this class of models, we considered alternative representations of costs of work (see Section 4.2 above). First, the model was estimated ignoring such costs altogether. The model fit statistic in this case is large (287.8) and the null hypothesis of similarity between model predictions and data is rejected. Hence, this model specification does not pass the first of our validation criteria. Furthermore, the estimated impacts of the welfare reform on employment and social assistance are substantially smaller than those obtained using RD.

The second model specification, within the class of models that assume away unobserved heterogeneity, includes an additional parameter designed to pick up monetary and non-monetary costs associated with different hours of work arrangements. Although the addition of this parameter substantially improves the model fit (the chi-square statistic drops from 287.8 to 74.5) the null hypothesis of similarity between model predictions and data is still rejected. The simulated impacts of the welfare reform on employ-

[17] Complete estimation results are available upon request.

[18] Like Lemieux and Milligan (2008), our RD results were obtained from regressions of the outcome variable (employment, welfare or hours of work) on a treatment indicator plus a function, $\delta(age)$, that captures the effect of age on the outcome. As mentioned by Lemieux and Milligan (2008), the key identification assumption is that $\delta(age)$ is a smooth function. The estimates reported in Table 3 were obtained using a linear spline specification for age.

ment and welfare use are substantially larger than for the model without a control for fixed costs of work. The impacts are also larger than those obtained using RD. Thus, although the addition of fixed costs to the model specification improved the performance of the model, it does not pass the validation criteria discussed above.

The remaining sets of results in Table 3 are all based on model specifications where we have added unobserved heterogeneity to the preferences for leisure, consumption and welfare use as discussed in Section 4.1 above. As is generally the case with these type of models, we start by assuming a distribution with two support points and then proceed if there is an improvement in AIC.

The results obtained with two support points suggest a dramatic improvement in model fit for the model specification with fixed costs.[19] According to all three selection criteria, the model with unobserved heterogeneity outperforms the one without these heterogeneity terms. As we add support points, performance improve and for the specification with four types, we fail to reject the null hypothesis that model predictions and data are similarly allocated across the seven hours classes. Moreover, the estimated effects of the welfare reform on employment and welfare participation are close to those obtained using RD and statistically significant. The employment rate is predicted to decrease with 4.2 percentage points (compared to 4.1 using RD) while social assistance is predicted to increase with 3.9 percentage points (compared to 1.3 using RD). Further, annual hours of work are predicted to drop with 38 hours using our model and with 26 hours using RD.

Finally, the AIC for the model with four support points is 3833. This is lower than the corresponding values for any of the models with fewer support points. Adding additional number of support points did not improve the model fit nor did it lower the AIC. Thus, using our validation criteria, we determined that the appropriate number of support points is four.[20]

[19] Given the inferior performance of the model without fixed costs above, we do not consider that model specification further.

[20] Cameron and Taber (1998) also use AIC to estimate the number of support points in

To summarize, the model specifications that include unobserved heterogeneity in preferences for leisure, consumption and welfare and that also incorporate some measure of non-monetary and monetary costs associated with working provide estimated effects of the 1989 welfare reform that resembles the true effects, as estimated by RD on our sample. It is perhaps remarkable, but certainly encouraging, how well this relatively simple model performs, especially given the large change in policy environment that the reform implied.

In addition to this external validation, we explored how the predicted hours distribution compared with the distribution observed in our data. As shown in the first two columns in Table 5, the estimated model matches the observed hours distribution very well. Another test of the model's performance is illustrated in the last two columns in Table 5 where we present predicted and observed employment in 1990. To obtain the predicted distribution, we utilized sample data from the 1991 Canadian Census and tax and benefit parameters of the 1990 income tax system.[21] This is clearly a challenging exercise since time does not enter our model. Notwithstanding this, the predicted distribution of hours of work resembles the observed distribution very well. We slightly over predict non-employment (16.1 percent in the data compared to 19 percent in our model) and under predict the fraction of males working more than 2,500 hours per year (8.3 percent in the data compared to 4.7 percent in our model).

6.2 Heterogeneity in Responses to the 1989 Welfare Reform

As noted above in Section 5.1, our sample includes both high school dropouts and high school graduates. However, we allow the preference parameters for consumption, leisure and welfare to depend on educational attainment and we can therefore obtain effects of the welfare reform separately for the two groups. This will enable us to compare our predicted responses to those reported in Lemieux and Milligan (2008) who used a

their Monte Carlo study of dynamic, discrete choice models.

[21] The sample selection criteria was the same as that for our estimating sample.

sample of high school drop-outs.

In Table 4 we present estimated impacts of the reform for different groups using estimated behavioral parameters from a model that accounts for fixed costs of work and that assumes a four-support distribution for unobserved heterogeneity. For high school drop-outs, the reduction in the employment rate is 5.0 percentage points. This figure is similar to the estimates reported by Lemieux and Milligan (2008), providing additional support for the validity of our structural model.[22] The predicted increase in social assistance use for high school drop-outs is 4.5 percentage points, which again compares very favorable to the increase reported in Lemieux and Milligan (2008).[23] Finally, the reform is predicted to reduce annual hours of work for this group by 44 hours which corresponds to a 3.7 percent reduction. We also note that all of the impact of the reform happens at the extensive margin (participation) instead of the intensive margin (hours of work conditional on working). Similar findings were reported in Lemieux and Milligan (2008).

Table 4 also shows the estimated effects of the reform for high school graduates and separately for natives and immigrants. The changes in labor supply and welfare use are smaller for high school graduates than for high school drop-outs. For example, the employment rate effect is about three quarters of that for high school drop-outs. It is also noteworthy that natives respond more to the benefit change than immigrants.

Finally, Table 4 shows the reform's impact on males located at different places in the distribution of net incomes.[24] As expected, the effects are largest among those with lowest incomes. Moving up the income distribution, the effects of the reform fade and among those in the top quartile, the benefit increase has no effect on either labor supply or welfare participation.

To summarize, the entries in Table 4 provide additional insights, be-

[22] Their estimated employment effects range from -0.056 to -0.038, depending on definition of employment and specification of the RD regression equation.

[23] Their estimated welfare effects range from 0.033 to 0.041.

[24] We use net incomes at predicted hours of work.

yond those obtained using RD estimates, into the expected effects of a very large benefit increase for a portion of the population. In addition to provide support for our simple structural model of labor supply behavior, the results in Table 4 show that the reform will impact behavior differently for different groups in the population.

6.3 Labor Supply and Welfare Responses to Alternative Benefit Changes

A major limitation of studies exploiting "natural" experiments is the inability of extrapolating the results to other policy changes. The responses estimated using RD methods apply to a particular welfare reform where benefits for those under the age of 30 increased with 175 percent. Not only are such policy changes extremely rare, they are also unlikely to provide substantive guidance to current policy makers as most modifications to important policy parameters are much smaller in magnitude. For example, our RD estimates cannot be used to infer how labor supply and welfare use would change if benefits would have increased with 50 percent instead of 175 percent. On the other hand, with estimated preference parameters from a structural labor supply model that has been validated, we can predict employment and welfare changes resulting from any change in benefit levels.

To illustrate how employment, welfare use and hours of work change as we change social assistance benefits, we have calculated percentage changes in these measures as we increase benefits by $20 increments up to a total increase of $300. The results are presented in Figure 2. As can be seen, the responses to benefit changes are highly non-linear. For example, for modest increases ($60 or less per month), there are small increases in welfare use and small reductions in labor supply. However, an increase of $100 per month is predicted to increase welfare use by 2 percentage points as well as reduce employment by a similar magnitude.

6.4 Labor Supply and Welfare Responses to Wage Subsidies and Tax Credits

Another advantage of structural models over experimental studies is the ability of the former to perform counterfactual policy simulations. Hence, they serve as a major tool for policy makers in forming and developing future policies. In this paper, we consider two simulations: a wage subsidy for part-time workers and a tax credit to low-wage workers. Both simulations are designed to encourage nonworkers to enter the workforce. For the wage subsidy simulation, we double per hour wage rates for those working 1,000 hours per year or less. For the tax credit simulation, we use the 2010 parameters of the Canadian Working Income Tax Benefit (WITB) and adjust them to 1985 dollars.[25]

The results from the simulations, obtained using estimates from our preferred specification (four support unobserved heterogeneity distribution and fixed costs), are presented in Tables 6 (wage subsidy) and 7 (WITB). The entries in Table 6 suggest that the subsidy would increase employment rates (by 3.5 percentage points on average), reduce welfare participation rates (by 3.1 percentage points on average) and increase hours of work (on average with 11 hours per year). Most of the labor supply response will come from an increase at the extensive margin.

Table 6 further shows how the subsidy is predicted to affect individual differently depending on their educational attainment and their position in the pre-reform income distribution. As expected, the largest reactions to the wage subsidy are found among the high school drop-outs and among those in the first income quartile. For individuals in these groups, the subsidy is expected to substantially increase labor supply and reduce welfare use.

In Table 7 we present the outcomes from a simulation of the WITB. The employment and welfare reactions are much smaller than those obtained for the wage subsidy case. On average, employment rates would only in-

[25]The Canadian WITB is a refundable tax credit that provides tax relief for eligible working low-income individual. The structure of the program is similar to the EITC in the U.S. The maximum benefit in Quebec for a single individual with no dependents was $1,560.87 in 2010.

crease by 0.2 percentage points and welfare participation rates would drop by the same magnitude. The reactions are significantly smaller since the simulated tax credit program is much less generous than the wage subsidy program. Like the wage subsidy example, estimated reactions are largest among those with low incomes and low educational attainment. Even though this is not an evaluation of the 2010 WITB, the results suggest that the employment impacts of the tax credit are quite modest, at least among single males with no dependents.

6.5 Elasticities

A common and convenient way to illustrate the behavioral impact of different counterfactual policies is to calculate labor supply elasticities. In Table 8 we report earnings elasticities for different groups at both the extensive and intensive margins.[26] The entries in columns one and four were obtained by increasing the gross (or before tax) hourly wage with 10 percent for everyone and record the estimated changes in labor supply. Overall, we find evidence of small elasticities at both the extensive and intensive margins (0.031 and 0.106, respectively). The entries in Table 8 show a great deal of heterogeneity in these elasticities. For example, for respondents in the first quartile of the (net) income distribution, the elasticities are 0.314 and 0.048 at the extensive and intensive margins, respectively.

Small elasticities are not uncommon in the literature on male labor supply. For example, for the U.S., MaCurdy et al (1990) report wage elasticities between -0.24 and 0.03 for married men, 25-55 years old.[27] Further, Devereux (2003) using a sample of single and married men, report intensive wage elasticities between -0.022 and 0.017 and extensive elasticities ranging from -0.061 to 0.001. Finally, studies based on European data, summarized in Bargain, Orsini and Peichl (2011), indicate that wage elasticities for single

[26] The extensive elasticity was obtained by dividing the percentage change in net earnings as a result of the policy change by the percentage change in employment. The intensive elasticity was calculated for those predicted to work before the policy change by dividing the earnings change with the change in hours of work.

[27] As far as we know, there are no reported wage or income elasticities for single Canadian men.

men in Europe are small but positive, around 0.1 at both margins.

We have also calculated elasticities for the two counterfactual policy changes discussed above and they are also presented in Table 8. For the wage subsidy experiment, the average extensive elasticity equals 0.524 and at the intensive margin it equals -0.081. Again, there is evidence of substantial heterogeneity since only respondents in the lowest income quartile react to the policy change.The elasticities for the WITB program are presented in columns three and six. Despite small reactions in response to the change in tax parameters, the elasticities are larger for this policy change than those obtained for the wage subsidy experiment.

7 Conclusions

In this paper, we formulate and estimate a relatively simple economic model of labor supply and welfare participation. We then validate our model using a welfare reform that took place in the province of Quebec in 1989 and that implied a dramatic increase in welfare benefits for those under the age of 30. In particular, we compare our model predictions with those obtained using Regression Discontinuity methods. We explore different model specifications and document the importance of accounting for costs associated with work as well as unobserved heterogeneity in preferences. The results are encouraging. Our models predict employment reductions that are close to the Regression Discontinuity estimates.

We show the existence of heterogeneity in responses to the welfare reform. The changes in labor supply and welfare use are smaller for high school graduates than for high school drop-outs. Moreover, the effects are largest among those with lowest incomes. For those in the lowest quartile, the benefit increase is predicted to reduce employment by 18 percent and increase welfare use with 64 percent.

We also illustrate how employment, welfare use and hours of work change as we marginally increase social assistance benefits. The responses to these benefit changes are highly non-linear. For modest increases ($60 or less per month), there are small increases in welfare use and small reduc-

tions in labor supply. However, an increase of $100 per month is predicted to increase welfare use by 2 percentage points as well as reduce employment by a similar magnitude. Thus, results from experimental studies may only be of limited value when evaluating policy changes that differ from those used in the experiment.

Indeed, a major limitation of studies exploiting natural experiments is the inability of extrapolating the results to other policy changes. For example, our Regression Discontinuity estimates show responses to a welfare reform where benefits for those under the age of 30 increased with 175 percent. Not only are such policy changes extremely rare, they are also unlikely to provide substantive guidance to current policy makers as most modifications to important policy parameters are smaller in magnitude. However, with estimated preference parameters from a structural labor supply model, we can predict employment and welfare utilization changes resulting from any change in benefit levels.

Finally, it is important to note that we do not claim that our findings should be regarded as an endorsement of any structural model for any particular group of the population. Instead, we argue that it is important to validate structural models with estimates obtained using experimental methods. The model we estimate in this paper works quite well in predicting labor supply behavior of single men but it may not work as well for other groups, such as single mothers.[28] But once a model has passed a rigorous examination of its predictive power, it can serve as an important tool for both researchers and policy makers.

[28]For example, Choi (2012) uses data on single mothers to estimate a model that is similar to the one used in this paper. Although her model's predictions compare favorably with outcomes in the estimation sample, it is unable to replicate observed treatment effects from experiments in Minnesota and Vermont. However, her model ignores important aspects of female labor supply, such as child care costs. Further, she incorporates unobserved heterogeneity in a very restrictive way. Generalizing her model along these important dimensions is likely to improve the performance of her model.

References

Aaberge Rolf, Ugo Colombino, and Steinar Strom. 1999. Labor supply in Italy: An empirical analysis of joint household decisions with taxes and quantity constraints. *Journal of Applied Econometrics* 14:403-22.

Bargain, Olivier, Kristian Orsini, and Andreas Peichl. 2011. Labor supply elasticities in Europe and the U.S. Working Paper No. 5820, IZA, Bonn, Germany.

Blundell, Richard, Alan Duncan, Julian McCrae, and Costas Meghir. 2000. Evaluating in-work benefit reform: the working families tax credit in the U.K. Working Paper 160, Joint Center for Poverty Research, Northwestern University and University of Chicago, IL.

Blundell, Richard, and Andrew Shephard. 2012. Employment, hours of work and the optimal taxation of low-income families. *Review of Economic Studies* 79:481-510.

Cameron, Stephen V., and Christopher R. Taber. 1998. Evaluation and identification of semi-parametric maximum likelihood models of dynamic discrete choice," Working Paper, Department of Economics, Northwestern University.

Choi, Eleanor J. 2012. Evaluating a structural model of labor supply and welfare participation: Evidence from state welfare reform experiments. Unpublished manuscript, U.S. Bureau of Labor Statistics.

Devereux, Paul J. 2003. Changes in male labor supply and wages. *Industrial and Labor Relations Review* 56:409-28.

Dickens, William T., and Shelly J. Lundberg. 1993. Hours restrictions and labor supply. *International Economic Review* 34:169-92.

Flood, Lennart, Jorgen Hansen, and Roger Wahlberg. 2004. Household labor supply and welfare participation in Sweden. *Journal of Human Resources* 39:1008-32.

Flood, Lennart, and Nurul Islam. 2005. A Monte Carlo evaluation of discrete choice labour supply models. *Applied Economics Letters* 12:263-66.

Fortin, Bernard, Guy Lacroix, and Simon Drolet. 2004. Welfare benefits and the duration of welfare spells: Evidence from a natural experiment in Canada. *Journal of Public Economics* 88:1495-1520.

Friedlander, Daniel, and Philip K. Robins. 1995. Evaluating program evaluations: New evidence on commonly used nonexperimental methods. *American Economic Review* 85:923-37.

Gong, Xiaodong, and Arthur van Soest. 2002. Family structure and female labour supply in Mexico City. *Journal of Human Resources* 37:163-91.

Hansen, Jorgen, and Xingfei Liu. 2011. Estimating labor supply responses and welfare participation: Using a natural experiment to validate a structural labor supply model. Working Paper No. 5718, IZA, Bonn, Germany.

Heckman, James J., and Burton Singer. 1984. A method for minimizing the distributional assumptions in econometric models for duration data. *Econometrica* 52:271-320.

Heckman, James J., and V. Joseph Hotz. 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. *Journal of the American Statistical Association* 84:862-74.

Hoynes, Hilary Williamson. 1996. Welfare transfers in two-parent families: Labor supply and welfare participation under AFDC_UP. *Econometrica* 64:295-332.

Keane, Michael P., and Robert Moffitt. 1998. A structural model of multiple welfare program participation and labor supply. *International Economic Review* 39:553-89.

Keane, Michael P., and Kenneth I. Wolpin. 2007. Exploring the usefulness of a non-random holdout sample for model validation: Welfare effects on female behavior. *International Economic Review* 48:1351-78.

LaLonde, Robert J. 1986. Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review* 76:604-20.

Lemieux, Thomas, and Kevin Milligan. 2004. Incentive effects of social assistance: A regression discontinuity approach. Working Paper No. 10541, National Bureau of Economic Research, Cambridge, MA.

Lemieux, Thomas, and Kevin Milligan. 2008. Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics* 142:807-28.

Lise, Jeremy, Shannon Seitz, and Jeffrey Smith. 2005. Equilibrium policy experiments and the evaluation of social programs. Working Paper, Queen's University, Kingston, Canada.

MaCurdy, Thomas, David Green, and Harry Paarsch. 1990. Assessing empirical approaches for analyzing taxes and labor supply. *Journal of Human Resources* 25:415-90.

McFadden, Daniel L. 1977. Demand model estimation and validation. Urban Travel Demand Forecasting Project, Phase 1, Final Report Series, Volume 5, Institute of Transportation Studies, University of California, Berkeley.

Milligan, Kevin. 2008. Canadian tax and credit simulator: Database, software and documentation, version 2008-1, University of British Columbia, Canada.

Moffitt, Robert. 1983. An economic model of welfare stigma. *American Economic Review* 73:1023-35.

Todd, Petra E., and Kenneth I. Wolpin. 2006. Assessing the impact of a school subsidy program in Mexico using experimental data to validate a dynamic behavioral model of child schooling. *American Economic Review* 96:1384-1417.

van Soest, Arthur. 1995. Structural models of family labor supply: A discrete choice approach. *Journal of Human Resources* 30:63-88.

van Soest, Arthur, and Marcel Das. 2001. Family labor supply and proposed tax reforms in the Netherlands. *De Economist* 149:191-218.

Table 1.

Employment and Social Assistance Rates

in the 1986 Canadian Census

	Employment Rate Last Year	Employment Rate at Census Date	Social Assistance Use
Full Sample	0.785	0.815	0.153
High School Drop-outs	0.607	0.638	0.252
High School Graduates	0.867	0.897	0.107

Table 2.

Sample Averages by Welfare Status

Variable	No Welfare	(st dev)	Welfare	(st dev)
Age	29.6	(5.2)	31.5	(5.4)
High School Graduate	0.72	(0.45)	0.48	(0.50)
Immigrant	0.07	(0.25)	0.04	(0.21)
Hours of Work	1,676	(787)	254	(636)
Wage/hour	12.1	(5.4)	11.2	(2.2)
Employment Rate	0.90	(0.3)	0.16	(0.37)
Number of Observations	876		158	

Table 3.
Estimated Reactions to the Welfare Reform (standard errors in parantheses).

Model	Model Fit	Change in Employment Rate	Change in Social Assistance	Change in Annual Hours of Work	AIC
No Unobserved Heterogeneity					
No Fixed Cost	287.8	-0.001 (0.002)	0.002 (0.003)	-0.003 (0.082)	4723
Fixed Costs	74.5	-0.090 (0.013)	0.100 (0.014)	-0.001 (0.290)	4007
With Unobserved Heterogeneity					
Two Supports	25.1	-0.075 (0.011)	0.072 (0.007)	-0.076 (0.012)	3881
Three Supports	16.4	-0.051 (0.019)	0.047 (0.018)	-0.054 (0.019)	3866
Four Supports	9.3	-0.042 (0.011)	0.039 (0.011)	-0.038 (0.010)	3833
Five Supports	61.1	-0.039 (0.025)	0.038 (0.026)	-0.044 (0.048)	3888
Regression Discontinuity	N.A.	-0.041 (0.073)	0.013 (0.061)	-0.026 (0.151)	N.A.

Table 4.
Estimated Reactions to the Welfare Reform Among Different Groups

Group	Change in Employment Rate	Change in Social Assistance	Change in Annual Hours of Work
High school drop-outs	-0.050 (0.014)	0.045 (0.013)	-0.044 (0.011)
High school graduates	-0.039 (0.011)	0.036 (0.011)	-0.036 (0.011)
Natives	-0.043 (0.012)	0.040 (0.011)	-0.039 (0.011)
Immigrants	-0.026 (0.010)	0.026 (0.013)	-0.023 (0.011)
Position in the net income distribution:			
First Quartile	-0.076 (0.021)	0.072 (0.022)	-0.062 (0.017)
Second Quartile	-0.081 (0.024)	0.073 (0.020)	-0.081 (0.023)
Third Quartile	-0.008 (0.006)	0.008 (0.006)	-0.008 (0.006)
Fourth Quartile	-0.003 (0.003)	0.003 (0.003)	-0.003 (0.004)

The reactions were obtained using estimates from the preferred specification (unobserved heterogeneity (four supports) and fixed costs). Standard errors appear in parantheses.

Table 5.

Comparisons of Observed and Predicted Outcomes

	1985		1991	
Hours	Data	Model	Data	Model
0	21.5	21.5	16.1	19.0
$0 < hours <= 500$	1.0	0.7	1.7	1.0
$500 < hours <= 1,000$	8.6	8.2	7.5	7.3
$1,000 < hours <= 1,500$	7.5	7.7	8.9	7.3
$1,500 < hours <= 2,000$	23.9	21.6	25.4	23.6
$2,000 < hours <= 2,500$	31.0	31.7	32.1	37.1
$2,500 < hours$	6.6	8.6	8.3	4.7

Note: The first two columns show the within sample fit (we estimate the model on 1985 data) while the last two columns show out-of-sample predictions (obtained using 1991 data).

Table 6.
Estimated Reactions - Wage Subsidy

Group	Change in Employment Rate	Change in Social Assistance	Change in Annual Hours of Work
Full sample	0.035 (0.008)	-0.031 (0.008)	0.011 (0.005)
High school drop-outs	0.053 (0.014)	-0.047 (0.013)	0.027 (0.010)
High school graduates	0.027 (0.006)	-0.025 (0.006)	0.003 (0.004)
Natives	0.035 (0.008)	-0.031 (0.008)	0.011 (0.005)
Immigrants	0.038 (0.017)	-0.037 (0.017)	0.010 (0.015)
Position in the net income distribution:			
First Quartile	0.141 (0.034)	-0.125 (0.032)	0.081 (0.022)
Second Quartile	0	0	-0.026 (0.010)
Third Quartile of the net	0	0	-0.007 (0.005)
Fourth Quartile of the net	0	0	-0.005 (0.005)

The reactions were obtained using estimates from the preferred specification (unobserved heterogeneity (four supports) and fixed costs).

Table 7.
Estimated Reactions - Working Income Tax Benefit

Group	Change in Employment Rate	Change in Social Assistance	Change in Annual Hours of Work
Full sample	0.002 (0.001)	-0.002 (0.001)	0.0002 (0.001)
High school drop-outs	0.003 (0.003)	-0.003 (0.003)	0.001 (0.002)
High school graduates	0.001 (0.001)	-0.001 (0.001)	-0.0002 0.001)
Natives	0.002 (0.001)	-0.002 (0.001)	0.0002 (0.001)
Immigrants	0.002 (0.002)	-0.002 (0.002)	0.0002 (0.003)
Position in the net income distribution:			
First Quartile	0.007 (0.005)	-0.007 (0.005)	0.004 (0.003)
Second Quartile	0	-0.00002 (0.002)	-0.003 (0.002)
Third Quartile of the net	0	0	-0.0001 (0.001)
Fourth Quartile of the net	0	0	0

The reactions were obtained using estimates from the preferred specification (unobserved heterogeneity (four supports) and fixed costs).

Table 8.
Elasticities

	Extensive margin			Intensive margin		
Group	Overall wage increase	Wage subsidy	WITB	Overall wage increase	Wage subsidy	WITB
All	0.031 (0.086)	0.524 (0.083)	2.815 (0.512)	0.106 (0.025)	-0.081 (0.085)	0.106 (0.060)
Position in the net income distribution:						
First Quartile	0.314 (1.549)	1.053 (0.145)	2.906 (0.549)	0.048 (0.022)	-0.118 (0.072)	0.099 (0.056)
Second Quartile of the net	0.009 (0.014)	0	0	0.175 (0.052)	-0.061 (0.306)	1.744 (0.034)
Third Quartile of the net	0.022 (0.021)	0	0	0.082 (0.034)	0.119 (0.519)	1.594 (0.027)
Fourth Quartile of the net	0.005 (0.012)	0	0	0.117 (0.039)	0.416 (0.773)	0

Note: The results for the overall wage increase were obtained by increasing everyones wage with 10 percent.

Appendix:

Table A1.

Estimated Parameters of the Preferred Model (specification with unobserved heterogeneity (four supports) and fixed costs of work)

Parameters	Estimates	Std Err	T-Stat
Preference for Leisure			
$\beta_L(age)$	-0.06	0.06	-1.00
$\beta_L(highschool)$	-3.15	0.77	-4.09
$\beta_L(immigrant)$	2.40	1.41	1.70
θ_{L1}	13.21	4.05	3.26
θ_{L2}	19.45	4.70	4.14
θ_{L3}	59.33	6.96	8.52
θ_{L4}	-0.26	0.56	-0.45
Preference for Consumption			
$\beta_{NI}(age)$	0.02	0.01	2.00
$\beta_{NI}(highschool)$	0.09	0.08	1.05
$\beta_{NI}(immigrant)$	0.09	0.12	0.75
θ_{NI1}	-17.54	2.26	-7.76
θ_{NI2}	-15.25	2.36	-6.47
θ_{NI3}	-5.69	1.73	-3.29
θ_{NI4}	-8.90	5.84	-1.52

Table A1.

Continued

Parameters	Estimates	Std Err	T-Stat
Preference for Welfare			
$\beta_{SA}(age)$	-0.003	0.03	-0.10
$\beta_{SA}(highschool)$	0.50	0.34	1.47
$\beta_{SA}(immigrant)$	1.13	0.71	1.59
θ_{SA1}	1.06	0.94	1.13
θ_{SA2}	16.97	2.04	8.32
θ_{SA3}	2.01	0.87	2.31
θ_{SA4}	14.39	1.63	8.83
Other Utility Parameters			
β_{NIsq}	-0.53	0.11	-4.82
β_{Lsq}	-30.73	3.31	-9.28
$\beta_{NI,L}$	3.74	0.59	6.34
FC	19.87	5.27	3.77
Other			
$p1$	-0.56	0.13	-4.31
$p2$	0.03	0.14	0.21
$p3$	-0.05	0.10	-0.50

Figure 1. Federal and Provincial Marginal Income Tax Rates 1985

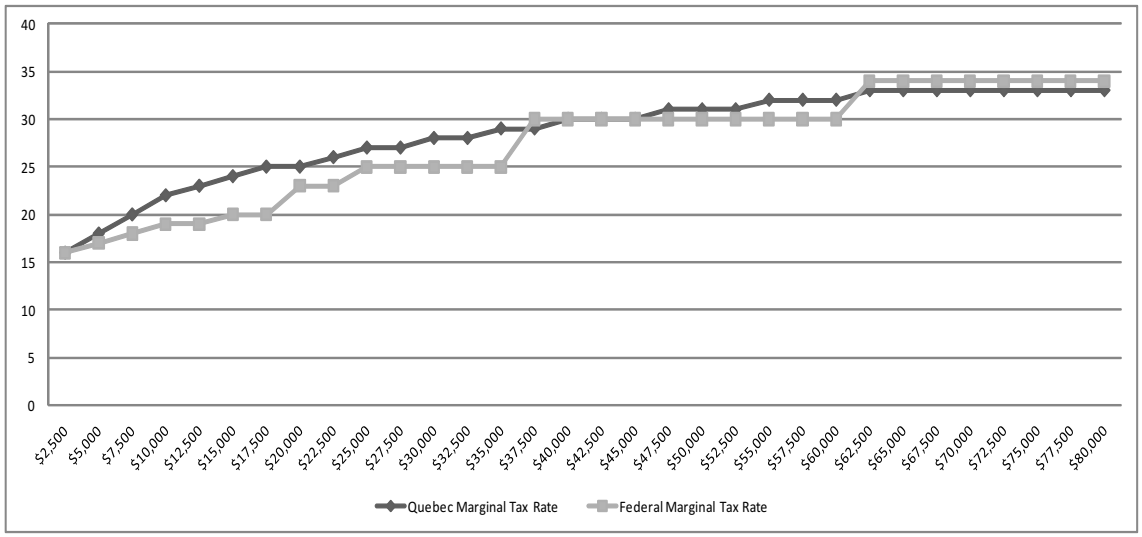

Figure 2.Estimated Reactions to Different Welfare Benefit Levels