

Hansen, Jorgen; Liu, Xingfei

Working Paper

Estimating labor supply responses and welfare participation: Using a natural experiment to validate a structural labor supply model

IZA Discussion Papers, No. 5718

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Hansen, Jorgen; Liu, Xingfei (2011) : Estimating labor supply responses and welfare participation: Using a natural experiment to validate a structural labor supply model, IZA Discussion Papers, No. 5718, Institute for the Study of Labor (IZA), Bonn, <https://nbn-resolving.de/urn:nbn:de:101:1-201106013313>

This Version is available at:

<https://hdl.handle.net/10419/51807>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 5718

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

Jorgen Hansen
Xingfei Liu

May 2011

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

Jorgen Hansen

*Concordia University,
CIRANO, CIREQ and IZA*

Xingfei Liu

Concordia University

Discussion Paper No. 5718

May 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

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

In this paper, we formulate and estimate an economic model of labor supply and welfare participation. The model is estimated on data on single men from Quebec drawn from the 1986 Canadian Census. Budget sets for each work-welfare combination – accounting for income taxes, tax credits and welfare benefit rules – are derived using a microsimulation model. We validate our model by comparing reactions to a welfare reform that implied a dramatic increase in welfare benefits predicted by our model to those obtained by using a regression discontinuity approach. The results show that our model is capable of recovering actual changes in labor supply and welfare participation. We also show the advantage of having estimated a structural model by illustrating how labor supply and welfare participation change when benefit levels change.

JEL Classification: J22

Keywords: labor supply, welfare participation, unobserved heterogeneity,
natural experiment, regression discontinuity, microsimulation

Corresponding author:

Jorgen Hansen
Department of Economics
Concordia University
1455 de Maisonneuve Blvd W
Montreal, Quebec, H3G 1M8
Canada
E-mail: Jorgen.Hansen@concordia.ca

* We thank seminar participants at Goteborg University and IZA for helpful comments and Kevin Milligan for providing detailed information on tax and benefit parameters.

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 the large changes. Indeed, as mentioned in Keane and Wolpin (2006), 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. Wise (1985) exploited a randomized social experiment involving housing subsidies to evaluate a model of housing demand. Lise et al (2003) used data from the Self-Sufficiency Project, a Canadian large-scale experiment designed to move people off of 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 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. Finally, Keane and Wolpin (2006) 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.

In this paper we exploit 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.¹ This form of age discrimination was terminated in 1989 when 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 et al, 2006) 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 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

¹Recipients were required to complete a form each month allowing officials to determine if respondents had reached age 30.

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 18 percent and increase welfare participation with 64 percent. Finally, we 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 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.

The remainder of this paper is organized as follows. Section 2 provides a description of the change in welfare benefits that took place in Quebec in 1989 as well as a brief description of the main features of the Quebec income tax system. 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 is 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 al-

lowed 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.² 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". There are many assistance programs available for Quebec residents. In order to be eligible for these programs, individuals will have to meet certain criterions. For social assistance, eligibility requires Quebec residency and that the applicant is 18 years or older, in addition to passing the basic needs test.

Regarding income taxes, the Quebec system was - and still is - composed of two parts, a provincial component and a federal component. In

²Although 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.

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. 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), and Flood et al (2004). Unlike the continuous labor supply model, the discrete choice model allows us to include as many details as possible regarding the budget set. Specifically, we assume that each individual can choose among all the alternatives in the choice set of net income (NI)-leisure (L) combinations (NI_j, L_j) , where $j = 1, \dots, J$. Further, $L_j = TE - h_j$ where TE denotes total time endowment and is set to 4,000 hours per year.³ Thus, the choice set for an individual contains J

³ TE can also be regarded as a parameter that can be estimated together with all other parameters.

different hours of work combinations. In the empirical part of the paper, we set $J = 7$.⁴

We assume that utility depends not only on income and leisure, but also on participation in welfare programs. We further assume that the utility function is increasing in income 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.⁵

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

$$U(NI, L) = \beta_{NI} \log(NI) + \beta_L \log(L) + \beta_{NI^2} (\log(NI))^2 + \beta_{L^2} (\log(L))^2 + 2\beta_{NI,L} \log(NI) \log(L) - \phi_{SA} d_{SA} \quad (1)$$

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

The individual chooses leisure time (L), social assistance status (d_{SA}) and consumption (or net income) by maximizing utility subject to the following budget constraint:

$$NI_j = Wh_j + Y - t(Wh_j + Y^T - D) + B_{SA}(NI_j) d_{SA} \quad (2)$$

where NI_j is the income net of taxes at hours combination j (h_j), W equals the before-tax hourly wage rate, Y denotes annual non-labor income, $t(\cdot)$ is a function that determines income taxes, Y^T is taxable non-labor income, and D represents deductions. Lastly, $B_{SA}(\cdot)$ is a function that

⁴We set $h_1 = 0, h_2 = 500, h_3 = 1,000, h_4 = 1,500, h_5 = 2,000, h_6 = 2,500$ and $h_7 = 3,000$.

⁵What 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.

determines social assistance benefits.

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 in (1) 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:

$$U_{j,r} = U(NI_j, L_j) + \varepsilon_{j,r} \quad (3)$$

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 $Pr(\varepsilon_{j,r} < \epsilon) = \exp(-\exp(-\epsilon))$. The 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

$$l = \sum_{r=1}^2 \sum_{j=1}^J p_{j,r} \delta_{j,r} \quad (4)$$

where

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

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

$$\begin{aligned}\beta_L &= \sum_{k=1}^K \beta_{L,k} x_k + \theta_L \\ \beta_{NI} &= \sum_{k=1}^K \beta_{NI,k} x_k + \theta_{NI} \\ \phi_{SA} &= \sum_{k=1}^K \beta_{SA,k} x_k + \theta_{SA}\end{aligned}\tag{5}$$

where the elements of the vector x are observed individual characteristics (age, education and immigrant status). K denotes the dimension of the vector x , while the θ '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 $(\theta_L, \theta_{NI}, \theta_{SA})$ that determine an individual's preferences, each observed with probability π_m (where $\pi_m > 0$ and $\sum \pi_m = 1$, $m = 1, \dots, M$). 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 util-

ity function, the contribution to the likelihood function for a given individual is

$$l = \sum_{m=1}^M \pi_m \left\{ \sum_{r=1}^2 \sum_{j=1}^J (p|\Theta)_{j,r} \right\} \delta_{j,r} \quad (6)$$

where

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

and where $\Theta = \{\theta_L, \theta_{NI}, \theta_{SA}\}$. 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 et al (1995, 1999) and Flood et al (2004)).

Aaberge et al (1995, 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 $\log(NI)$ for employed individuals with $\log(NI) - \log(FC)$ 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.⁶

⁶In addition to adding heterogeneity in preferences and fixed costs of work, we have

5 Data

5.1 Description of the Data and Sampling Procedures

The data employed in this study are drawn from the 1986 Canadian Census.⁷ The data are extracted from a sample of 500,434 individuals who accounted for approximately 2% of the Canadian population at the time.⁸ The data contain detailed 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 males without children residing in the province of Quebec in 1985. Further, 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.⁹ 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. Furthermore, since we are

also explored the possibility that wages are measured with errors and that wages may be endogenous. In both cases, we find no evidence suggesting that incorporating these issues into our model improve its performance. This finding is consistent with the results in van Soest (1995).

⁷Specifically, we utilized the public use micro-data file on individuals from the 1986 Census of Canada.

⁸The target population for the 1986 Census includes all individuals except for institutional residents and residents of incompletely enumerated Indian reserves or settlements.

⁹Including married individuals would necessitate a different economic model that accounts for joint household decisions.

also modeling individuals' welfare decisions, we need information on welfare participation as well. 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. These calculations are based on a small micro-simulation model that we constructed using details from the 1985 Quebec and Federal Income Tax Schedules.¹⁰

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). It 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. We excluded males who worked less than 13 weeks and those who worked less than 4 hours during the reference week. We further 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.

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,

¹⁰We utilized information on tax and benefit parameters kindly provided by Kevin Milligan.

young males, we believe that this measure provides a reasonable measure of social assistance receipt. As mentioned in Lemieux & Milligan (2008), "... social assistance benefits are by far the largest component of the "other transfers" variable".¹¹

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, 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 grad-

¹¹Lemieux and Milligan (2004) show that 85 percent of income included in "other transfers" in the Census is social assistance payments.

uates). 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.¹² 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 of the introduction of higher welfare benefits for those under the age of 30. The results are presented in Table 3.¹³ 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 Regression Discontinuity methods. Finally, we use the Akaike information criteria 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

¹²This 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.

¹³Complete estimation results are available upon request.

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 employment 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.¹⁴ 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 partic-

¹⁴Given the inferior performance of the model without fixed costs above, we do not consider that model specification further.

ipation are close to those obtained using RD. The employment rate is predicted to decrease with 3.8 percentage points (compared to 4.1 using RD) while social assistance is predicted to increase with 3.6 percentage points (compared to 1.3 using RD). Further, annual hours of work are predicted to drop with 3.4 percentage points using our model and with 2.6 percentage points using RD.

The effects on both the employment and welfare participation rates are similar to those reported in Lemieux and Milligan (2008). Their RD estimates are 4.9 and 4.1 percentage points for employment and welfare, respectively. Hence, this model specification performs quite well, suggesting that it may be considered for evaluation of a range of benefit changes. The standard errors for both our RD estimates as well as for our model effects are however relatively large. The limited sample size is a likely reason for these inflated standard errors.

Finally, the AIC for the model with four support points is 3972. 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.¹⁵

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 closely resembles the true effects, as estimated by RD on our sample and by Lemieux and Milligan (2008). In addition to this external validation, they also generate hours distributions that are similar to the observed distribution. 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.

¹⁵Cameron and Taber (1998) also use AIC to estimate the number of support points in their Monte Carlo study of dynamic, discrete choice models.

6.2 Heterogeneity in Responses to the Welfare Reform

As noted above in Section 5.1, our sample includes both high school drop-outs 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 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 4.9 percentage points (or 7.7 percent). This figure is identical to that reported by Lemieux and Milligan (2008), providing additional support for the validity of our structural model. The predicted increase in social assistance use for high school drop-outs is 4.3 percentage points, which compares very favorably to the 4.1 point increase reported in Lemieux and Milligan (2008). Finally, the reform is predicted to reduce annual hours of work for this group by 43 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 two thirds of that for high school drop-outs. It is also noteworthy that natives respond more to benefit changes than immigrants.

Finally, Table 4 shows the reform's impact on males located at different places in the distribution of net incomes. As expected, the effects are largest among those with lowest incomes (those in the first quartile). 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, beyond those obtained using Regression Discontinuity 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 how the reform will impact behavior among 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.

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 also illustrate 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.

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.

We 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 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. Thus, results from experimental studies may only be of limited value when evaluating policy changes that differ from those used in the experiment.

References

- [1] Aaberge R., J. Dagsik and S. 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.
- [2] Blundell, R., A. Duncan, J. McCrae and C. Meghir, 2000, "Evaluating In-Work Benefit Reform: the Working Families Tax Credit in the UK," London: Institute for Fiscal Studies, working paper.
- [3] Cameron, S.V. and C.R. Taber, 1998, "Evaluation and Identification of Semi-parametric Maximum Likelihood Models of Dynamic Discrete Choice," Working paper, Department of Economics, Northwestern University.
- [4] Dickens, W. and S. Lundberg, 1985, "Hours Restrictions and Labor Supply," *International Economic Review* 34, 169-92.
- [5] Flood, L., J. Hansen and R. Wahlberg, 2004, Household Labor Supply and Welfare Participation in Sweden," *Journal of Human Resources* 39, 1008-32.
- [6] Fortin, B., G. Lacroix and S. Drolet, 2006, "Welfare Benefits and the Duration of Welfare Spells: Evidence from a Natural Experiment in Canada," *Journal of Public Economics* 88, 1495-1520.

- [7] Heckman J. and B. Singer, 1984, "A Method for Minimizing the Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52, 271-320.
- [8] Hoynes, H., 1996, "Welfare Transfers in Two-parent Families: Labor Supply and Welfare Participation under AFDC_UP," *Econometrica* 64, 295-332.
- [9] Keane, M. and K. 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.
- [10] Lemieux, T. and K. Milligan, 2004, "Incentive Effects of Social Assistance: A Regression Discontinuity Approach," National Bureau of Economic Research, Working Paper No. 10541.
- [11] Lemieux, T. and K. Milligan, 2008, "Incentive Effects of Social Assistance: A Regression Discontinuity Approach," *Journal of Econometrics* 142, 807-28.
- [12] Lise, J., S. Seitz and J. Smith, 2005, "Equilibrium Policy Experiments and the Evaluation of Social Programs," Working Paper, Queen's University, Kingston, Canada.
- [13] MaCurdy, T., D. Green and H. Paarsch, 1990, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *Journal of Human Resources* 25, 415-90.
- [14] McFadden, D., 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.
- [15] Milligan, K., 2008, Canadian Tax and Credit Simulator. Database, software and documentation, version 2008-1.
- [16] Moffitt, R., 1983, "An Economic Model of Welfare Stigma," *American Economic Review* 73, 1023-35.

- [17] Todd, P. and K. 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.
- [18] van Soest, A., 1995, "Structural Models of Family Labor Supply: A Discrete Choice Approach," *Journal of Human Resources* 30, 63-88.
- [19] van Soest, A. and M. 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 paranthesis).

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	-0.0004 (0.037)	0.003 (0.062)	-0.003 (0.082)	4723
Fixed Costs	74.5	-0.100 (0.300)	0.110 (0.312)	-0.001 (0.290)	4007
With Unobserved Heterogeneity					
Two Supports	25.1	-0.077 (0.267)	0.074 (0.262)	-0.076 (0.290)	3881
Three Supports	16.4	-0.051 (0.221)	0.047 (0.212)	-0.054 (0.253)	3866
Four Supports	9.3	-0.038 (0.191)	0.036 (0.186)	-0.034 (0.183)	3833
Five Supports	61.1	-0.037 (0.188)	0.035 (0.183)	-0.044 (0.237)	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	-0.049 (-7.7%)	0.043 (17.3%)	-0.043 (-3.7%)
High school graduates	-0.033 (-3.8%)	0.033 (28.9%)	-0.030 (-1.8%)
Natives	-0.039 (-5.0%)	0.037 (23.8%)	-0.036 (-2.4%)
Immigrants	-0.015 (-2.0%)	0.015 (14.3%)	-0.012 (-0.8%)
Respondents in the:			
First Quartile of the net income distribution	-0.143 (-18.0%)	0.138 (63.6%)	-0.134 (-10.9%)
Second Quartile of the net income distribution	-0.014 (-1.5%)	0.019 (36.4%)	-0.008 (-0.4%)
Third Quartile of the net income distribution	-0.005 (-0.5%)	0.005 (50.0%)	-0.004 (-0.2%)
Fourth Quartile of the net income distribution	0	0	0

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

Table 5.

In-sample Model Fit

Hours	Data	Model
0	21.5	21.5
$0 < \text{hours} \leq 500$	1.0	0.7
$500 < \text{hours} \leq 1,000$	8.6	8.2
$1,000 < \text{hours} \leq 1,500$	7.5	7.7
$1,500 < \text{hours} \leq 2,000$	23.9	21.6
$2,000 < \text{hours} \leq 2,500$	31.0	31.7
$2,500 < \text{hours}$	6.6	8.6

Table 6.

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	0.08	-0.83
$\beta_L(highschool)$	-3.15	0.83	-3.78
$\beta_L(immigrant)$	2.40	1.61	1.49
θ_{L1}	13.21	10.23	1.29
θ_{L2}	19.45	7.64	2.55
θ_{L3}	59.33	6.48	9.15
θ_{L4}	-0.26	17.10	-0.01
Preference for Consumption			
$\beta_{NI}(age)$	0.02	0.02	1.24
$\beta_{NI}(highschool)$	0.09	0.11	0.81
$\beta_{NI}(immigrant)$	0.09	0.15	0.61
θ_{NI1}	-17.54	8.47	-2.07
θ_{NI2}	-15.25	8.00	-1.91
θ_{NI3}	-5.69	4.28	-1.33
θ_{NI4}	-8.90	25.33	-0.35

Table 6.**Continued**

Parameters	Estimates	Std Err	T-Stat
Preference for Welfare			
$\beta_{SA}(age)$	-0.003	0.05	-0.06
$\beta_{SA}(highschool)$	0.50	0.40	1.25
$\beta_{SA}(immigrant)$	1.13	0.82	1.38
θ_{SA1}	1.06	1.56	0.68
θ_{SA2}	16.86	3.78	4.46
θ_{SA3}	2.01	1.41	1.43
θ_{SA4}	14.47	3.72	3.88
Other Utility Parameters			
$\beta_{NI_{sq}}$	-0.53	0.29	-1.84
$\beta_{L_{sq}}$	-30.73	5.49	-5.60
$\beta_{NI,L}$	3.74	2.00	1.87
FC	19.87	6.18	3.21
Other			
$p1$	-0.56	0.21	-2.72
$p2$	0.03	0.34	0.08
$p3$	-0.05	0.24	-0.21

**Figure 1. Federal and Provincial(Quebec) Marginal Income Tax Rates
1985**

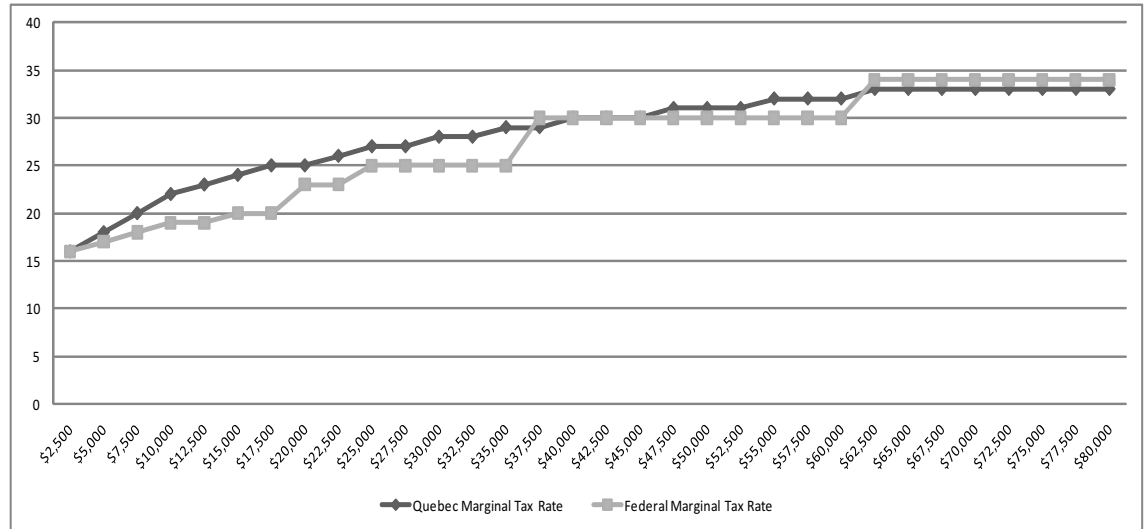


Figure 2. Estimated Reactions to Different Welfare Benefit Levels

