Essays on Structural Labour Supply and Government Policies

Arezoo Banihashem

A Thesis

in

The Department

of

Economics

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Economics) at

Concordia University

Montréal, Québec, Canada

September 2024

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Miss Arezoo Banihashem**

Entitled: **Essays on Structural Labour Supply and Government Policies**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________ Chair
Dr. Huan Xie

______________________________ External Examiner
Dr. Nizamul Islam

______________________________ Examiner
Dr. Damba Lkhagvasuren

______________________________ Examiner
Dr. Gabrielle Vasey

______________________________ Examiner
Dr. Jan Victor Dee

______________________________ Supervisor
Dr. Jörgen Hansen

Approved by ______________________________
Dr Jörgen Hansen, Chair
Department of Economics

__________ 2024 ______________________________
Dr Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Essays on Structural Labour Supply and Government Policies

Arezoo Banihashem, Ph.D.

Concordia University, 2025

This thesis uses a structural modeling approach to assess labour supply and evaluate policy programs. The first chapter compares labour market outcomes for high school dropouts to graduates in Quebec, Ontario, Alberta, and British Columbia. Results show that dropouts face worse outcomes across all provinces, with Quebec having a significantly higher proportion of male dropouts. Simulations aimed at boosting employment incentives for low-skilled individuals emphaise the importance of long-term strategies that enhance skill acquisition and reduce financial barriers. Current welfare eligibility criteria offer limited incentives to transition from welfare to work at modest wages.

The second chapter focuses on modeling individual heterogeneity, particularly unobserved characteristics, using random coefficients. It uses Monte Carlo simulations across six scenarios with varying shapes and variances for the distribution of unobserved characteristics. Findings reveal that methods accounting for heterogeneity perform well when variances are small, but become sensitive to distribution shapes as variances increase, indicating the need for more flexible models in high-variance contexts.

The final chapter examines the labour supply of single mothers, with a focus on childcare utilisation and social assistance participation. Contrary to traditional views, the study finds that childcare costs are no longer a significant barrier to employment, with access to childcare being a more critical issue. Policies targeting direct employment incentives may be more effective in increasing labour force participation. The chapter also highlights the role of unobserved preferences in shaping work decisions, suggesting that current programs may be limited by not fully addressing these behavioural factors.

Acknowledgments

I am deeply grateful to my supervisor, Professor Jörgen Hansen, for his guidance, patience, and unwavering support. His friendly and approachable attitude made it easy to reach out whenever I needed help. Beyond his research expertise, he has been an invaluable mentor in my professional life, offering advice and support that I truly appreciate.

I would also like to thank the members of my thesis committee—Dr. Nizamul Islam, Professor Damba Lkhagvasuren, Professor Gabrielle Vasey, Professor Jan Victor Dee and Professor Huan Xie—for their helpful comments and feedback, as well as their approachable and supportive attitudes throughout this process.

The analysis presented in Chapter One and Chapter Three was conducted at the Québec Interuniversity Centre for Social Statistics (QICSS) which is part of the Canadian Research Data Centre Network (CRDCN). Special thanks to Danielle Forest at the McGill-Concordia branch for her patience and support in vetting requests. The views expressed in these chapters are those of the author(s) and do not necessarily reflect those of the CRDCN, the QICSS, or their partners.

A huge thank you to my mom, dad, and brother for believing in me when I doubted myself and for offering unconditional support in ways only family can.

To The Guy, thank you for your endless patience, encouragement, and support, along with all the fun and laughter. I'm so happy to finally celebrate this milestone with you!

And of course, a heartfelt thanks to the Hapolonians, Tooties, Michelle, Corduroy, Moosie, Torti, Morti, Hapal and Piposh for being a constant source of fun and insanity. You've been the best kind of annoying—always entertaining, in your uniquely supportive ways! I'm sure you'll keep playing that role!

Finally, I would like to thank my friends for their support and encouragement.

The authors of the paper in Chapter One, are very grateful to Raphael Jananji for excellent research assistance on this chapter. The services and activities provided by the QICSS in this chapter are made possible by the financial or in-kind support of the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, the Fonds de recherche du Québec and the Quebec universities.

Contribution of Authors

Chapter one was coauthored by Dr Jörgen Hansen and Dr Christian Belzil. I was responsible for collecting and compiling the data and implementing the estimation procedures. This was all done at a Statistics Canada data centre. I also researched the existing, relevant literature and drafted the section on literature review. This work was completed under the supervision and guidance of Dr Hansen. Dr Belzil contributed to the modelling development and derivation of the estimation procedure. He also assisted in preparing the final manuscript. Dr Hansen oversaw all aspects of the project, developed the model and estimation methods, discussed data collection and variable definitions with me, and was the principal author of the manuscript.

Contents

List of Figures

List of Tables

Chapter 1

High School Dropouts in Québec: Outcomes and Policy Simulations

1.1 Introduction

For decades, the high school dropout rate in Quebec has been among the highest in Canada and it has been consistently higher than the one in Ontario.[1] In 2018, the proportion of men aged 25 to 34 with less than high school was 11 percent in Quebec - the highest of all provinces, see Statistics Canada (2019). By comparison, the corresponding proportion for Ontario was 6 percent. Further, the dropout rate in Quebec was significantly lower for females (6 percent). The persistently high incidence of dropping out of high school has attracted attention of policy makers as well as academics and other members of the Quebec society. For example, the Action Group on Student Retention and Success in Quebec was formed in 2008 to examine the issue and they produced a report in 2009 that both documents the situation and provides an action plan for improving student retention and high school graduation.

In the paper, and consistent with evidence elsewhere, high school dropouts experience lower wages, fewer employment opportunities, poorer health outcomes and higher incidences of criminal activities and welfare use than those with more education.[2] This adds up to a significant cost to

[1] Throughout this paper, we will use the term high school dropout for persons that have not obtained a high school diploma or a *Diplôme D.E.S* in Québec.

[2] We will use the terms welfare and social assistance interchangeably in this paper.

society through loss of tax revenues and increased expenses on government transfers. The report estimates the cost to be as high as $120,000 per person in Quebec. Results reported in Lemieux and Milligan (2008) suggest that this cost may be higher in Quebec than in other provinces as the employment rates for dropouts tend to be lower and welfare participation rates higher in this province.

In this paper, we will focus on labour market outcomes for those without a high school diploma, recognising that there may also be other aspects associated with lower educational attainment that negatively impact individuals over their life-cycle, such as health and criminal activities. In particular, we will utilize recent data on young male high school dropouts in Quebec, collected in the 2016 Canadian Census, and provide an in-depth analysis and assessment of different policy options, designed to improve and stimulate labour market attachments for this group.[3]

Using these data and when comparing outcomes across four different provinces (Quebec, Ontario, Alberta and British Columbia), we find that the average hourly wage among high school dropouts (expressed in 2015 dollars) was lowest in Quebec (at $20) and highest in Alberta (at $29). In Ontario and British Columbia, the average wage was $22. However, when comparing average wages for high school dropouts relative to high school graduates, dropouts from Quebec are doing better than dropouts in other provinces because the average wage of high school graduates is also lowest in Quebec.

In terms of employment rates, dropouts in Quebec and Ontario fare worse than those in Alberta and British Columbia, and the same is true for welfare participation. When comparing these measures for high school dropouts relative to high school graduates, dropouts from Quebec have lower relative employment rates than dropouts from Ontario, Alberta and British Columbia. In terms of welfare, participation rates are substantially higher among dropouts than among high school graduates in all provinces (the participation rates are 2.5 times higher for dropouts in Quebec, Ontario and Alberta and even higher in British Columbia). Overall, in all measurable outcomes (wages, employment, and welfare participation) and for all four provinces, high school dropouts are substantially disadvantaged when compared to high school graduates.

[3]We limit attention to male individuals without a high school diploma primarily because the dropout rate is substantially higher for males than for females. Further, the labour market activities of young males are easier to analyse since we can ignore fertility decisions.

It is important to analyse the weak performance of high school dropouts in the labour market and the role for government policies. While each province aims to improve living standards of low-skilled workers and limit the incidence of poverty, they must also ensure that the measures do not eliminate work incentives. This is a difficult balancing act and different provinces have adopted different approaches. Estimating and evaluating the responsiveness of this low-skill group to a variety of current and potential welfare policies can aid policymakers in making better-informed, evidence-based, decisions when trying to influence the labour supply and welfare participation decisions of this group.

This issue may be more important than ever before as technological improvements in the labour market are expected to dramatically change the landscape for low-skilled workers in ways never seen before. It is therefore critical to investigate and analyse alternative policy measures designed to improve outcomes for this group. In this paper, we will develop and estimate a simple economic model of labour supply and welfare take-up and use it to conduct a series of counterfactual policy experiments. A similar model was used in Hansen and Liu (2015) who showed that such a model was able to accurately estimate the treatment effect of a large increase in social assistance benefits.

The experiments we consider include both carrots and sticks. For example, we simulate the effects of lowering the benefit reduction rate and of expanding earned income tax credits. These exemplify policies designed to improve economic incentives to work and have been discussed and/or implemented both in Quebec and in other jurisdictions. We also simulate the effects of changes to the social assistance program that are designed to make reliance on welfare less appealing, such as reducing benefit levels.

As mentioned above, we focus on the labour supply incentives, manifested through wages and government transfers, facing this group. Because the dropout decision itself is important we also simulate the potential impact of a mandatory high school reform where the decision to drop out of high school is removed on employment and welfare use. The findings from this exercise illustrates the importance of permanent idiosyncratic heterogeneity (such as preferences for education and scholastic aptitudes) for success in the labour market, especially when compared to (marginal) changes in the economic environment.

The results in this paper highlight the importance of long-term preferences for work and leisure as well as financial incentives. In order to reduce the number of welfare caseloads among low-skilled youth and improve their labour market outcomes, policy makers need to adopt a long-term strategy to improve skill acquisition in school, starting at young ages, and reduce economic barriers to enter employment in jobs with low earnings. The current welfare eligibility criteria in Quebec offer limited reasons to give up welfare payments in exchange for part-time work or full-time work at modest wages. This policy environment risks cementing individuals in welfare dependence and permanently exclude them for the labour market. This has serious consequences at both the individual and society levels and with rapidly changing labour markets, addressing these concerns are more important than ever.

The rest of this paper is organized as follows. The next section reviews the related literature on the effect of government policies on labour supply and welfare participation. In section 1.3, we describe the data in detail, and this is followed by a description of the methodology framework, both the economic model and its estimation. Section 1.6 contains a discussion of the results and in particular the counterfactual simulations together with a summary of the findings and policy recommendations. The paper ends with concluding remarks in Section 1.7.

1.2 Literature

Most of the previous studies on the impact of government policies on labour supply and welfare use, both in Canada and internationally, have focused on single mothers or two-adult households. Only limited evidence is available for the at-risk group of young, unattached and low-skilled individuals, despite the fact that unemployment and welfare participation is, and has persistently been, very high for this group. The analysis in this paper aims to address this vacuum in the literature.

This section briefly reviews the labour market and welfare participation outcomes of social-assistance (SA) and income-support programs for low-income households in other studies, as well as the relationship between these programs and human capital accumulation in Canada-specific studies.

Social assistance benefits are intended as a last resort for needy families and financial assistance provides them with a minimum level of wellbeing. Nonetheless, they are often criticised for creating work disincentive effects and welfare dependency. Several studies looking at the effect of benefit transfers in various programs (see for example Hoynes (1996) and Moffitt (2003) for the US; Allen (1993), Lemieux and Milligan (2008) and Hansen and Liu (2015) for Canada and Bargain and Doorley (2011) and Bargain and Doorley (2017)for France) show that these programs are associated with a significant reduction in the employment level of participants.

An important parameter in the social assistance calculation is the rate at which benefits are reduced if recipients work. This rate has been set at as high as 100 percent to avoid attracting more people to the program.[4] However, since it is an implicit tax on work, it is also criticised for reducing work incentives and increasing welfare dependency. The empirical evidence varies. Some studies (see Hoynes (1996), Keane and Moffitt (1998) and Huffman* and Jensen (2005)) show that the benefit reduction rate has a very small effect on labour supply and actually increases participation in other parallel income support programs such as the food stamps program in the US. Other studies show that the size of the effect depends on family type heterogeneity (Moffitt (1986)) and on the skills and education level of the treated population (Bargain and Doorley (2017)).

To encourage work, policies were designed to reward work by using some form of wage or earnings subsidy. Some of these programs are available to a specific group of workers (often families with children, such as the UK Working Family Tax Credit (WFTC)) while others are general and based on demographic status (such as the US Earned Income Tax Credit (EITC), the Canadian Work Income Tax Benefit (WITB) and the Quebec Work Premium (WP)).

In addition to economic incentives, government can consider imposing work requirements in order to receive social assistance. Riddell and Riddell (2014) noted however that one of the concerns with a work requirement component in welfare programs is their potentially negative effect on educational attainment. Some studies investigate whether skills and educational enhancements can keep people off welfare through government-training programs (Fortin et al. (1999) and Gilbert et al. (2011)) or through further schooling (Barrett (2000)). The results suggest that most training programs and further schooling are unsuccessful in keeping participants away from welfare and in increasing employment, especially among men.

[4] This is the current rate in Quebec and in France's Revenue minimum d'insertion (RMI). It was the rate that applied for the American Aid to Families with Dependent Children (AFDC) program.

Experimental evidence from the Canadian Self-Sufficiency Program (SSP) show that people respond to economic incentives and the provision of a generous wage subsidy increased employment and reduced welfare participation among the participants.[5] These positive responses persisted throughout the respondents' participation in the program but faded out a few years after the payments ended. SSP has generated numerous studies on related topics, including the heterogeneous effects of the subsidy on earnings, government transfers and income distribution (see Bitler et al. (2008)) that should be considered when evaluating the true effects of any program. The Action Emploi (AE), a pilot project inspired by the SPP and implemented in Quebec in 2001, was successful in both increasing the time off welfare and reducing the time on welfare (see Lacroix and Brouillette (2011)).

In addition to experimental evidence, there is a large literature evaluating the American Earned Income Tax Credit (EITC) program and the UK Working Families Tax Credit program (WFTC) using survey or administrative income data.[6] Most of these focus on single parents and families with children, as the EITC is substantially more generous for families with children and the WFTC is only available to this group. These studies generally found that: (i) the tax credits significantly increase the labour force participation of single parents; (ii) the overall effect on couples tends to be small and in some cases negative; (iii) the effect varies across gender (Brewer et al. (2006) and Eissa and Hoynes (2004)) and initial employment status (Blundell and Hoynes (2004)). Generally, the results on hours of work are ambiguous and tend to be small for those in the labour force (Athreya et al. (2014); Gregg et al. (2009) and Yang (2018)).

The corresponding tax credit programs in Canada, (the federal Working Income Tax Benefit, WITB and the Quebec Work Incentive Program, WP), have not been analysed as extensively as the EITC and the WFTC. Annabi et al. (2013) show that labour market participation among low- and medium-skilled single parents and single persons increases significantly in response to the WITB. Similarly, Brouillette and Fortin (2008) and Moffette et al. (2013) show that the WP program increases the labour supply among single parents and single persons. Also, Moffette et al. (2013) find

[5]See for example, Card and Robins (1996), Card and Robins (2005) and Card and Hyslop (2006).

[6]For studies of the EITC, see Hoynes (1996), Hotz (2003), Eissa and Hoynes (2006), Eissa and Hoynes (2004), Chyi (2012), Athreya et al. (2014) and Yang (2018). For the WFTC, see Brewer et al. (2006), Blundell and Hoynes (2004), Blundell et al. (2000), Gregg et al. (2009) and Francesconi and Van der Klaauw (2007).

that this policy increases labour supply at both the intensive and extensive margins for singles and couples, although single mothers respond negatively at the intensive margin.

1.3 Data

In the empirical analysis we use data from the Canadian Census for the years 2006 and 2016. The Census is a mandatory survey with a cross section design, conducted by Statistics Canada every five years to provide information on demographic, social and economic characteristics of the population. The personal demographic information from the survey is linked to personal income tax and benefits record from the previous year collected by Canada Revenue Agency (CRA).

Most of the analysis is using information from the confidential master files which can only be accessed at one of the Statistics Canada data centers (one of which is located at McGill University). Each data file contains millions of observations representing about 25 percent of the population and is sufficiently large for separate analyses on sub-populations, such as young male high school dropouts in Quebec. The research that was conducted after the outbreak of Covid-19, which closed our access to the Research Data Center, used data from the public use files of the 2006 Census instead.[7] This includes estimation and simulation of responses to the following two policies: (i) mandatory high school graduation and (ii) replacing the Quebec income tax and transfer system with that of Ontario.

The 2006 public use file contains information on close to one million records, representing 2.7 percent of the Canadian population, which is still large enough to produce reliable results. However, unlike the 2016 Census, the data on welfare benefits is incomplete in the 2006 Census. Specifically, welfare benefits are merged with other government transfers, including not only social assistance transfers but also different refundable credits. To obtain a welfare benefit amount for everyone, we subtracted the estimated amount of refundable credits each individual in our sample would receive, given his income level, from the government transfers variable. Information is provided in the notes to each table below that clearly identifies what data source has been used.

[7] We were not able to use the public file of the 2016 Census since there is no information on hours of work in that file. While we can reasonably estimate the amount of welfare transfers individuals received, it is not possible to do this for hours of work.

1.3.1 Selected Labour Market Outcomes

In this section, we present descriptive evidence on different labour market outcomes for relevant groups in Canada. For instance, in Table 1.1 we present high school dropout rates for males in 2006 and in 2016. We describe how the rates differ across ages and provinces. The entries in the top panel are for unattached males while those in the bottom panel are for all males, regardless of attachment. While the dropout rates are generally higher among the latter group, the pattern of decreasing rates with age, differences between Census years and across provinces remains. Since our empirical analysis focus on unattached males, our discussion will focus on the dropout rates for this group.

The first two columns provide information for Quebec and as can be seen, the dropout rate decreases as we increase the age of the group. For instance, among 21-22 years olds in the 2016 Census, 19.9 percent had not completed high school, while this number is 14.1 percent among males 23-24 years old. There are many possible reasons for the significant reduction, including a delayed entry in primary school, a return to complete secondary school at a later than normal age (after having left school) and a delay in grade progression during primary and/or secondary levels. As we increase the age of the cohort, the dropout rate decreases until age 25-26.

Table 1.1: Male High School Dropout Rates in 2006 and 2016, by Age and Province

Age	Quebec		Ontario		Alberta		British Columbia	
	2006	2016	2006	2016	2006	2016	2006	2016
	Unattached Males							
21-22	20.2	19.9	16.9	8.5	16.9	11.1	12.1	7.7
23-24	17.0	14.1	12.3	7.0	16.2	9.1	11.1	6.8
25-26	13.1	12.8	9.6	7.3	15.1	8.9	10.9	6.7
27-28	12.7	13.3	9.4	7.4	13.0	7.9	10.6	7.4
	All Males							
21-22	21.3	18.5	15.4	8.6	21.5	12.9	13.0	8.3
23-24	18.6	16.6	13.0	8.4	20.3	12.4	12.7	8.0
25-26	16.8	15.9	11.8	9.3	19.1	12.1	12.6	8.8
27-28	15.7	15.2	11.4	9.2	17.2	11.7	13.0	9.0

Note: All entries are weighted and rounded. 2006 and 2016 Canadian Census, master data.

A second finding from the first two columns of Table 1.1 is the stability in dropout rates between 2006 and 2016. With the exception for the age group 23-24, where the rate dropped from 17 percent in 2006 to 14.1 percent in 2016, the dropout rates were similar in both years. The entries in the

Table 1.2: Female High School Dropout Rates in 2006 and 2016, by Age and Province

Age	Quebec 2006	Quebec 2016	Ontario 2006	Ontario 2016	Alberta 2006	Alberta 2016	British Columbia 2006	British Columbia 2016
	Unattached Females							
21-22	10.1	12.1	10.8	5.4	13.1	7.9	9.3	4.7
23-24	6.9	8.2	6.3	4.6	8.9	5.1	6.7	4.4
25-26	6.3	6.3	6.1	4.2	9.0	5.6	5.8	4.5
27-28	5.1	5.4	4.5	3.8	7.2	5.0	6.3	4.4
	All Females							
21-22	13.8	12.3	11.3	6.5	18.5	10.1	10.9	5.9
23-24	12.1	11.7	9.7	6.4	16.1	10.1	9.8	6.1
25-26	11.1	10.5	9.1	6.8	15.6	9.8	9.0	6.4
27-28	11.1	9.9	8.4	7.0	14.0	9.7	10.2	6.6

Note: All entries are weighted and rounded. 2006 and 2016 Canadian Census, master data.

remaining columns of Table 1.1 (dropout rates for Ontario, Alberta and British Columbia) show, similar to Quebec, a pattern with decreasing dropout rates as we increase the age. However, unlike Quebec, the dropout rates at any age were significantly higher in 2006 than in 2016.

Table 1.2 shows corresponding high school dropout rates for females in 2006 and in 2016. Like Table 1.1, we see a significant decrease in high school dropout rates for all provinces as we increase age. The table entries also show higher dropout rates in 2016 for Quebec than the other provinces. And finally, comparing Table 1.2 to Table 1.1, it is evident that dropout rates are significantly higher for males than for females. This is true for all provinces, for both years and for all age groups.

In tables 1.3, 1.4 and 1.5, we illustrate differences in selected labour market outcomes for male high school dropouts and graduates. We again compare Quebec experiences with those of the three other major provinces in Canada as well as how average outcomes have changed over time. Table 1.3 shows male average hourly wages (in constant 2015 dollars) for the two educational groups in 2005 and in 2015. Table 1.4 shows average employment rates for the same years and Table 1.5 presents information on welfare participation for 2015.[8]

The hourly wage for each worker, used in Table 1.3, was obtained by dividing annual earnings from the year before the survey year (that is, 2005 and 2015, respectively) with annual hours of

[8]There was no direct question on social assistance or welfare in the 2006 Census which makes a comparison between 2005 and 2015 difficult. For this reason, we limit the presentation to data from the 2016 Census.

Table 1.3: Average Hourly Wages in 2005 and 2015 for Males, age 25-30, by Province, in 2015 dollars

High School	Quebec		Ontario		Alberta		British Columbia	
	2005	2015	2005	2015	2005	2015	2005	2015
Dropouts	18.3	20.8	20.2	22.1	24.5	31.4	18.8	22.5
difference%		13.3%		9.7%		28.1%		19.7%
Graduates	20.5	22.4	22.7	24.0	24.7	32.8	21.1	25.5
difference%		9.3%		5.5%		32.9%		20.6%
Graduates/Dropouts	11.5%	7.5%	12.7%	8.4%	0.7%	4.5%	12.4%	13.3%

Note: Sample of unattached males. All entries are weighted. 2006 and 2016 Canadian Census, master data.

Table 1.4: Average Full-time Employment Rates in 2005 and 2015 for Males, age 25-30, by Province

High School	Quebec		Ontario		Alberta		British Columbia	
	2005	2015	2005	2015	2005	2015	2005	2015
Dropouts	83.8	85.1	84.2	82.4	95.6	93.0	89.3	89.5
difference%		1.5%		-2.1%		-2.7%		0.2%
Graduates	93.7	93.8	95.2	95.2	98.7	97.8	93.2	97.0
difference%		0.1%		0.0%		-0.9%		4.1%
Graduates/Dropouts	11.8%	10.2%	13.2%	15.5%	4.3%	5.1%	4.3%	8.4%

Note: Sample of unattached males. All entries are weighted and rounded. 2006 and 2016 Canadian Census, master data.

Table 1.5: Average Welfare Participation Rates in 2015 for Males, age 25-30, by Province

	Quebec	Ontario	Alberta	British Columbia
Dropouts	19.0	22.1	5.7	9.4
Graduates	7.3	7.0	2.1	2.5
Graduates/Dropouts	0.38	0.32	0.36	0.27

Note: Sample of unattached males. All entries are weighted and rounded. 2016 Canadian Census, master data.

work. The latter measure is constructed using information on the number of weeks worked in 2005 and 2015 as well as data on average weekly work hours at the time of the survey. This is a standard approach to construct hourly wages in survey data.

For Quebec, the average hourly wage in 2015 was \$20.8 for high school dropouts and \$22.4 for high school graduates. This represents a 7.5 percent wage premium to complete high school. The average wages are higher in Ontario than in Quebec for both educational groups but more so for high school graduates. As a consequence, the high school wage premium is a bit higher in Ontario, 8.4 percent, than in Quebec. In Alberta, the average wages are similar for both high school dropouts and graduates and substantially higher than the corresponding wages in Quebec and Ontario ($31.4 and $32.8 for dropouts and graduates, respectively. Finally, wages in British Columbia are higher than in Quebec and Ontario ($22.5 and $25.5 for dropouts and graduates, respectively) but below those in Alberta and the high school premium is 13.3 percent. Overall, the average wage for dropouts is lowest in Quebec among the four provinces suggesting that this group is doing worse in terms of wages in Quebec than in the other three provinces. The average wage for high school graduates is also lowest in Quebec and, with the exception of Alberta, the relative wage (high school graduates vs dropouts) is lowest in Quebec. Finally, a comparison of average wages over time shows that real wages increased 9-13 percent per year for both dropouts and high school graduates in Quebec, a bit less for high school graduates in Ontario and significantly more in Alberta and British Columbia.

Table 1.4 presents average full-time employment rates for 2005 and 2015 for dropouts and high school graduates. The rates are calculated using information on any full-time employment during the year. There are two clear and common patterns in the table. First, the employment rate is always lower for dropouts.[9] Second, the employment rates, in all provinces except for Quebec, are lower in 2015 than in 2005. In 2015, just over 85 percent of dropouts worked full-time in Quebec. Among high school graduates, 94 percent worked in Quebec that year. Overall, the data conclusively show that a high school education serves as an employment insurance when compared to dropping out and the difference in full-time employment rates between the two groups was similar between 2005 and 2015 in Quebec and increased in Ontario, Alberta and British Columbia .

Average welfare participation rates in 2015 are shown in Table 1.5 and not surprisingly, the rates are significantly higher for dropouts than for high school graduates, in all provinces.[10] For Quebec, 19 percent of single, unattached males 25-30 years old received welfare. Among high school graduates, the corresponding figure is 7.3 percent. A similar pattern is observed for Ontario where the participation rates are 22 percent and 7 percent, respectively. For the western provinces, participation rates are lower for both educational groups, especially in Alberta.

[9] Similarly, the unemployment rates are always higher for dropouts, results are available upon request.

[10] The participation rates show the proportion of the sample who received any social assistance or welfare payments during 2015.

To sum up, young male Quebeckers that do not complete their secondary education are behind similar individuals in Ontario in terms of full-time employment and hourly wages conditional on working. Further, they are well behind similar individuals in British Columbia and especially in Alberta.

1.3.2 Sample Selections and Characteristics of the Sample

The sample criteria used to construct the tables discussed above are similar to the sample used to estimate the economic models. Our main analysis will be based on a sample of dropouts drawn from the 2016 Census Master file and selected features of this sample are presented in Table 1.6. The first column shows entries for those who received welfare while the second column presents information for those with no welfare.[11] As expected, hours of work and earnings are strongly correlated with welfare participation status. A vast majority (95 percent) of welfare recipients were not working at all and among those who worked, none worked more than 1,200 hours during the year. For the other group, those without welfare, the proportion of non-workers is much lower, 3.4 percent and more than half of these respondents worked full-time (more than 1,800 hours per year).

Overall, in this sample, the proportion who received welfare during 2015 was 15.4 percent. This is slightly lower than the corresponding number in Table 1.5 and due to the fact that we include those 19-24 years old in Table 1.6. Further, the proportion of foreign-born respondents is slightly lower among welfare recipients, 6.4 percent compared to 8 percent for those without welfare. The average age is similar for both groups, around 25. The mean hourly wage among those who work is $19.1 for the non-welfare group. We are unable to report a mean wage for the welfare group because of sample size restrictions.

As mentioned above, we also use data from the public use files of the 2006 Census and the main features of this sample are presented in Table 1.7. Because of sample size limitations, we included high school graduates in the 2006 sample and this may explain some of the differences compared to Table 1.6. For example, the welfare participation rate is lower (7.3 percent compared to 15.4 percent in Table 1.6). Further, average hours of work among those who work is higher in the 2006 data.

[11] Welfare participation is determined using information on receipt of social assistance benefits during 2015. Anyone who received any benefits is recorded as a welfare recipient.

Table 1.6: Characteristics for the Quebec Sample of Male Dropouts, age 19-30, in 2016

	Receiving welfare	Not receiving welfare
Proportion with annual hours (h):		
$h = 0$	94.8	3.4
$h < 1,200$	5.2	18.2
$1,200 \leq h < 1800$	-	20.9
$1,800 \leq h$	-	57.6
Welfare Participation	15.4	
Proportion Immigrant	6.4	8
Average Age	25.4	25.1
Average Annual Hours of Work	38	1,654
Average Hourly Wage	-	19.1
Number of Observations	1,165	6,385

Note: Characteristics for sample of unattached, males. All entries are weighted. 2016 Canadian Census, master data. There were too few wage observations for the welfare-receiving group.

The entries in Table 1.7 show that education is clearly correlated with welfare receipt. The proportion of high school dropouts among welfare recipients is 43.5 percent while it is 14.9 percent among the non-welfare group. The reverse is observed for high school graduates with a certificate while for high school graduates without a certificate, the proportions are similar in both categories. The real average hourly wage is slightly higher for the non-welfare group in the 2006 data than for the 2016 data. This suggests a negative real wage growth for this group between 2006 and 2016. However, as mentioned before, the groups are not comparable as the 2006 sample includes high school graduates.

1.4 Model Description

We derive and estimate an economic model of labour supply and welfare participation in order to counterfactually simulate behavioral changes from alterative labour market policies. Estimation of a structural model of this sort is needed to compare predicted responses to such policy changes. We follow a well-established literature and model the two outcomes using a discrete choice model. Specifically, we reduce the dimension of annual hours of work (h) into the following four categories:

Table 1.7: Characteristics for the Quebec sample of males, age 20-34, in 2006

	Receiving welfare	Not receiving welfare
Proportion with annual hours (h):		
$h = 0$	80.4	3.5
$h < 1,200$	19.6	8.1
$1,200 \leq h < 1800$	-	16.2
$1,800 \leq h$	-	72.3
Age Categories		
20-24	16.3	20.1
25-29	34.8	42.5
30-34	48.9	37.4
Education		
HS Dropouts	43.5	14.9
HS Grads + Dropouts with a certificate	23.9	27.0
HS Grads with a Certificate	32.6	58.1
Welfare Participation	7.3	
Proportion Immigrant	13.0	7.5
Average Annual Hours of Work	110	1,887
Average Hourly Wage	-	21.5
Number of Observations	92	1,164

Note: Characteristics for sample of unattached males in Quebec. 2006 Canadian Census, public files. Hourly wage is in 2015 dollars. There were too few wage observations for the welfare receiving group.

$$\begin{cases} \text{hour} = 0, \\ 0 < \text{hour} < 1200, \\ 1200 \leq \text{hour} < 1800, \\ 1800 \leq \text{hour}. \end{cases}$$

The categories with positive hours correspond to two forms of part-time work as well as full-time work and this discretisation or grouping is supported by the observed distribution of annual hours in the data.

We assume that individuals choose how much to work by maximising their utility, which depends not only on income and leisure (the opposite of working), but also on welfare participation. We also assume that the utility function decreases with welfare participation and increases with both leisure and income. We assume that the disutility of participating in a social assistance program primarily reflects the non-monetary costs, such as the "stigma" associated with participation. This factor is included to explain why some eligible individuals choose not to participate. Following van Soest (1995) and Hansen and Liu (2015), We use a trans-log specification of the direct utility function, where for a given individual, we have

$$U(C_j, L_j) = b_c ln(C_j) + b_l ln(L_j) + b_{csq}(lnC_j)^2 + b_{lsq}(lnL_j)^2 + 2b_{cl} ln(C_j) ln(L_j) - b_{SA} d_{SA}, \quad j = 1, ..., 4 \quad (1)$$

where C_j is after tax income for hours combination j, L_j is the annual hours of leisure time at j, d_{SA} is equal to one if the person receives social assistance and is zero otherwise. Finally, b_c, b_l, b_{csq}, b_{lsq}, b_{cL} and b_{SA} are all preference parameters to be estimated.

We assume that (b_{SA}) or the disutility of receiving welfare is separable from the utility of disposable income and leisure (following Moffitt (1983), Hoynes (1996) and Hansen and Liu (2015)).

The individual chooses leisure time (L), consumption (C) and welfare participation (d_{SA}) by maximising the utility function subject to the following budget constraint:

$$C_j = I_j + B_{SA}(I_j) \cdot d_{SA}, \quad (2)$$

where

$I_j = wh_j + Y_j - t(wh_j + Y_j(taxable) - D_j)$,

$B_{SA}(I_j) = 7,392 - \mu(I_j - 2,400)$ for $0 \leq \mu(I_j - 2,400) < 7,392$,

$B_{SA}(I_j) = 0$ for $7,392 \leq \mu(I_j - 2,400)$,

and where w equals the before-tax hourly wage rate, Y_j denotes annual non-labour income (not including social assistance). $Y_j(taxable)$ represents taxable non-labour income. The function $t(\cdot)$ determines income taxes and D_j is deductions. Lastly, $B_{SA}(\cdot)$ is a function that calculates welfare benefits.

In Quebec in 2015, single unattached individuals could receive a maximum $616 per month (or

$7,392 per year) and income up to $200 per month ($2,400 per year) were exempted and had no impact on welfare benefits. Any income above this amount was reduced by the parameter μ, which in Quebec was (and still is) equal to 1. This effectively means a 100 percent implicit tax rate on any income exceeding $2,400 per year for welfare eligible workers.

Including the disutility of social assistance participation means that an individual has $4+2$ possible work and welfare combinations to consider. There are four work options without welfare and two with welfare ($h = 0$ and $0 < h < 1200$). We do not include the two options with more hours (including full-time) since the earnings associated with them exceed the eligibility threshold for social assistance. Some welfare states may remain infeasible if the individual's income is high enough to disqualify them from social assistance. To solve the optimization problem, the utility function must be evaluated for each possible combination of work hours and social assistance participation, with the individual choosing the state that provides the highest utility.[12]

1.5 Estimation

To estimate the behavioural parameters of the model, we add random shocks to the utilities of all choice options:

$$U_{j,r} = U(C_j, L_j, SA_r) + e_{j,r}. \tag{3}$$

Here, j is the individual's labour supply choice, r denotes their welfare participation state, and $U_{j,r}$ is the person's actual utility from choosing the combination (j, r). We assume that the error term $e_{j,r}$ follows a type I extreme value distribution. This error term can be viewed either as an optimisation error (i.e., a mistake in the person's assessment of the utility from selecting a particular work-welfare combination (j, r)) or as an unobserved, alternative-specific utility component. Based on the assumed distribution of the stochastic terms in the utility function, the contribution to the likelihood function for each individual is given by

[12] An important advantage of the discrete approach to labour supply is that it allows us to include very detailed information on income. In addition to the hourly wage, described above, we use detailed information on income tax rates at both provincial and federal level, tax credits, deductions and government transfers. For each labour supply choice, we calculate a net or disposable income which is assumed to govern individuals' labour supply decisions. Information on these calculations are available in Table A.1 in appendix.

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

where $p_{j,r} = \frac{exp(U_{j,r})}{\sum_{r=1}^{2}\sum_{j=1}^{4} exp(U_{j,r})}$ and $d_{j,r}$ is an indicator of each individual's observed state.

Previous research has highlighted the significance of accounting for heterogeneity in preferences (e.g., Hoynes (1996) and Hansen and Liu (2015)). In this study, we follow the work of Hansen and Liu (2015) by incorporating heterogeneity in preferences related to leisure, consumption, and welfare through the parameters

Previous research has highlighted the significance of accounting for heterogeneity in preferences (e.g. Hoynes (1996) and Hansen and Liu (2015)). Here, we follow Hansen and Liu (2015) and incorporate heterogeneity in preferences related to consumption, leisure, and social assistance through the parameters b_l, b_c and b_{SA}. These parameters are modeled as functions of observable characteristics such as age, immigrant status, language (English, French) and bilingual ability (English and French).

To generate a predicted distribution of hours that closely resembles the observed one, we include a term that accounts for the fixed costs associated with work. These costs may include work-related expenses, such as transportation. When such costs are present, some individuals may not find it optimal to work only a few hours. We follow van Soest (1995) and replace $ln(C)$ with $ln(C) - ln(FC)$ for workers. FC is a parameter to be estimated and represents fixed costs of work. Since utility increases with income, these positive costs (FC) decrease the utility of working while not impacting the utility of not working.

The likelihood function described above needs to be adjusted to incorporate wages. In the model, we compare all six options for each individual and assume that the chosen option is the one that provides the highest utility. To do this, we need a wage for every individual, including those who do not work and have no observed wage. Consequently, we estimate a wage function along with the choice model above and use it to generate wage predictions. We use information on age, immigrant status, language and area of residence (urban vs rural) in estimating the wage parameters, and specify the estimating equation as

$$ln(w_i) = \mathbf{X_i}b + v_i, \tag{5}$$

where w_i represents the individual's hourly wage (before income taxes), X_i includes individual observed characteristics and v_i is a random error term. We estimate the vector b using data on workers and predict hourly wages for non-workers using this wage equation. Specifically, the likelihood contribution from estimating the wage parameters is given by

$$l(w) = (\frac{1}{s}f(\frac{v}{s}))^{d(w)}, \tag{6}$$

where s is the standard deviation of the wage error term, v and $f(\cdot)$ is the standard normal density function. $d(w)$ is an indicator function that takes the value of one if a wage is recorded for a specific individual and zero if it is not. The overall likelihood function is therefore,

$$l = l(w)\sum_{r=1}^{2}\sum_{j=1}^{4} p_{j,r}d_{j,r}. \tag{7}$$

In the primary analysis, we limit our sample to high school dropouts. We also estimate a version of the model where we include high school graduates and estimate the probability of graduation. This extension allows us to simulate the effect of policies designed to improving graduation rates. It also allows us to compare impacts of economic incentives with those obtained by changing behaviour.

1.6 Results

The model presented above is estimated on a sample of high school dropouts in Quebec taken from the 2016 Census whose main features were described in Table 1.6. A table with the estimated parameters are available in Appendix (Table A.2) and we will briefly comment on some of them before discussing results from a range of simulations. As described in section 1.4, we allow some preference parameters in the utility functions to depend on observable characteristics, such as age, immigrant status, language and education. The estimates suggest that preferences for leisure is higher among those who speak English only and decreases with age. Consumption preferences on the other hand are not linked to age but are lower among immigrants. For welfare, the parameters show the extent of disutility perceived by different individuals and the estimates suggest lower disutility or "stigma" for older respondents and higher "stigma" for those who speak English only.

The wage estimates show a wage penalty to English only and a positive wage effect from graduating from high school. The high school graduation equation results suggest that students that speak French only are less likely to graduate while immigrants and those who speak English have higher graduation probabilities, as do those living in an urban area.

In Table 1.8, we present the difference between predicted and actual proportions in each of the six hours classes using results from the main model applied to the 2016 Master file. In the first column, we show the differences for those predicted to receive welfare while the entries in the second column show the corresponding differences for those not on welfare. In all cases, the differences are small. The largest difference is for males working between 1,200 and 1,800 hours and who do not receive welfare at 1.6 percentage points.[13] This means that the model assigns a larger proportion of the sampled respondents to this category than what is observed in the data. However, a 1.6 percentage point difference is small, and the table entries indicate that the model predictions are similar to those in the data.

Table 1.8: Difference between predicted and actual proportions in different hours in 2016

	Receiving welfare	Not receiving welfare
Hours Category:		
$h = 0$	-0.88	-0.04
$0 < h < 1,200$	0.88	-1.49
$1,200 \leq h < 1,800$	-	1.56
$1,800 \leq h$	-	-0.04

Note: Predictions based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, male dropouts. Actual and predicted proportions were not disclosed by the data center for confidentiality reasons.
Negative means under-predicted. 2016 Canadian Census, master data.

[13]It should be noted that the numbers in Table 1.8 show differences in unweighted sample proportions in each category and we are not allowed to disclose the actual proportions. This also means that these entries are not directly comparable to the sample proportions in each hours class in Table 1.6.

1.6.1 Policy Simulations

We use the model and the estimated preference parameters to generate counterfactual distributions of outcomes following different hypothetical policy experiments. For each simulation, we present changes in the employment rate, annual hours of work, welfare participation and disposable income. In order to assess the public finance impact of each reform, we complement this list with changes in 'net taxes'. In all cases, predictions generated after a policy reform is (counterfactually) implemented are compared to predictions obtained using the model given the actual policy environment.

Changes to the Social Assistance Program

A common feature of many welfare programs is a very high implicit tax rate on earnings, often the rate is 100 percent. The implicit tax rate, or equivalently the benefit reduction rate (μ above), shows how much benefits a welfare recipient must give up as earnings increase. For example, assume a person who has no earnings and receives \$600 per month in welfare. If this person starts working 10 hours per week at an hourly wage of \$10, his earnings of \$400 will reduce his welfare payments with the same amount (abstracting from any taxes and credits). Adding fixed costs of working (such as transportation to and from work), working may not provide any financial gains. Moreover, leisure is generally valued by individuals and if so, this will further reduce the willingness to work under this benefit structure. Thus, low-wage, part-time work (often the type of work that is available for welfare recipients and other low-skilled workers that are weakly attached to the labour market) is not attractive. However, a reduction in the benefit reduction rate may also increase the number of welfare recipients since more workers are eligible for welfare.

In Table 1.9 we present results from counterfactually reducing the benefit reduction rate from the current 100 percent to 50 percent. The table entries are generated by predicting outcomes (hours of work, wages and welfare participation) for everyone in the data using our model and the estimated parameters. The top panel shows predicted proportions of workers and welfare participants as well as average annual hours of work and net taxes under the 2015 policy environment. When we take the average across everyone in the sample, the model predicts an employment rate of 83.1 percent and a welfare participation rate of 14.4 percent. These numbers compare very favorably to those observed in the data (82.5 and 15.4 percent, respectively; see Table 1.6). Further, average annual hours overall equal 1,491 and as expected, there is a large difference in labour supply depending on predicted welfare recipiency (1,736 hours for those predicted to be off welfare versus 38 for those predicted to receive welfare). In the last column we present average net taxes per individual, defined as income taxes minus tax credits and government transfers, to shed light on the public finance aspect of each policy simulation. Evaluated using the 2015 policy environment, those who did not receive welfare contributed on average $5,571 per year while those on welfare had a negative contribution of $8,164.

Table 1.9: Simulated Impact of Changes to Social Assistance Regulations

Policy	Subgroups	Employment Rate	Welfare Participation	Anuual Hours of Work	Average Change in Disposable Income	Net Taxes
2016 Actual	Overall	83.1%	14.4%	1491	-	3590
Policy	No welfare	96.3%	0.0%	1736	-	5571
Environment	Welfare	5.0%	100.0%	38	-	-8164
Reducing the Benefit	Overall	83.7%	17.3%	1469	-1750	3338
Reduction Rate	No welfare	96.4%	3.3%	1707	-3928	5291
to 50%	Welfare	8.1%	100.0%	61	3540	-8251
A 25%	Overall	83.4%	13.6%	1499	-1407	3893
Benefit	No welfare	96.3%	0.0%	1736	0	5571
Reduction	Welfare	7.0%	94.6%	93	-1407	-6062

Note: Predictions based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, male dropouts. 2016 Canadian Census, master data.

In the bottom panel, we show predictions generated under the counterfactual scenario that the benefit reduction rate is 50 percent instead of 100 percent. All other aspects of the model and economic environment are identical to those used to predict outcomes in the top panel. Consistent with the theoretical predictions, the lower reduction rate substantially increases employment among welfare participants – the employment rate is 8.1 percent after the change compared to 5.0 percent before (a 62 percent increase). The change also increases welfare participation by 3.3 percentage points as more people become eligible. In terms of consumption or net income, those who were initially on welfare and responded to the decreased reduction rate saw an increase with \$3,540. This is entirely coming from retained earnings that in the current regulatory framework is clawed back. For those who were not on welfare before the change, there is a similarly sized decrease in disposable income among those who react to the policy. This results from a reduction in their hours of work as they exchange some earnings for social assistance payments.

In Table 1.9, we also present predictions generated when we reduce the basic social assistance benefit with 25 percent (that is, from \$7,392 per year to \$5,544). In the bottom panel, we show outcomes under the hypothetical policy environment with significantly lower benefit levels. This change only impacts welfare participants and their employment increases from 5 percent to 7 percent and welfare participation drops to 94.6 percent. The average net income among those who start working due to the benefit reduction is reduced by \$1,407. However, this is lower than the change for those who remain out of the labour force as they see a reduction in their benefits with \$1,848 (\$7,392 - \$5,544).

Changes to Federal and Provincial Tax Benefits

We also use the model to explore labour supply effects of two income tax benefit systems: the federal Working Income Tax Benefit (WITB) program and the provincial Work Premium (WP) program. Both programs supplement earnings after a minimum income level is achieved and increase proportionally until a certain income level, at which the credits are reduced by a constant rate. The maximum benefit for a single person in Quebec in 2015 was \$1,634 from WITB and \$558 from WP.[14] These maximums are achieved at the following income levels: \$11,903 for WITB and \$10,370 for WP. Although the Quebec version of this program is less generous than the federal one, the combination of the two implies a significant earnings subsidy for low-wage workers and we will use our model to simulate the impact they have on the willingness to work in our low-skill sample. Since both programs were in effect in 2015, we simulate the effect of eliminating them (instead of adding them) on the same labour market outcomes that we discussed above.

The results are in Table 1.10 and again, for ease of comparison, we include predicted outcomes

[14] The maximum eligible incomes, after which no benefits are paid out, were \$20,072 for WITB and \$15,949 for WP in 2015. For additional details, see appendix.

using the 2015 policy environment at the top of the table. In the second panel, we remove WITB for everyone and this reduces employment slightly for those predicted not to receive welfare and has no effect on employment for those predicted to receive welfare.A very small fraction will start receiving welfare (0.1 percent) and a few will leave welfare (0.8 percent). Average annual hours of work increase for both groups and disposable income among those who work more increase on average by $748 for those who were initially on welfare and by $1,040 for those who were not on welfare initially. The impact on net taxes is an overall saving of $4,093 per person for the government. The increase in hours of work from the removal of the tax credit may seem counterintuitive. However, it is precisely what the theoretical model predicts. Individuals who were working and qualified for the WITB will see a reduction in their overall income and an increase in their hourly marginal wage (because the tax credit is no longer reduced hours for hours). The income reduction is predicted to increase hours of work (assuming leisure is a normal good), this is referred to as the income effect in the labour supply literature. Similarly, the increase in the marginal wage will also increase hours of work and this is referred to as the substitution effect in the labour supply literature. Thus, removing the tax credit increases hours of work for those who already work both theoretically as well as empirically in our model. Tax credits of this type are designed to encourage non-workers to find employment but also reduce hours of work among those who already work and who are impacted by the credits.

The results from removing the provincial Work Premium (WP) program are presented in the third panel and the findings are consistent with those obtained when removing WITB. However, since the WP program is less generous than the WITB, the effects in all categories are also smaller.

The last set of results show outcomes when we remove both the WITB and the WP. The direction of the effects is the same as those discussed above when only WITB was removed but the two programs reinforce each other, and the magnitudes of the responses are greater. Nevertheless, both programs are relatively modest and the changes in labour supply and welfare use are consequently also limited.

Table 1.10: Simulated Impact of Changes Tax Benefit Regulations

Policy	Subgroups	Employment Rate	Welfare Participation	Anuual Hours of Work	Average Change in Disposable Income	Net Taxes
2016 Actual	Overall	83.1%	14.4%	1491	-	3590
Policy	No welfare	96.3%	0.0%	1736	-	5571
Environment	Welfare	5.0%	100.0%	38	-	-8164
Elimination of	Overall	83.1%	14.4%	1543	1032	4093
WITB	No welfare	96.2%	0.1%	1795	1040	6141
	Welfare	5.0%	99.2%	49	748	-8063
Elimination of	Overall	83.1%	14.4%	1507	433	3729
Work Premium	No welfare	96.3%	0.0%	1753	431	5725
	Welfare	5.0%	99.6%	44	478	-8118
Elimination of	Overall	83.1%	14.4%	1552	1184	4172
WITB and	No welfare	96.2%	0.1%	1805	1187	6230
Work Premium	Welfare	5.0%	98.8%	52	1057	-8039

Note: Predictions based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, male dropouts. 2016 Canadian Census, master data.

Mandatory High School Graduation

In addition to policies that directly impacts the financial situation of individuals, we also simulated effects of the role of high school graduation. More specifically, we added high school graduates to our sample (but not those who had acquired any education beyond high school) and allowed both preference and wage parameters to vary across educational attainment. We also estimated the probability of graduating from high school. After having obtained these estimates, we simulated two changes. Because we are unable to access the Master file of the 2016 Census data because of covid-19, we used the public use file of the 2006 Census as described in section 3 above for these simulations.

Before describing the simulation results, we will illustrate the model's capacity to replicate the hours distribution and the proportion of welfare users. In Table 1.11 we present predicted proportions in each of the labour supply and welfare groups. For those not receiving welfare, the labour supply proportions are very similar to those observed in the data (shown in Table 1.7). For those predicted to receive welfare, the model somewhat over-predicts non-employment (88 percent versus

80) but we believe the differences are not substantial enough to invalidate the use of the model to perform simulations.

Table 1.11: Actual hours distribution of low educated males in Quebec, age 20 to 34 in 2006

	Receiving welfare	Not receiving welfare
	Data	
Hour Category (h):		
$h = 0$	80.4	3.5
$h < 1,200$	19.6	8.1
$1,200 \leq h < 1800$	-	16.2
$1,800 \leq h$	-	72.3
	Model	
Hour Category (h):		
$h = 0$	88.3	3.3
$h < 1,200$	11.7	7.6
$1,200 \leq h < 1800$	-	15.2
$1,800 \leq h$	-	73.9

Note: Actual and predicted proportions. Predictions based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, male dropouts, high school graduates and those with a certificate. Negative means under-predicted. 2006 Canadian Census, public files.

In the first simulation, we replaced the wage parameters for the dropouts with the corresponding parameters for high school graduates. The purpose of this simulation is to see the effect on labour supply and welfare use among dropouts when they face the same wage function as high school graduates. The results are shown in Table 1.12. The first row shows average predicted outcomes for the sample when wage and utility parameters differ across the two groups. The model predicts an employment rate of 77 percent and a welfare participation rate of 15 percent. Average annual hours of work are 1,531 and the average net tax equals $5,011 (this number shows the average amount the two governments collect in income taxes minus what they spend on benefits and social assistance).

In the second row, we present the corresponding numbers when we replace the wage function for high school dropouts with that for high school graduates. This implies an increase in predicted hourly wages for the dropouts and, as expected, this generates an increase in employment and hours worked and a reduction in welfare participation. However, the changes are modest.

The second simulation is based on applying both wage and utility parameters for high school graduates to the dropout group. The results are shown in the bottom panel of Table 1.12 and the

Table 1.12: Simulated Impact of Mandatory High School Completion

Employment Rate	Welfare Participation	Annual Hours of Work	Average Change in Disposable Income	Net Taxes
2006 Actual Policy Environment				
77.1%	14.9%	1,531	-	5,011
Same Wage Function For Graduates and Drop-outs				
78.0%	14.5%	1,514	1,245	5,695
Same Wage Function and Preference Parameters for Graduates and Drop-outs				
87.4%	4.2%	1,752	3,426	7,458

Note: Predictions based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, male dropouts, high school graduates and those with a certificate. 2006 Canadian Census, public files.

changes are substantial. The employment rate increases with 10 percentage points (or 13 percent) from 77.1 percent to 87.4 percent and average hours worked increase with 221 hours per year or by 14 percent. Similarly, welfare participation changes substantially with a reduction from 15 percent to just over 4 percent. Consequently, the average change in disposable income and net taxes are significant. Average disposable or net income per year increase with $3,426 compared to the case when outcomes were generated using separate parameters for high school dropouts and graduates. Net taxes collected by the governments also increase significantly, from $5,011 to $7,458. That is, with this change, money in the pockets of the high school dropouts increase and government transfers decrease.

Comparing the effects from this last simulation with the one when only wage parameters were modified, it is clear that educational upgrading (for high school dropouts) will only generate measurable improvements in labour market outcomes if preferences for work and welfare are altered as well as wage offers. This may however be very difficult to achieve and may require significant investments in schools and support systems. Mandating students to complete high school by forcing them to stay in school or by offering them a high school “equivalent” diploma may not be sufficient to change their preferences for work.

However, these findings point to an important aspect of high school dropouts, namely that even if policy makers could force every student who are at risk of dropping out to obtain their high school diploma, there are no guarantees that this would bring them success in the labour market. One reason for this is selection where students are sorted into educational groups based on a range of factors and to counterfactually or forcefully move students up one level does not automatically make them share the same characteristics and preferences as those in that group. This is clearly illustrated in the finding that when we only modify the economic return to a high school diploma (that is increase the hourly wage), we find only modest adjustments on labour supply and welfare use. On the other hand, when we also adjust preferences for work and welfare to align them with those of high school graduates, significant changes present themselves.

Ontario Tax and Transfer System

As a final policy simulation, we used the model and the Quebec sample of high school dropouts and graduates from the public use file from 2006 but replaced the income tax and transfer system with the one that was in place for Quebec in 2015 as well as the one for Ontario in 2015. The reason for this simulation is to explore how labour supply and welfare use of low-skilled workers in Quebec would change if they instead faced the tax and benefit rules that applies in Ontario.

Table 1.13: Predicted proportions in different hours classes in 2015 using estimated Quebec Parameters from the 2006 Census and Different Tax and Transfer Systems

	Receiving welfare	Not receiving welfare
	Quebec Tax and Transfer System	
Hour Category (h):		
$h = 0$	90.28	3.13
$h < 1,200$	9.72	8.70
$1,200 \leq h < 1800$	-	15.46
$1,800 \leq h$	-	72.72
	Ontario Tax and Transfer System	
Hour Category (h):		
$h = 0$	70.21	3.10
$h < 1,200$	29.79	7.83
$1,200 \leq h < 1800$	-	16.35
$1,800 \leq h$	-	72.72

Note: Predictions are based on the model presented in Sections 4 and 5 using estimates from a sample of unattached, low-educated males in Quebec using public files from the 2006 Canadian Census. All the dollar values in tax and benefits schedule are adjusted to 2015 using CPI.

The results are shown in Table 1.13 with predicted proportions under the Quebec tax schedule in the top panel and similar proportion under the Ontario tax schedule in the lower panel. It is important to keep in mind that everything except for the tax schedule is held constant when generating the two sets of results. If Ontario taxes and transfers applies, the average employment rate among welfare recipients would increase from 9.7 percent to 29.8 percent. This is a substantial difference and may be due to parameters governing welfare eligibility and incentives as well as parameters determining income taxes and tax credits. The table entries for those not receiving welfare are however very similar for both tax schedules. The non-employment rate is 3.1 percent in both cases and the full-time rate is also the same at 72.7 percent.

The similarity in predicted labour supply for those not on welfare suggest that differences in welfare eligibility rules between the two provinces is the main reason for the higher predicted employment rates for welfare recipients under Ontario rules, not differences in income taxes. And as shown in Appendix, one of the main differences in welfare eligibility parameters between the two provinces is the benefit reduction rate, which was 100 percent in Quebec and only 50 percent in Ontario. Even if some portion of earnings is exempted, the very high implicit tax rate clearly reduces any incentives to work, especially at jobs with low pay, those that are mostly available to this group.

1.7 Conclusion

For decades, increased trade and technological advances have reduced jobs available for low-skilled workers in most industrialized countries. This development is likely to persist as advances in automation and robotics continue to change labour markets.[15][16]. These developments put pressure on policy makers and particular attention has been given to low-skilled youths who already face limited demand for their skills. This group of low-skilled youths is the focus in this research. We provide a descriptive analysis of their labour market attachment and then analyse the effectiveness of different policies aimed at increasing their involvement and success in the labour market.

Consistent with previous evidence for Canada and other countries, we show that high school

[15] OECD (2016) assess that half of todays work activities can be automated by 2055 and in some scenarios even by 2035. Company and Manyika (2017) reaches similar conclusions.

[16] The negative impact of Covid-19 on the labour market was concentrated on low-skilled jobs, which were lost at a higher rate than other jobs, see Lemieux et al. (2020). The long-term impact of this is unknown, see Jones et al. (2020)

dropouts have weaker connections to the labour market than individuals with higher educational attainment. Unlike much of previous research on welfare use and employment, which has focused on single mothers, we analyse the behaviour of single, unattached males without a high school diploma. Unlike the U.S., these individuals are eligible for welfare in Canada and as we show, a significant proportion of them receive social assistance benefits, and more so in Quebec and Ontario than in Alberta and British Columbia. The welfare participation rates drop substantially when we consider similar males with a high school diploma.

An instinctive policy response may therefore be to impose a mandatory high school degree and prevent any students from dropping out. However, as we also show, this is not guaranteed to improve outcomes for those who will be forced to remain in school. This finding is consistent with recent research using Danish data, see Hansen et al. (2020). It is also in line with a large volume of research on educational attainment using structural approaches, see for instance Cameron and Heckman (1998) and Eckstein and Wolpin (1999). Many of these papers report that pre-market factors, such as cognitive and non-cognitive skills of students, are important. However, improving these skills is challenging and likely to be more successful at younger ages.

While enhancing the skill levels of high school dropouts may increase the demand for their labour, we argue that the supply side also matters. In particular, safety nets in place to prevent poverty may come with strong disincentive effects and reduce the willingness to work. We show that the benefit eligibility rules for social assistance in Quebec are not designed to encourage recipients to work. Any earnings above $200 per month reduce welfare benefits dollar for dollar. This means that it is not economically worthwhile for a welfare recipient to take a part-time job at a low wage, even if such a job can lead to more hours and higher wages in the future. Our simulations show that if Quebec were to adapt Ontario's benefit regulations, where additional earnings are reduced at 50 cents per dollar instead, the proportion of welfare recipients who work will increase by a factor of three. This will increase their net income and only marginally reduce public finances (more money spent on social assistance benefits, but also more income tax collected).

We also examined the effects of removing the tax credit programs currently in place that are designed to encourage employment. This exercise shows the expected impact on labour supply from these programs and our findings suggest limited changes in employment, welfare participation and hours worked. The small effects do not however imply that tax credits are ineffective policy instruments. Instead, these reactions are expected because of the small magnitudes of tax credits in the current system.

To sum up, the findings in this paper highlight the importance of long-term preferences for work and leisure and financial incentives. To reduce the number of welfare caseloads among low-skilled youth and improve their labour market outcomes, policy makers need to adopt a long-term strategy to improve skill acquisition in school, starting at young ages, and reduce financial obstacles to enter employment in jobs with low earnings. The current eligibility criteria offer very limited reasons to give up welfare checks in exchange for part-time work or full-time work at modest wages. This policy environment risks cementing individuals in welfare dependence and permanently exclude them from the labour market. This has serious consequences at both the individual and society levels and with rapidly changing labour markets, addressing these concerns are more important than ever.

Chapter 2

The Role of Unobserved Preference Heterogeneity in Labour Supply Outcomes: A Monte Carlo Simulation Study

2.1 Introduction

The structural discrete labour supply model is a popular tool to predict labour supply responses to tax and welfare reforms in a partial equilibrium framework. This model uses a direct utility function where an agent chooses a combination of income and leisure that maximises utility. The researcher can include a detailed tax and credit schedule that makes these models an attractive method for ex-ante evaluation of specific tax policies.

In the discrete structural labour supply literature, it is commonly assumed that the error terms in the utility function have extreme value distributions and that they are uncorrelated among the alternatives. This assumption significantly simplifies the estimation process because the model can be estimated using conditional logit. However, the main drawback is that it assumes that the relative preferences or odds between two alternatives are not affected by introducing a third alternative (the independence of irrelevant alternative property). This is a restrictive assumption in many areas including labour supply. For example, let's assume that an individual faces two alternatives: not working and working 30 hours per week and that he has a preference for work. Now, if this person is given an additional choice of working 40 hours per week, it is reasonable to assume that this new alternative is a much closer substitute for the latter option rather than the former one. As a result of introducing the new option, while it is still more likely that the person chooses one of the work alternatives, the probability of choosing 30 hours would more likely be affected by the new option than the probability of not working. That is, the relative preference between the two original alternatives would not stay the same.

This assumption could be relaxed if we allow individuals' unobserved preferences to play a role in the utility function. One common method is through random coefficients where individual preferences for each utility component can deviate from the population average. In this setup, the randomness in preferences is combined with the independent error terms in the utility function which makes the residuals across alternatives correlated through the unobserved heterogeneity component. In other words, by doing so, the person's taste (or distaste) for work is reflected when evaluating the available alternatives.

Theoretically, when unobserved heterogeneity is not correctly incorporated into the model, the model is misspecified and the estimates are biased and inconsistent. Several parametric and semiparametric models are suggested to account for unobserved preferences through random coefficients and each method has theoretical and computational complexities.

The most common parametric approach developed by McFadden and Train (2000), assumes a continuous distribution of the unobserved heterogeneity of the parameters. This distribution is commonly assumed to be normal, and the researcher estimates the means and covariance matrix of the distribution. This method allows for a continuous range of unobserved heterogeneity, the estimation process is smooth, and estimates are more stable and less sensitive to the initial values used to start the optimisation process. However, the normality assumption is restrictive and not based on any real information. Therefore, the model could still suffer from misspecification.

On the other hand, semiparametric methods remove the normality assumption by estimating the probability distributions. For example, in the latent class (LC) approach, researchers arbitrarily choose a certain number of types and estimate the type coefficients and probabilities corresponding to each type. In this method, unobserved heterogeneity is limited to a few types and allowing for more types could raise identification problems not to mention that in this approach results could be more sensitive to the initial values. In more flexible variations of this method (Bajari et al. (2007); Fox et al. (2011)), the researchers arbitrarily choose the values of discrete mass points and estimate the probability density of these points. In this method, selecting appropriate mass points is a challenge, and convergence becomes increasingly difficult with a higher number of mass points. In addition, Fosgerau and Bierlaire (2007) proposed a method to approximate any continuous distribution and Train (2016) proposed the Logit-Mixed-Logit (LML) which is the generalised method of all and is claimed to be computationally convenient.[1] In all the methods, the estimation process becomes significantly more complex and time-consuming as the number of alternatives as well as the number of random coefficients increase, especially when correlation among these coefficients is also allowed.

Most applications of discrete choice models in labour supply use a random coefficient specification, such as mixed logit, to represent preference heterogeneity. Two specifications have been commonly used. The most common approach has been adopting a parametric approach assuming normal distribution and the other one is using latent class (LC) where the researcher uses a specific number of types. However, little is known about the ability of these models to capture the heterogeneity and the accuracy of model outcomes in these methods. Given the limitations and complexity of these methods, we aim to determine whether accounting for unobserved heterogeneity significantly enhances the outcomes of the labour supply model. This study evaluates their effectiveness and performance under various distributional assumptions, including differences in shape and variance of the underlying distribution. By understanding how these unobserved characteristics affect the outcomes, we can determine their significance and identify the situations in which they become particularly important. Answering this question could shed some light on the importance of this issue, especially within the structural labour supply model where the focus is not necessarily on the estimated coefficients themselves but rather on their ability to conduct *ex-ante* policy evaluations.

In this study, we use a Monte Carlo simulation to further investigate the role of unobserved

[1] Interested researcher can check Bansal et al. (2018a) and Bansal et al. (2018b) for a review of these methods.

heterogeneity in preferences through random coefficients in the structural labour supply context. Given the focus of our study on unknown heterogeneity, the Monte Carlo simulation is a suitable approach since it enables us to simulate a precisely defined and controlled environment for evaluating the desired assumptions. We generate data by assuming three types of unobserved heterogeneity distributions: normal, bimodal and discrete distributions, each with small and large variances. That is six scenarios. In this framework, we examine the performance of the four common model specifications in each scenario. The four specifications are Standard Conditional Logit (SCL), parametric with normal distribution (MCLN), the semi-parametric approach or latent class (LC) and a more flexible approach (LCN) that combines the types from LC and continuous coverage from MCLN. Given the assumptions of this study, the results show that the variance of unobserved heterogeneity significantly influences the model outcomes and the chosen specification. When the variance is small, there is less sensitivity to the choice of specification, and both the MCLN and LC specifications perform well. However, in scenarios with large variance, it is important not only to account for heterogeneity in estimation but also to consider the distribution's shape when selecting the appropriate approach. For distributions with discrete types and limited variations within each type, the LC approach is preferred. Conversely, in distributions with substantial overlap between types (ideally approaching a normal distribution), the MCLN approach produces better results. Given the unknown shape of the underlying distribution, more flexible approaches like LCN may be more suitable in such scenarios, despite the potential complexity trade-offs.

The remaining paper is organised as follows: Section 2 briefly discusses the studies that evaluate the performance of some of these methods in labour supply field. Section 3 presents the Monte Carlo plan in this study which includes the data generation process and the estimation method in each specification. Section 4 compares the performance of these specifications. Section 5 concludes.

2.2 Previous Studies

A random coefficient model is a widely used method in industrial organisation, health, marketing and labour economics to limit the extent of the IIA property. Several parametric (McFadden and Train (2000)) and semi-parametric methods (Bajari et al. (2007); Fosgerau and Bierlaire (2007);

Train (2008); Fox et al. (2011); Bastin et al. (2010); Fosgerau and Mabit (2013)) are suggested by researchers. While each method offers advantages and disadvantages, there is no clear agreement on which one is better. More importantly, only a few of these methods were adopted in structural labour supply studies because of the additional complexities involved in this model. For example, Bansal et al. (2018a) in a Monte Carlo simulation study use two explanatory variables to show that the Logit-Mixed-Logit (LML) by Train (2013) performs well in mimicking different shapes of the underlying unobserved heterogeneity distributions. The advantage of this method is its flexibility in generalising many of the other proposed methods and its computational convenience since the likelihood equation does not require the computation of choice probabilities in iterative optimisation, but this requires assuming random coefficients for all the parameters. However, in another study, the authors (Bansal et al. (2018b)) show that the computational time increases substantially, by a factor of 40, when there is a combination of fixed and random parameters in the reduced regression. Consequently, these findings suggest that implementing this method in a structural labour supply framework would be highly time-consuming due to the complexity of these models. For instance, studies (van Soest (1995); Löffler et al. (2018)) indicate that including fixed costs in the model and simultaneously estimating the wage equation alongside the utility model significantly enhance model performance. However, these additions quickly escalate the model's complexity, making the advantageous features of a flexible model like the LML model cumbersome and disadvantageous.

The most common method in structural labour supply to deal with unobserved heterogeneity assumes that some of the taste parameters in the utility function are normally distributed. This method estimates the mean and the parameters in the variance-covariance matrix (ex. van Soest (1995); Van Soest and Das (2001); Soest et al. (2002); Choi (2018); Clavet et al. (2013)). On the other hand, a few papers used the latent class method (Wrohlich (2011); Hansen and Liu (2015)). However, it is not quite clear how these methods perform and whether the results are different from a standard model with fixed parameters. As mentioned earlier, more flexible approaches are not common in this field, due to the complexities that they introduce to the model, outweighing the flexibility they provide, thereby making them less appealing methods in this domain.

Keane and Wasi (2013) compare several methods using empirical datasets from marketing and health, finding that more flexible approaches for incorporating heterogeneity outperform common methods based on normality assumptions or latent class models. These flexible methods are better at capturing extreme and random behaviours when present in the data. Therefore, the preferred method ultimately depends on the context and data being analysed. To our knowledge, only a few studies have investigated the role of unobserved heterogeneity on the outcomes of the structural labour supply model. Haan (2006) compares the performance of a discrete structural labour supply model using labour supply elasticities with different specifications for unobserved heterogeneity. He uses the German Socio-Economic Panel (SOEP) data and assumes that unobserved heterogeneity only exists for the income variable. He then uses the parametric with normal unobserved heterogeneity distribution and a non-parametric specification with two latent classes to account for preference heterogeneity in income. His results show that there is not any significant difference between the elasticities from the standard conditional logit model (without unobserved heterogeneity) and the two mixture models. In a similar study, Pacifico (2013) uses the European panel of income and living conditions (EU-SILC) with a similar setup to Haan's study and confirms his findings. However, his study goes a step further and shows that labour supply elasticities are, in fact, significantly different when unobserved heterogeneity is allowed for all the preference parameters in the model instead of only one parameter.

2.3 Monte Carlo Simulation Plan

In this study, we use the Monte Carlo simulation method to further investigate the role of unobserved heterogeneity in model outcomes in the labour supply context. The advantage of this method compared to the previous studies, is that the real population parameters and characteristics are known which allows us to compare the model's outcome to the true values. Also, we can generate scenarios under different assumptions which enables us to investigate the effect of each of those assumptions. Neither is possible using real datasets. Moreover, in using real data even if we can assume that the model specification is correct, the results could be simply sensitive to the sample used.

2.3.1 Labour Supply Model

We use the static discrete choice labour supply model that is based on a direct random utility function to generate labour supply distributions with different assumptions. We assume that individuals get utility from combinations of disposable income (C) and leisure (L), measured by the deterministic part of the utility function, $V(C, L)$. Also, their utility is affected by a random component, ϵ.

For the deterministic part, each individual faces a budget set from which they choose a combination of C and L that maximises their utility. We assume that the utility function is a log transformation of the Cobb-Douglas function with an order of two.[2] Similar variations of this function are commonly used in the literature (among all, van Soest (1995); Keane and Moffitt (1998); Van Soest and Das (2001); Soest et al. (2002); Flood et al. (2004); Haan (2006); Choi (2018); Pacifico (2013); Clavet et al. (2013); Hansen and Liu (2015)).

$$U(C_j, L_j) = \beta_c log(C_j) + \beta_l log(L_j) + \beta_{csq}(logC_j)^2 + \beta_{lsq}(logL_j)^2 + \beta_{cl} log(C_j) log(L_j) + \epsilon_j, \quad j = 1, ..., J = 16 \quad (8)$$

where C_j and L_j are disposable income and leisure time for individuals from alternative j.

We further assume that the annual time endowment available to each individual is $T = 4,000$ hours that can be allocated between work (h_j) and leisure (L_j), where $L_j = T - h_j$. The annual hours of work start from zero and go up to 3,000 hours in increments of 200 hours (ie, h_j = 0, 200, 400,..., 3000). This approximates the continuous distribution of work hours in real data. Then, the disposable income corresponding to each h_j is calculated using a detailed taxes and credits schedule (ie, each individual faces $j = 16$ alternatives or combinations of (C_j, L_j) from which to choose.)

Disposable income (C_j), or the budget constraint is calculated as

$$C_j = wh_j + y - t(wh_j + y_j^T - D_j) + TC_j + d \cdot SA_j, \quad (9)$$

[2] van Soest et al. (2002) used a flexible polynomial utility function, exploring model performance for orders 1 to 5, and found that outcomes at order 2 closely matched higher orders. We adopt the same assumption here.

where,

w: Before-tax hourly wage rate

h_j: Annual hours of work corresponding to alternative j

y: Annual non-labour income extracted from the Census data

t: Federal and provincial income tax schedule

D_j and TC_j: Federal and provincial deductibles and tax credits, respectively

SA_j: Social assistance benefits and $d = 1$ if eligible to receive the benefits, zero otherwise.

In order to estimate disposable income based on the number of hours worked, we use the hourly wage distribution and non-labour income, along with a detailed tax and credit schedule. The natural logarithm of the wage distribution follows a normal distribution with a mean of \$21.2 and a standard deviation of \$10.1, as observed in real data. Additionally, the tax and credits schedule is created using CTaCS[3] parameters. The average estimated disposable income for each hour category is presented in Table B.1 in the appendix. The generated disposable income is incorporated into the model exogenously based on the hour category.

Following van Soest (1995), we assume that there is a fixed cost of working. This can be interpreted as transportation or any work-related cost that occurs from working. In empirical studies, the prediction of the proportion of non-workers improves by including this term; otherwise, the proportion would be underpredicted. The fixed cost enters the model as $log(C_j) - log(fc)$, where $j > 1$.

Moreover, we assume that the utility is also affected by a stochastic component, ϵ_j. This counts for idiosyncratic preferences for certain jobs/alternatives or an error in individuals' evaluation of each alternative. We assume that ϵ_j is drawn from a type I extreme value distribution. This is a common assumption in discrete choice models as it provides a much easier and faster estimation process.

Furthermore, we include individual-specific preferences for disposable income and leisure by assuming heterogeneity through β_c and β_l. The next section describes this in more detail.

[3] Milligan (2016)

2.3.2 Data Generation

We generate datasets assuming that the preference parameters for leisure and disposable income are distributed normally and bimodally. The bimodal distributions consist of two normal distributions with partially overlapping and discrete distributions. We divide the proportion of each distribution unequally, 30 percent and 70 percent, to mimic an asymmetric distributional shape while also incorporating bimodality. Therefore, we have three scenarios: a normal distribution, bimodal distributions, and a discrete bimodal distribution. For each scenario, we consider low and high variances for the entire distribution, resulting in six scenarios. We generate 100 datasets for each scenario by taking random draws from the specified distributions. The sample size for each dataset is 2110. Characteristics of the distributions of unobserved preference heterogeneity are as follows:

a. Normal distributions with small variances (scenario a)

$$\begin{bmatrix} \beta_l \\ \beta_c \end{bmatrix} \sim BVN(\begin{bmatrix} b_l = 22.0 \\ b_c = 15.5 \end{bmatrix}, \Omega = \begin{bmatrix} 2.250 & -1.905 \\ -1.905 & 2.250 \end{bmatrix}), \rho = -0.85$$

b. Bimodal distributions with small variances (scenario b)

$$\begin{aligned} &\text{with probability } 0.7: \quad \beta_c \sim N(14, 0.7),\ \beta_l \sim N(17, 0.7) \\ &\text{with probability } 0.3: \quad \beta_c \sim N(11, 0.7),\ \beta_l \sim N(20, 0.7) \\ &\text{ie:} \quad \begin{bmatrix} \mu_{\beta_l} = 17.90 \\ \mu_{\beta_c} = 13.09 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 2.38 & -2.01 \\ -2.01 & 2.38 \end{bmatrix}, \quad \rho = -0.85 \end{aligned}$$

c. Discrete distributions with small variances (scenario c)

$$\text{with probability } 0.7: \quad \beta_c \sim N(14, 0.2),\ \beta_l \sim N(17, 0.2)$$

$$\text{with probability } 0.3: \quad \beta_c \sim N(11, 0.2),\ \beta_l \sim N(20, 0.2)$$

$$\text{ie:} \quad \begin{bmatrix} \mu_{\beta_l} = 17.90 \\ \mu_{\beta_c} = 13.09 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 1.94 & -1.42 \\ -1.42 & 1.93 \end{bmatrix}, \quad \rho = -0.75$$

aa. Normal distributions with large variances (scenario aa)

$$\begin{bmatrix} \beta_l \\ \beta_c \end{bmatrix} \sim BVN(\begin{bmatrix} b_l = 36 \\ b_c = 29 \end{bmatrix}, \Omega = \begin{bmatrix} 49 & -42 \\ -42 & 49.7 \end{bmatrix}), \rho = -0.85$$

bb. Bimodal distributions with large variances (scenario bb)

$$\text{with probability } 0.7: \quad \beta_c \sim N(29, 3),\ \beta_l \sim N(33, 3)$$

$$\text{with probability } 0.3: \quad \beta_c \sim N(15, 3),\ \beta_l \sim N(47, 3)$$

$$\text{ie:} \quad \begin{bmatrix} \mu_{\beta_l} = 37.22 \\ \mu_{\beta_c} = 24.77 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 50.23 & -42.70 \\ -42.70 & 50.11 \end{bmatrix}, \quad \rho = -0.85$$

cc. Discrete distributions with large variances (scenario cc)

$$\text{with probability } 0.7: \quad \beta_c \sim N(29, 0.85),\ \beta_l \sim N(33, 0.85)$$

$$\text{with probability } 0.3: \quad \beta_c \sim N(15, 0.85),\ \beta_l \sim N(47, 0.85)$$

$$\text{ie:} \quad \begin{bmatrix} \mu_{\beta_l} = 37.22 \\ \mu_{\beta_c} = 24.78 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 42.00 & -31.16 \\ -31.16 & 41.93 \end{bmatrix}, \quad \rho = -0.74$$

We assume that β_c and β_l are negatively correlated. In other words, higher utility from disposable income corresponds to lower utility from leisure.

Other utility parameters of the population are slightly adjusted to generate the desired hour distribution described above. Table 2.1 shows the parameters used to generate the data. Unlike reduced-form regressions, most of the structural labour supply model parameters are not informative by themselves. We set the parameters such that (1) the generated datasets satisfy the theoretical assumptions of labour supply (ie, the marginal utility of disposable income and leisure are positive, and the utility function is concave), and (2) the generated hour distributions resemble what we anticipate observing in actual data. (3) the generated hour distributions from each scenario are similar to facilitate easier comparison of model predictions[4]. Furthermore, we assume that β_{cl} is positive, which means leisure and disposable income are substitutes.

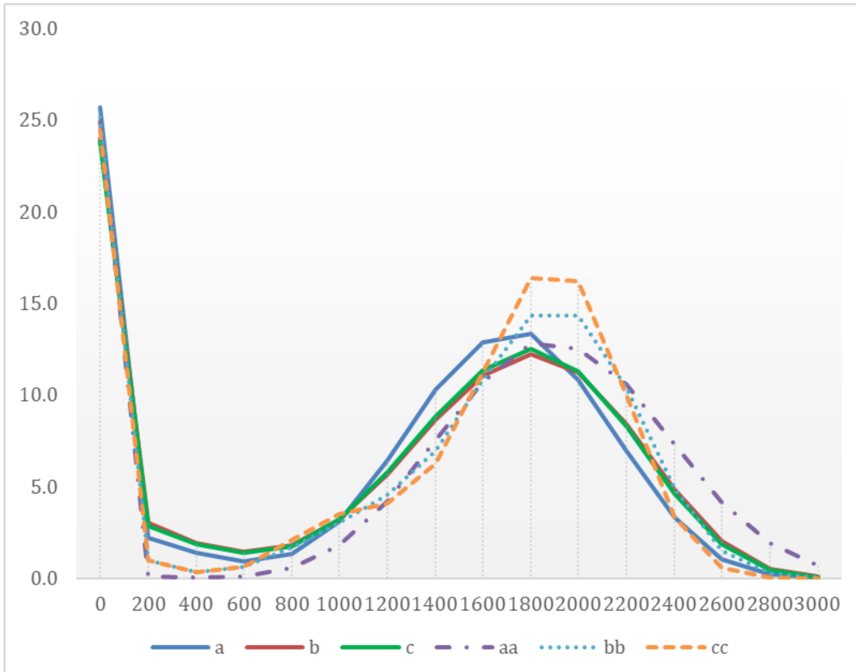
Table 2.1: Population Parameters for the Overall Distributions

	Small Variance			Large Variance		
Population Parameters	a.Normal	b.Bimodal	c.Discrete	aa.Normal	bb.Bimodal	cc.Discrete
β_{csq}	2.87	2.60	2.60	1.87	1.65	1.65
β_{cl}	0.35	0.33	0.33	0.42	0.42	0.42
$\beta_{l_{sq}}$	-2.15	-2.10	-2.10	-1.95	-1.95	-1.95
fc	2.25	2.40	2.40	2.30	2.00	2.00
β_l	22.00	17.90	17.90	36.00	37.22	37.22
β_c	15.50	13.09	13.09	29.00	24.77	24.78
std(β_l)	1.50	1.54	1.39	7.00	7.09	6.48
std(β_c)	1.50	1.54	1.39	7.05	7.08	6.48
cov(β_c, β_l)	-1.91	-2.00	-1.42	-42.00	-42.70	-31.11
corr(β_c, β_l)	-0.85	-0.85	-0.75	-0.85	-0.85	-0.74
Observations			2110			

Figure 2.1 (and Tables B.8-B.10 in the appendix) shows the generated-hour distributions in the six scenarios. The distributions generated in these scenarios share many general characteristics. The non-employment rate is between 24 and 26 percent, and the hours of work peak around full-time hours, which is between 1600-2200 annual worked hours. This distribution pattern resembles the labour supply distribution of women with lower education levels resulting in varied hour distribution with larger elasticities. This makes the analysis and comparison of model outcomes for different

[4]For the small variance scenarios, we also used coefficients similar to those with large variances. This slightly altered the generated hour distribution but did not affect the results of the study.

Figure 2.1: The Generated Hour Distribution



specifications clearer and more meaningful, which would not be possible with very small or near-zero elasticities.

2.3.3 Estimation

Standard Conditional Logit (SCL)

Consider the following utility function where a representative decision maker chooses a choice set (C_j, L_j) that maximises utility U_j.

$$U_j = V_j(C_j, L_j) + \epsilon_j, \quad j = 1, ..., J \tag{10}$$

V_j is the deterministic part of the utility function as described in equation (8). J is the number of alternatives which is equal to 16. ϵ_j is assumed to be *iid* and follows a type I extreme value distribution, which leads to a logistic distribution of differences in the error terms between any two alternatives with $(0, \sigma_\epsilon^2)$. The logistic probability density function of alternative j for each individual is

$$p_j = Pr(h_j) = \frac{exp(V_j)}{\sum_j exp(V_j)}. \tag{11}$$

Then, the likelihood function for standard conditional logit (SCL) is

$$L = \sum_j \delta_j \cdot p_j, \qquad j = 1, ..., J \tag{12}$$

where δ_j is equal to 1 for the observed state of j and 0 otherwise.

The parameters to estimate in this specification are $\theta = (\beta_{csq}, \beta_{lsq}, \beta_{cl}, \beta_c, \beta_l, fc)$. These parameters are estimated using the Maximum Likelihood Estimation method.

Mixed Conditional Logit - Normal Unobserved Heterogeneity Distribution (MCLN)

The standard conditional logit ignores preference heterogeneity and assumes that the error terms across alternatives are uncorrelated while in this setup, these error terms are correlated through unobserved preference heterogeneity. By ignoring them, as in SCL, unobserved heterogeneity remains in the error term and, therefore, the utility error terms are correlated through the common unobserved heterogeneity.

The most common parametric approach to include unobserved heterogeneity in this framework is to assume a bivariate normal (BVN) distribution for selected preference parameters. We assume that

$$\begin{cases} \hat{\beta}_l = b_l + e_l \\ \hat{\beta}_c = b_c + e_c, \end{cases} \tag{13}$$

where $\begin{bmatrix} \hat{\beta}_l \\ \hat{\beta}_c \end{bmatrix} \sim BVN\left(\begin{bmatrix} b_l \\ b_c \end{bmatrix}, \Omega \right)$.

The choice probabilities are calculated by a given set of (e_c, e_l) where this condition is integrated out using the BVN probability density function $\varphi(e_c, e_l)$. The choice probability for set j is

$$p_j = \iint_{-\infty}^{+\infty} Pr(h_j|e_c, e_l)\cdot\varphi(e_c, e_l)\, de_c\, de_l = \iint_{-\infty}^{+\infty} \frac{exp(V_j)}{\sum_j exp(V_j)}\cdot\varphi(e_c, e_l)\, de_c\, de_l, \quad j = 1, ..., J. \tag{14}$$

The integration is approximated by taking several random draws from the BVN density function, $\varphi(e_c^r, e_l^r)$ where r is the r^{th} draw from this distribution. Therefore, the simulated choice probability is

$$\hat{p}_j = \frac{1}{R}\sum_{r=1}^{R} Pr(h_j|e_c^r, e_l^r) = \frac{1}{R}\sum_{r=1}^{R} \frac{exp(V_j|e_c^r, e_l^r)}{\sum_j exp(V_j|e_c^r, e_l^r)}, \qquad r = 1, ..., R. \tag{15}$$

We take 100 draws from a Halton sequence[5][6]. Finally, we substitute $\hat{p}_j$ in (12) and that gives us the simulated likelihood function (SL)

$$SL = \sum_j \delta_j \cdot \hat{p}_j, \qquad j = 1, ..., J. \tag{16}$$

The parameters to estimate in this specification are $\theta = (\beta_{csq}, \beta_{lsq}, \beta_{cl}, b_c, b_l, fc, \Omega)$. These parameters are estimated using the Simulated Maximum Likelihood Estimation method.

Mixed Conditional Logit - Latent Class (LC)

In this approach, unobserved heterogeneity is integrated by allowing discrete types and estimating the types and their probabilities. This method does not require imposing a parametric structure on the probability density function of unobserved heterogeneity. In this study, we incorporate two points for both β_l and β_c, and allowing for the consideration of all four combinations of these points. This yields four distinct types $(S = 4)$ of individuals. Therefore, the probability of each alternative in equation (4) is calculated given types (s),

[5] Train (2009) explains that two issues matter in an accurate approximation of an integral: good coverage of the domain and negative covariance both across draws for each observation and across observations. Negative covariance across draws leads to better coverage and also reduces the variance of the simulator. Also, negative covariance across observations reduces the variance of the SML maximand that results in a more efficient estimator. He explains that, unlike similar methods, Halton draws satisfy all these characteristics. There are several empirical studies that show the superiority of the results based on Halton draws. These studies show that 100 Halton draws provide more precise results than 1000 pseudo random draws.

[6] We have tried 50, 100 and 200 draws and estimates are consistent.

$$p_{j|s} = Pr(h_j|s) = \frac{exp(V_j|s)}{\sum_j exp(V_j|s)}, \tag{17}$$

where $S = \{\beta_{l,m}, \beta_{c,k}\}$ for $k = 1, 2$ and $m = 1, 2$. Then, the likelihood function in (12) is adjusted to

$$L = \sum_s \pi_s(\sum_j \delta_j p_{j|s}), \quad j = 1, ..., J; s = 1, ..., S. \tag{18}$$

where π_s is the probability of type and is estimated using a logistic function. The estimated parameters in this approach include the fixed utility parameters, parameters for the probability types and heterogeneity points. That is, $\theta = (\beta_{csq}, \beta_{lsq}, \beta_{cl}, fc, \pi_s(i.e.pr_1, pr_2, pr_3), b_{c1}, b_{l1}, b_{c2}, b_{l2})$. These parameters are estimated using the Maximum Likelihood Estimation method.

Mixed Conditional Logit - Latent Class and Normal Distribution (LCN)

This particular approach combines elements from the previous two approaches. It operates under the assumption that there are multiple types, each of which follows a normal distribution. This approach is quite flexible as it allows for a continuous range of values for unobserved characteristics and can capture underlying multimodal distributions.

Equations (14) and (15) are computed for each type. Therefore, the simulated probability of choosing alternative j given type (s) is

$$p_{j|s} = \frac{1}{R}\sum_{r=1}^{R} Pr(h_j|e_c^r, e_l^r, s) = \frac{1}{R}\sum_{r=1}^{R} \frac{exp(V_j|e_c^r, e_l^r, s)}{\sum_j exp(V_j|e_c^r, e_l^r, s)}, \quad r = 1, ..., R. \tag{19}$$

Then, we use the simulated probability in (19) to define the simulated likelihood function similar to equation (18) to include the types.

$$SL = \sum_s \pi_s(\sum_j \delta_j \hat{p}_{j|s}), \quad j = 1, ..., J; s = 1, ..., S. \tag{20}$$

π_s is probability of type.

In this approach, we assume that we have two types $(S = 2)$. Therefore, the estimated parameters are the fixed utility parameters, parameters for type probabilities (pr_s) and parameters of normal distribution for each type (i.e. $\theta = (\beta_{csq}, \beta_{lsq}, \beta_{cl}, fc, \pi_s(i.e.pr_1), b_{cs}, b_{ls}, \Omega_s)$). These parameters are estimated using the Simulated Maximum Likelihood Estimation method.

2.4 Analysis of Results

The evaluation of the performance of four specifications (SCL, MCLN, LC, and LCN) is done using the standard performance metrics of Monte Carlo studies such as percentage bias of the estimates and their mean squared errors (MSE), loglikelihood function value and Bayesian Information Criterion (BIC) (section 2.4.1). As mentioned earlier, the greatest attraction of structural labour supply models is their application in the ex-ante analysis of policy changes. Therefore, we also assess how these specifications perform in predicting the hour distribution (section 2.4.2) and the labour supply response to a policy change (section 2.4.3). The formula for these metrics is detailed in Appendix A.

2.4.1 Model Estimates

Tables (2.2-2.7) compare the true and estimated parameters of variables with heterogeneous coefficients (i.e. the natural logarithm of consumption and leisure in the utility function) across four specifications (SCL, MCLN, LC, and LCN). In this table, columns 'Mean', 'Bias', and 'MSE' are the average estimate from 100 datasets, the corresponding bias in percentage, and mean squared errors. The mean log-likelihood and Bayesian Information Criteria (BIC) are also reported at the bottom of these tables. The detailed parameter estimates are in Tables (B.2-B.7) in the appendix.

Table 2.2: Estimates (DGP: (a) Normal Distribution - Small Variance)

	True	Means			Bias%			MSE		
	Parameter	SCL	MCLN	LC	SCL	MCLN	LC	SCL	MCLN	LC
β_l	22	17.25 (1.20)	21.87 (2.16)	22.07 (2.25)	-21.6	-0.6	0.3	23.99	4.65	5.02
β_c	15.5	11.96 (0.42)	15.63 (0.92)	15.78 (1.00)	-22.8	0.8	1.8	12.70	0.86	1.06
Loglikelihood		-3976	-3952	-3954						
Estimated Parameters		6	9	11						
BIC		7997	7973	7993						

Values in parentheses are standard deviations.

Table 2.3: Estimates (DGP: (b) Bimodal Distribution - Small Variance)

	True Parameter	Means			Bias%			MSE		
		SCL	MCLN	LC	SCL	MCLN	LC	SCL	MCLN	LC
β_l	17.90	13.80 (0.82)	18.02 (1.68)	18.12 (1.82)	-22.9	0.7	1.2	17.53	2.82	3.33
β_c	13.09	9.78 (0.43)	13.15 (0.78)	13.32 (0.92)	-25.3	0.5	1.7	11.16	0.60	0.89
Loglikelihood		-4361	-4333	-4329						
Estimated Parameters		6	9	11						
BIC		8768	8734	8741						

Values in parentheses are standard deviations.

Table 2.4: Estimates (DGP: (c) Discrete Bimodal Distribution - Small Variance)

	True Parameter	Means			Bias%			MSE		
		SCL	MCLN	LC	SCL	MCLN	LC	SCL	MCLN	LC
β_l	17.90	14.80 (0.84)	18.36 (1.62)	18.21 (1.55)	-17.3	2.5	1.7	10.32	2.81	2.49
β_c	13.09	10.36 (0.42)	13.22 (0.75)	13.33 (0.78)	-20.8	1.0	1.8	7.61	0.57	0.66
Loglikelihood		-4299	-4276	-4272						
Estimated Parameters		6	9	11						
BIC		8644	8622	8628						

Values in parentheses are standard deviations.

Tables (2.2-2.4) present the estimated unobserved heterogeneity parameters for distributions with small variances. The results indicate that SCL shows the poorest performance in these three scenarios compared to MCLN and LC. Both the percentage bias and the MSE measures from these two specifications are significantly lower than those from SCL. For example, biases in MCLN and LC are up to about 2 percent in the three scenarios, whereas in SCL they range between 17 and 26 percent. In these three scenarios, biases and mean squared errors from both MCLN and LC are small, with a slight preference for MCLN.

Tables (2.5-2.7) present the estimated parameters for the three scenarios with large variances. Similar to the previous findings, SCL has the largest biases in all three scenarios, with biases ranging from 50 and 90 percent, which is more than double those of the second-worse specification in each scenario. Furthermore, comparing the results from the SCL specification to Tables (2.2-2.4), it is clear that the biases, and consequently, MSEs are much larger when the unobserved heterogeneity distributions have larger variances, indicating that SCL does not capture the variation in preferences.

Next, we compare the performance of MCLN and LC in scenarios (aa-cc) of Table (2.5-2.7). We observe that MCLN performs best in scenario (aa) with normal unobserved heterogeneity. This is expected, as the underlying distribution aligns with the specification of the model. However, the biases and MSEs from this method increase in scenarios (bb) and (cc) with bimodal distributions, becoming even larger as the distribution becomes more discrete (i.e., bb versus cc). On the other hand, biases and MSEs, in the LC method is the largest in scenario (aa) with normal distribution and smallest in scenario (cc) with discrete bimodal distribution. This could be related to the bimodality of the underlying distributions with distinct types in this scenario, which is better captured in the LC specification.

An additional observation involves comparing the LC approach in scenario (cc), where each sub-normal distribution has a standard deviation of 0.85, and in scenario (bb), where each sub-normal distribution has a standard deviation of 3. The biases from the LC approach in (cc) are only 1.2 and 3.3 percent, whereas in scenario (bb), they are 13.3 and 12.9 percent. Additionally, in Table B.7, we can see that each point estimate closely approximates the modes of the underlying bimodal distributions in scenario (cc), but the biases are larger for scenario (bb). This shows that the two mass points used for each of consumption and leisure preferences in scenario (cc) with small variances, accurately capture the bimodality of the underlying distribution with small biases. However, in scenario (bb) with larger variances, a single-point estimate for each sub-distribution is not as effective. This further emphasises that the size of the variance of unobserved heterogeneity plays a crucial role in obtaining accurate estimates.

Table 2.5: Estimates (DGP: (aa) Normal Distribution - Large Variance)

	True Parameter	Means				Bias%				MSE			
		SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN
β_l	36.00	9.50 (0.55)	35.90 (4.27)	27.54 (4.81)	35.37 (4.35)	-73.6	-0.3	23.5	-1.8	702.41	18.09	94.44	19.16
β_c	29.00	3.44 (0.38)	29.05 (2.99)	24.67 (4.96)	29.20 (3.16)	-88.1	0.2	-14.9	0.7	653.46	8.84	43.13	9.90
Loglikelihood		-4361	-4107	-4158	-4103								
Estimated Parameters		6	9	11	15								
BIC		8768	8282	8401	8321								

Values in parentheses are standard deviations.

Table 2.6: Estimates (DGP: (bb) Bimodal Distribution - Large Variance)

	True	Means				Bias%				MSE			
	Parameter	SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN
β_l	37.22	14.38 (0.75)	44.41 (4.45)	32.27 (4.70)	36.67 (4.05)	-61.4	19.3	-13.3	-1.5	521.44	71.32	46.34	16.54
β_c	24.77	2.53 (0.39)	31.09 (2.62)	21.58 (3.62)	24.86 (3.08)	-89.8	25.5	-12.9	0.4	496.16	46.74	23.09	9.39
Loglikelihood		-4366	-4239	-4203	-4189								
Estimated Parameters		6	9	11	15								
BIC		8777	8548	8489	8492								

Values in parentheses are standard deviations.

Table 2.7: Estimates (DGP: (cc) Discrete Bimodal Distribution - Large Variance)

	True	Means				Bias%				MSE			
	Parameter	SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN	SCL	MCLN	LC	LCN
β_l	37.22	17.52 (0.81)	48.11 (4.84)	36.78 (2.49)	36.93 (2.91)	-52.9	29.2	-1.2	-0.8	388.83	1466.9	6.33	8.50
β_c	24.78	3.25 (0.42)	35.15 (2.76)	23.97 (2.09)	25.08 (2.76)	-86.9	41.8	-3.3	1.2	463.93	782.71	4.98	7.64
Loglikelihood		-4247	-4099	-3986	-3988								
Estimated Parameters		6	9	11	15								
BIC		8540	8267	8056	8091								

Values in parentheses are standard deviations.

The current findings indicate that the choice between the LC and MCLN specifications is important when there is considerable variation in unobserved heterogeneity, and it relies on the shape of the underlying distribution. However, the true shape of the underlying distribution is unknown. Therefore, we next apply the LCN specification for scenarios (aa-cc). This more flexible approach can mimic both bimodality and coverage, as it accounts for both discrete types and continuous variation in preference heterogeneity within each type. The findings, in Tables (2.5-2.7), show that the estimates generated by LCN have minimal biases (less than 2 percent) and mean squared errors (MSEs) across all three scenarios, which makes this a preferred method under these circumstances. Also, biases and MSEs show significant improvements in scenario (bb) compared to the MCLN and LC approaches. In scenario (aa) biases and MSEs are comparable to those of MCLN, while in scenario (cc) they are similar to those of LC.

These tables also present loglikelihood and BIC values. In all six scenarios, there is a significant improvement in loglikelihood values observed in MCLN, LC, and LCN specifications compared to SCL. The BIC values from MCLN, LC, and LCN are lower than those from SCL, suggesting that the additional parameters used in these specifications contribute to a better model fit. Furthermore, among the scenarios with small variances, the BIC value from MCLN is consistently the lowest. In scenarios with large variances, the BIC from MCLN in the (aa) scenario with a normal distribution is the smallest, as expected, while the BIC from LC in the (bb) and (cc) scenarios with bimodal distributions is the smallest. In these scenarios, the BIC values from the LCN approach are also comparable to the preferred approaches above.

In summary, methods that incorporate unobserved heterogeneity (i.e. MCLN, LC and LCN) lead to significantly improved estimates of bias size, mean squared error (MSE), log-likelihood, and Bayesian Information Criterion (BIC) measures of model fit compared to the standard conditional logit (SCL). When the variance is small, the choice of specification is less sensitive to the underlying distribution type. Both the conditional logit with a normal distribution (MCLN) and the latent class (LC) models show similar performance, with MCLN performing slightly better. However, when the variance of the unobserved heterogeneity is large, the choice of specification becomes more sensitive to the shape of the distribution. If there are distinct types with less variations in each type, then LC is preferred; otherwise, more flexible approaches should be used. The estimates from the LCN specification indicate that it could be the preferred alternative due to the added advantage of accommodating both types and continuous coverage.

2.4.2 Model Fit - Predicting the Hour Distribution

In this section, we examine how well these specifications predict the hour distribution by using the chi-squared test on each dataset. Table 2.8 presents the number of datasets in each scenario that successfully pass the chi-squared test at a 5 percent significance level in each specification. This shows that the predicted hour distribution in these datasets closely matches the actual distribution. The benefit of this test is that it assesses each simulated dataset individually rather than based on an average bias[7]. Tables (B.8-B.10) in the appendix show the extended predicted hour distribution with

[7] It is worth noting that when dealing with a large number of hour categories, conducting the Chi-squared test can be more challenging. This is because the probability of misallocation decreases when there are only a few alternatives available. For instance, if the hour distribution is divided into numerous categories, some may have a small number of observations. Even a few misallocated observations can rapidly increase the chi-squared value due to this disparity. Conversely, categories with high density might not experience a significant change in the chi-squared value, even with larger misallocations. Therefore, it is better to use this method with caution along with other considerations when evaluating model fit.

mean proportions in each category, as well as the bias (in percentage points) for each specification in all six scenarios.

Table 2.8: Chi-squared Test for Predicted Hour Distribution

	Small Variance			Large Variance		
Specification	a.Normal	b.Bimodal	c.Discrete	aa.Normal	bb.Bimodal	cc.Discrete
SCL	79	27	53	50	0	0
MCLN	100	99	100	98	17	0
LC	95	89	81	43	81	80
LCN	-	-	-	97	82	91

At 5 percent significance level with 16 categories.

Table 2.8 shows that the SCL specification demonstrates the lowest number of datasets passing the chi-squared test compared to other specifications in each scenario. It is also sensitive to both the variance and shape of the underlying distribution. For instance, 70 percent of the datasets pass the test in scenario (a) with a normal distribution and small variance but only 50 percent pass when the variance is large, and none pass when distributions are bimodal with large variances. Furthermore, in scenarios with small variances both MCLN and LC perform well, with MCLN demonstrating better performance as close to 100 percent of datasets pass the test.

In comparing MCLN and LC in scenarios with large variances significant differences become apparent. In scenario (aa), characterised by a normal distribution, 98 percent of datasets pass the chi-squared test when utilising the MCLN specification, which is expected due to the matching distribution and model specification. In contrast, only 43 percent of datasets pass the test when using LC in this scenario. Conversely, in scenarios (bb) and (cc) with bimodal distributions, 81 and 80 percent of datasets, respectively pass the model fit test using LC, while only 17 and zero percent pass with the MCLN specification. In these three scenarios, 97 percent for normal distribution and 82 and 91 percent for bimodal distributions, of datasets pass the chi-squared test when using the LCN specification.

These results align with previous findings in the study and emphasise the importance of including unobserved heterogeneity in the estimation. Additionally, the choice of specification becomes more critical when dealing with large variances. The LCN specification is a more reliable approach for the hour distribution in datasets with large variations in unobserved heterogeneity.

2.4.3 Policy Evaluation - Labour Supply Response

After analysing the above results, it is clear that the SCL approach shows poor performance across all scenarios, especially those with large variances. Based on the current assumptions, it is evident that the MCLN and, to some extent, LC are suitable when variances are small. Additionally, when the variance is large, MCLN performs well when the distribution is normal, while LC is preferred when the distribution is bimodal. LCN performs well in all these scenarios. This section examines how these preferred specifications predict labour supply responses to policy changes.

We use the introduction of an EITC-like refundable tax credit programme (EITC[8]) as a policy change. We calculate the average employment rate change in percentage points and percentage change in intensive hours of work among workers. Table 2.9 summarises the results. The 'True' column shows the average responses from the generated data after implementing the policy change. The employment rate increases in response to implementing this programme as it involves a wage increase for non-workers. On the other hand, there is small overall reduction in hours of work in four of the six scenarios, which suggests that the overall income effect of the transfers dominates the substitution effect of the subsidised wages in the phase-in region of the programme. For two of the scenarios, there is a small increase in average worked hours of workers. These results coincide with labour supply theory and suggest that the generated datasets are reliable for this type of analysis.

Table 2.9 reports the mean predicted responses and the associated bias in percent across datasets for each selected specification, as well as the MSEs. An overall review of the results shows that the selected specifications presented here, produce small biases in predicting labour supply responses and effectively estimating them.

More specifically, results in Table 2.9 show that in scenarios with small variations in the unobserved heterogeneity, both MCLN and LC have small biases regardless of the shape of the unobserved heterogeneity. Comparing MCLN and LC, we see that MCLN tends to have smaller biases

[8]Individuals should have minimum $3500 earnings to be eligible. Then, the EITC payment is 20 percent of their earnings above $3500. The maximum annual payment is $1,300 where it decreases by 25 percent of their earnings exceeding $15,000. This policy has a structure similar to the actual WITB introduced in 2007, but it is slightly more generous.

Table 2.9: Labour Supply Response to an Introduction of an EITC-like Program

	Scenarios		True	Mean (Bias%) Selected Specification		MSE Selected Specification	
Small Variance				*MCLN*	*LC*	*MCLN*	*LC*
	a. Normal	Employment	2.72	2.70 (0.8)	2.32 (14.7)	0.08	0.40
		Hours*	-0.01	0.01 (0.0)	0.13 (0.1)	0.64	0.88
				MCLN	*LC*	*MCLN*	*LC*
	b. Bimodal	Employment	2.28	2.17 (5.0)	1.94 (14.9)	0.15	0.31
		Hours*	0.16	0.17 (0.0)	0.27 (0.1)	0.91	0.89
				MCLN	*LC*	*MCLN*	*LC*
	c. Discrete	Employment	2.30	2.19 (5.1)	2.10 (9.0)	0.16	0.26
		Hours*	0.20	0.16 (0.0)	0.18 (0.0)	0.76	0.61
Large Variance				*MCLN*	*LC*	*MCLN*	*LC*
	aa. Normal	Employment	1.79	1.74 (2.7)	1.67 (6.9)	0.08	0.15
		Hours	-1.05	-1.05 (0.3)	-1.00 (4.9)	0.04	0.04
				LC	*LCN*	*LC*	*LCN*
	bb. Bimodal	Employment	2.90	2.74 (5.4)	2.79 (3.7)	0.34	0.27
		Hours	-1.53	-1.59 (4.2)	-1.36 (10.8)	0.54	0.32
				LC	*LCN*	*LC*	*LCN*
	cc. Discrete	Employment	3.19	3.09 (3.0)	3.24 (1.7)	0.14	0.15
		Hours	-1.50	-1.24 (17.0)	-1.31 (12.4)	0.26	0.41

The table shows the mean values across 100 datasets. Values in parentheses are biases in percent
*In this case, the bias is measured in percentage points due to the very small base value.

and MSEs than the LC specification, however, the differences are marginal. This further confirms that under these assumptions the results are not sensitive to the shape of the unobserved heterogeneity.

Next, we observe scenarios with significant differences in unobserved heterogeneity. In scenario (aa) with a normal distribution, both MCLN and LCN perform well, showing minimal biases and MSEs. The biases from both methods are less than 7 percent and MSEs are between 0.04 and 0.15. As expected, results from MCLN are marginally superior. In scenarios (bb) and (cc) with bimodal distributions, the biases from both specifications are similar across most outcomes. However, when comparing the MSEs, the values from the LCN in scenario (bb) and LC in scenario (cc) are slightly more stable than the alternative method, making these methods the preferred choice at the margin.

2.5 Conclusion

The discrete choice structural labour supply model is well-known for its ability to use for ex-ante policy evaluations. However, the IIA assumption in this model is an overly simplifying assumption which assumes alternatives are uncorrelated. A common method to remove this assumption is to include unobserved preference heterogeneity through random coefficients. Several methods are suggested to deal with this issue in discrete choice modeling. Within the structural labour supply framework, two commonly used methods are the parametric specification with normal distribution and the semi-parametric latent class specification.

The normality assumption in parametric specification is rather ad hoc and lacks an empirical basis. On the other hand, the latent class assumes the existence of distinct types and estimates the coefficient and probability associated with each type. However, determining the number of types is done arbitrarily, and selecting too many types can lead to convergence and identification issues. While more flexible variations of semi-parametric methods have been proposed, their increased flexibility comes at the cost of an intense estimation process, particularly when applied to structural labour supply models, making them less appealing methods in these models.

In the labour supply literature, the impact of including unobserved preference heterogeneity in models remains uncertain. While ignoring unobserved heterogeneity can lead to theoretical model misspecification, it's unclear if accounting for unobserved heterogeneity significantly enhances model outcomes or if common specifications like MCLN and LC can effectively improve results despite methodological challenges. This study addresses these questions by evaluating how well prevalent methods perform under various unobserved heterogeneity distributional assumptions, including different shapes and variances, in estimating model outcomes.

In this study, we set up a Monte Carlo study in a structural labour supply model to evaluate the performance of the standard conditional logit (SCL), the parametric mixed conditional logit (MCLN) with a normal distribution and semi-parametric approach with latent classes (LC). We use the common framework of the discrete choice labour supply with the random utility model to generate six scenarios. We allow unobserved heterogeneity on consumption and leisure preference parameters where the unobserved heterogeneity has normal distribution, bimodal distribution and discrete distribution, each with small and large variances. To assess the performance of each specification, we compare the model estimates and Bayesian Information Criteria (BIC).

In all six scenarios, the results indicate that SCL underperforms in comparison to other specifications, particularly when the variance of unobserved heterogeneity is large. This suggests that using a single point to approximate a wide range of unobserved characteristics is ineffective and highlights the importance of integrating heterogeneity into the estimation process.

Analysis of estimated coefficients, mean squared errors (MSEs), loglikelihood values, and Bayesian Information Criterion (BIC) shows a significant improvement in model performance across all scenarios when incorporating unobserved heterogeneity in the estimation process. Particularly, when variance is low, MCLN performs marginally better than LC in all three scenarios in this category. This suggests that when the variance is small results are not very sensitive to the shape of the underlying distribution. But as variance increases, the choice of specification becomes more critical. MCLN tends to outperform with a normal underlying unobserved heterogeneity, while LC performs better with discrete unobserved heterogeneity.

As researchers are uncertain about the shape of the underlying distribution, we use the LCN approach, which accommodates both discrete types and the continuous scope of unobserved characteristics. The results from this approach show that LCN is effective in the three scenarios examined, with minimal biases and mean squared errors (MSEs) compared to other preferred approaches.

In this field, structural modeling is primarily used for ex-ante policy evaluation. Therefore, we proceeded to compare the performance of these specifications in predicting the hour distribution using a chi-squared test and estimating labour supply response to a policy change. The results of this analysis are consistent with previous findings. The chi-squared test indicates that the SCL exhibits inferior performance, particularly in scenarios with large variations in unobserved heterogeneity. However, the chi-squared test based on LCN and especially MCLN confirms a good model fit for almost all datasets with small variations in unobserved heterogeneity, regardless of the shape of the distribution. In scenarios with substantial variations in unobserved heterogeneity, MCLN, and

LCN for normal underlying distribution, and LC and LCN for bimodal underlying distributions, demonstrate significantly better results, passing the test for most datasets. We then assessed the labour supply response to the introduction of an EITC-like program for the preferred specifications, and the findings are consistent with those from previous analyses in this study.

In summary, the findings of this study indicate that MCLN can be a reliable approach when dealing with small variances, regardless of the shape of the underlying distribution. This can be easily confirmed by estimating the model with a simple MCLN specification and estimating the variances of unobserved heterogeneity. However, when dealing with large variances, it becomes more important to use specifications that appropriately account for unobserved heterogeneity. In such cases, it is important to investigate the underlying shape of the unobserved heterogeneity distribution and select the appropriate specification accordingly. One way to do this is by using the LC method to determine if there are distinct types based on the size of the estimated coefficients and the probability of the types. These two factors should provide a reasonable insight into the shape of the underlying distribution and help the researcher choose the most appropriate specification. If distinct types are identified, either the LC or a more flexible approach such as LCN should be used. If no distinct types are identified or if the estimated variances from the simple MCLN specification are small, then the MCLN approach should be sufficient.

To our knowledge, this study is the first to employ the Monte Carlo method to address these research questions. There are numerous avenues for extending this study to validate these findings and explore alternative options. This could involve generating data with different distributional assumptions or incorporating unobserved heterogeneity into other variables of the utility function. Additionally, future studies can explore the concept of variance size and what constitutes large or small variance. Such studies can provide insights into the importance of unknown unobserved heterogeneity and its implications, offering guidance on future directions. This could enhance the credibility of these models and increase their appeal among researchers and policymakers.

Chapter 3

The Joint Decision on Employment, Social Assistance Participation and Paid Childcare Utilisation for Single Mothers

3.1 Introduction

Many developed countries have significantly invested in providing low-cost childcare, with Scandinavian countries being the main role model because of having a high subsidised childcare coverage and large labour market participation of women, largely attributed to their daycare policies. There are often two primary objectives for subsidising childcare. First, to make childcare affordable, as high costs are often seen as a barrier to entering the labour market. Second, to create a reliable environment that supports children's development, particularly for those from disadvantaged backgrounds.

Canada and in particular Quebec, has also implemented policies that have generated significant improvements in the labour force participation of mothers. Following the substantial subsidised childcare experience of Quebec, other provinces such as Ontario, British Columbia, and Alberta adopted similar approaches and are moving to expand subsidised childcare. To further support families with children and fight poverty, different cash transfer programs, whether universal or means-tested, have been used at the federal and provincial levels. However, the structure of these programs may discourage work and have counterproductive effects.

Despite the availability of affordable childcare and various child benefit programs, low employment rates, high welfare dependency, and poverty continue to be significant issues for mothers, particularly single mothers with younger children. In this context, could the cost of childcare still be the primary barrier to employment? Could alternative forms of financial support better promote work participation? What other factors, besides childcare costs, might influence individuals' labour market decisions? It seems that encouraging work, strengthening labour market attachments, and avoiding long periods of market inactivity and skills erosion especially in today's rapidly changing economies, are the key strategies to reduce poverty and should be the focused of policymakers.

This paper focuses on a sample of low-educated single mothers in Quebec who are between 25 and 45 years old and have at least one child below the age of 13. The paper investigates their decision regarding labour supply jointly with the decision about formal childcare utilisation and social assistance participation using a structural framework. Many single mothers in this group rely on social assistance and it is important to consider their decisions about work and childcare utilisation jointly with social assistance participation. This paper aims to make the following contributions:

(1) It investigates whether childcare costs pose a significant barrier to work and how individuals respond to changes in childcare expenses. This is achieved by simulating two scenarios: free childcare (in-kind transfer) versus an equivalent cash transfer which gives the individuals the flexibility to allocate the funds as they see fit. The findings show that, first, the positive effect of free childcare on labour supply is marginal, and second, when given the option, fewer mothers choose to use the funds for formal childcare. Instead, many prefer to maintain their informal childcare arrangements and allocate the funds to other expenses. This shows that free childcare generates a strong demand for formal subsidised childcare and does not necessarily increase labour supply, crowd out informal care, lead to long waits for subsidised care, and potentially create additional challenges for those with a greater need for the service.

(2) It analyses various financial policies to determine what motivates work participation, leading to higher income and reduced poverty. The findings show that policies promoting employment are more effective in increasing labour supply of single mothers compared to free childcare. Specifically, reducing the dollar-for-dollar reduction rate in social assistance significantly increases the labour market participation among welfare participants, with a small negative effect on non-participants and no additional to the government.

(3) The study also investigates additional factors that may contribute to labour supply for single mothers. Specifically, it assesses how mothers' preferences for work-leisure, formal childcare utilisation, and welfare participation influence their decision-making process. The simulation analysis shows that both observed and unobserved characteristics influence preferences, but particularly the latter play a large role in labour supply decisions of single mothers.

The findings of this study provide valuable insights into the challenges, incentives, and preferences of this population and clear guidance for policy development.

This study is organised as follows. Section 2 reviews the related literature on childcare provision and maternal labour supply. Section 3 provides some institutional background on childcare and child benefit programs in Canada, particularly in Quebec. Section 4 discusses the methodological background. Section 5 describes the dataset and sample used. Section 6 analyses variations in policies and discusses the results. Section 7 compares the results to similar studies in the field and Section 8 concludes.

3.2 Literature

3.2.1 Childcare and Maternal Labour Supply

Many studies have investigated the relationship between childcare provision, childcare cost and maternal labour supply and child well-being. Most of these studies employ quasi-experimental methods with different identification techniques to measure the impact of a policy change on certain outcomes. Some also use a structural framework (ex. Ribar (1995); Andrén (2003); Wrohlich (2011); Haan and Wrohlich (2011)) to simulate a specific scenario. The findings show significant variations. These variations have been attributed to methodological issues (Cascio (2009); Blau and Currie (2006)), other changes coinciding with the policy shift (Bettendorf et al. (2015); Baker et al. (2008)) or the environment in which the policy was implemented such as the availability of childcare services, pre-existing labour market conditions, labour demand and general economic growth during policy implementation (Lundin et al. (2008); Nollenberger and Rodríguez-Planas (2015); Haeck et al. (2015)). Additionally, it has been shown that childcare policy responses have significant heterogeneous effects across sub-groups of the population.

Blau and Currie (2006) reviewed studies on childcare cost and employment and noticed a wide range of labour supply elasticities, ranging from nearly zero to -1.3. They attribute most of these variations to identification issues, the treatment of unpaid childcare and childcare cost estimations in the model, rather than the data sources and sample characteristics. For example, they explain that the few studies that appropriately included unpaid childcare in the choice set tend to estimate lower elasticities.

In a comprehensive analysis of studies, Del Boca et al. (2015) examined the relationship between labour market participation and childcare costs. The study finds that childcare accessibility, rather than childcare cost, significantly influences the size of the labour supply response. In countries, where childcare is widely accessible, the labour supply elasticities tend to be larger. Conversely, in countries with limited availability and rationed public childcare, the elasticities are smaller. The review of these studies also shows heterogeneous responses based on the income, skills and education levels of the target groups. Similarly, Wrohlich (2011) use a structural model of labour supply and childcare utilisation on German data and finds a stronger labour supply response to the availability of childcare compared to the effect of childcare costs in Germany where childcare is subsidised and rationed. Also, using a similar structural approach, Haan and Wrohlich (2011) simulated the effect of a proposed childcare subsidy set to take place in Germany in 2013. Their findings indicated only a one percentage point increase in the employment rate of married and cohabiting women.

Many quasi-experimental studies assess the effectiveness of different childcare policies. These studies suggest that in countries where childcare subsidies rarely existed the policies were more effective in increasing labour market participation of mothers. Baker et al. (2008) and Lefebvre and

Merrigan (2008) examined the effect of the introduction of the universal \$5/day childcare program in Quebec in 1997, which was available for children up to age 4. According to Baker et al, the labour market participation of women in two-parent families increased by 7.7 percentage points (14.5 percent) and the childcare enrollment rate increased by 14.6 percentage points. The difference between childcare enrollment and labour force participation also indicates a considerable substitution away from informal care into formal care. Similarly, Lefebvre and Merrigan (2008) and Lefebvre et al. (2009) found a significant increase in maternal labour supply, both at the intensive and extensive margins. The authors also find varying effects of this policy by education levels. Mothers with lower educational backgrounds respond more to this policy. The authors explain that this difference could be because of lower reservation wages in this group or the lower labour market participation rate prior to this change. Haeck et al. (2015) investigated the long-term effects of this policy approximately ten years after its introduction. They find persistent positive effects of the program on labour supply of single mothers, particularly for highly educated mothers. A significant positive effect on this group was observed from early in the program, while there was a delay for the less educated group. The authors explain that this difference might be due to the higher-educated group benefiting from a higher reduction in costs, as they were not eligible for low-cost childcare before this change. The authors attribute the greater impact of this policy to the dynamic nature of the Canadian labour market compared to that of Europe.

Gelbach (2002) and Fitzpatrick (2012) study the effect of free preschool enrollment in the US (Oklahoma and Georgia) during different periods. Using data from 1980, Gelbach shows a significant improvement in labour supply and other social-income measures of single mothers with five-year-old children. The findings show a 4 to 5.1 percentage point (6 percent to 10 percent) increase in their employment, as well as a 10 to 16 percent increase in hours and weeks worked, and a 10 percent decrease in social assistance participation. On the other hand, in a similar study using 2000 data, Fitzpatrick finds that only employment of single mothers without additional young children increased by 15 to 20 percentage points (22 percent to 29 percent). The effects on other labour supply measures and married women were insignificant. The author concludes that the differences in results between these two studies are likely related to changes in labour supply behaviour and lifecycle patterns over time.

In another study on the US, Cascio (2009) examines the expansion of slots in public school kindergartens between 1960 and 1980. The results show that single mothers with five-year-old children increase their participation by 6.9 percentage points (12 percent) and their hours of work by 2.4 (11 percent). The impact on married women and mothers with younger children is small and statistically insignificant. Additionally, the study finds a significant shift from private to public school kindergartens.

Studies on European countries have similar findings. For example, Goux and Maurin (2010) find a significant increase in labour market participation and childcare utilisation among single mothers after the introduction of universal public preschool for three-year-old children in France. The effects are insignificant for married mothers and variations were observed across education levels. Specifically, mothers with lower education have a 5.1 percentage points increase in their labour supply, while those with higher education levels only experienced an increase of 0.2 percentage points.

Nollenberger and Rodríguez-Planas (2015) analyse the impact of expanding publicly subsidised full-time childcare for three-year-old children in Spain during the early 1990s. Before this reform, maternal labour supply was around 30 percent with low childcare utilisation. The authors find a significant increase of 2.8 percentage points (9.6 percent) in the labour force participation of mothers of three-year-olds. They also found heterogeneity in the response across age and family size. Employment of mothers above the age of 30 and those with two or more children increased by 4.7 (15.3 percent) and 3.9 percentage points (14.9 percent), respectively. The authors mention that this reform occurred during the slow economic growth in Spain, suggesting that the program could have been more effective under different economic conditions.

In a study on Italy, Carta and Rizzica (2018) exploit labour market outcomes of highly subsidised universal daycare reform in the mid-2000. This reform provided early access to daycare for 2-year-olds and resulted in a decrease in reservation wages, leading to a greater willingness to work for a lower return. Consequently, maternal labour market participation and employment increased by 6 and 5 percentage points (roughly 10 percent), respectively. The study also shows evidence of heterogeneity in the effects based on regional labour market demand, education level, marital status, family income and the presence of a younger child. Specifically, the effects are more pronounced in regions with higher vacancy rates, among mothers with high school degrees, married mothers, higher-income families, and those without younger children at home.

Brewer et al. (2016) examined the effect of offering free childcare in England during the 2000s to mothers with children aged 3 or 4. Their study distinguishes between part-time and full-time care. They find that providing part-time care does not affect the labour supply of mothers significantly, but providing full-time care increases their labour force participation by 5.7 percentage points (8.7 percent). The study also finds that effects vary by mother's education and marital status. Mothers with lower education levels and single mothers were more responsive to this change.

Then we have the studies that evaluate the maternal labour market outcomes of a policy change in countries with high childcare provision or already high labour market participation rates. The findings in these studies generally indicate a modest impact. For example, Lundin et al. (2008) studied the effect of a 2002 reform in Sweden that imposed a cap on how much municipalities could charge for subsidised childcare. The study finds that the reform has nearly zero impact on the maternal labour supply of women in two-parent families. It is worth noting that the baseline employment before the reform was 70 percent with over 80 percent working full-time, leaving little room for improvement. These results do not significantly vary by education level or the age of children. Havnes and Mogstad (2011) investigated the effect of a large expansion of full-time free childcare across municipalities in Norway which began in mid 1970's for children between the ages of 3 and 6 years old. At the time of this expansion, Norway had a high maternal employment rate of about 70 percent, and the expansion only led to an increase of 1.1 percentage point (4.5 percent) in the employment of married mothers. Additionally, the study finds a large substitution between informal and formal care. The authors conclude that the policy was ineffective in increasing maternal employment with a significant net cost of childcare subsidies. Similarly, Givord and Marbot (2015) find that the significant increase in the childcare subsidy in France in 2004 had limited effects on increasing the maternal labour supply of young mothers. The employment increased by 1 to 2 percentage points, and the childcare take-up increased by 1.7. Given the cost of the program, these effects are considered modest.

The mean effect of some of these policies could be small when considering a larger population, but more studies have shown that the response could vary among different groups. For instance, Bettendorf et al. (2015) study the effect of generous childcare subsidies in the Netherlands in 2015.

At that time, the employment rate of mothers in 2004 was among the highest in the OECD and comparable to that of Norway, Sweden and the US. However, the majority of mothers worked part-time. The study finds that the maternal employment rate only increased by 2.3 percentage points (3 percent) and the average weekly hours worked increased by 1.1 hours (6.2 percent). This suggests that a large share of the increased subsidies went to parents who were already using formal childcare at a higher cost prior to the policy change, indicating a substantial shift from informal to formal childcare. However, consistent with most other studies, they find a larger positive effect on single mothers compared to couples. Single mothers increased their labour supply by 4.7 percentage points (8.4 percent) and by 1.7 hours per week (12 percent) at the extensive and intensive margins, respectively, while couples show a smaller increase of 2.3 percentage points (3.2 percent) and 1.1 (7.2 percent) hours per week.

Huebener et al. (2020) study the effect of abolishing the small private contribution made by families to parental labour supply in Germany starting in 2006. Their results show no evidence of change at the extensive margins, and only a 1 percentage point (7.2 percent for the very low full-time rate) increase in full-time workers. However, they found the response is stronger by single mothers, mothers without younger children, those who live in urban areas perhaps with dense local labour markets, highly educated mothers and low-income households, where the increase in employment ranges from 2.5 to 6 percentage points. The authors consider the program to be inefficient if the goal of the policy is boosting labour supply. They suggest that there are more efficient programs through targeted plans linked to employment and childcare expenses.

Busse and Gathmann (2020) exploit the effect of the introduction of free childcare slots for children between 2 and 6 years old in Germany. Their results show heterogeneity across the children's age. The labour force participation of mothers of young (2-3 years old) children increased by 8 percentage points (18 percent) and daycare take-up by 6 percentage points (12 percent). But the increase in labour supply of mothers of older children (4-5 years old) is only significant at the intensive, increasing weekly hours by 10 percent. Moreover, their results show that formal and informal care are complementary.

Ito and Yamamoto (2022) find heterogeneity by type of employment. They found that employment among non-regular employees increased by approximately 14 percentage points as a result of a series of childcare policies implemented in Japan in the early 2000s. But regular employment only responds at the intensive margin by increasing work approximately by an equivalent of a day per week. The authors mention that the lack of effect at the extensive margin for the latter group could be difficulty finding this type, regular employment, of job and access to childcare are not the main determinants of that.

In some countries, unlike childcare subsidy, a different type of policy was implemented aiming at motivating homecare through forms of cash transfers. For example, in Norway since 1999, parents with children between the age of one to three years old, became eligible for a cash-for-care (CFC) subsidy if they chose not to use publicly subsidised daycare regardless of their employment status. Schøne (2004) finds a slight reduction of 3 percentage points (4 percent) in maternal labour market participation and a decrease in annual work hours of 3 percent. Naz (2004) examined the effect of this policy on labour force participation of couples and their relative work hours. The study finds that the husbands' labour supply did not change significantly, while women's labour supply decreased, leading to an increase in the gap between the hours worked by couples. The negative effect is larger for women with higher levels of education. The study by Kornstad and Thoresen (2007) is one of the few studies that uses a structural approach to estimate the effect of this reform on two-parent households with preschool children. They estimate that the labour supply of mothers reduced by 9 percent which is greater than the findings in the previous two studies. However, they argue that the effects measured in the previous studies were immediate, whereas the results from their model could be considered long-term responses, allowing families more time to adjust.

Kosonen (2014) study the effect of a homecare policy in Finland, similar to Norway's Cash-For-Care but more generous and designed specifically for stay-at-home care, unlike the Norwegian one which was aimed at non-use of public care. The study finds a significant decrease in the participation of women, with no effect on fathers. The estimates indicate that a 100 euro increase in monthly supplement, reduces the participation of women by 3 percentage points (10 percent).

Gathmann and Sass (2018) analyse the impact of a generous cash incentive for homecare offered in Thuringia-East Germany in 2006. Their findings show an-8 percentage point drop in formal childcare attendance. While no immediate effect on the labour supply of mothers was observed, labour force participation decreased by 4-5 percentage points (5 percent) in the full sample after a couple of years. They also noted a substantial decline in labour force participation of single parents, low-skilled parents and low-income parents, ranging from 13 to 26 percentage points. In another study, Collischon et al. (2020) studied the effect of a different homecare policy introduced in 2013 for parents with children under the age of three in Germany. They find a 1.4-percentage point (1.7 percent) decrease in maternal employment and a 5-percentage point (10 percent) reduction in subsidised childcare take-up.

3.2.2 Labour Supply and Social Assistance

In addition to Gelbach's study in 2002 that is discussed above, there are a few other studies investigating the relationship between childcare cost, employment and social assistance participation. Connelly and Kimmel (2003) review the papers in this field and find that there is consensus in the direction of the relation between social assistance reliance and childcare cost, yet the size of the estimates vary. The authors conducted a study using the US data and find a significant drop in welfare dependency in response to childcare subsidy. They also find that the estimated elasticity of AFDC recipiency with respect to childcare cost is more sensitive to model specification and the estimates are between -1 to -2, but childcare cost elasticity of employment is relatively consistent and is around -0.42 and -0.42 for single mothers.

Andrén (2003) conducts a study on the relationship between childcare expenses and the employment decisions of single mothers using data from Sweden. In the study, he developed a structural labour supply model that accounted for the joint decisions regarding labour force participation, childcare utilisation, and reliance on social assistance. Through a microsimulation exercise, Andren finds that increasing social assistance support did not significantly impact childcare utilisation but had a relatively larger negative effect on labour supply at both the intensive and extensive margins. On the other hand, reducing childcare costs has a minor positive effect on labour supply, particularly at the intensive margin, but a more pronounced impact on social assistance participation. He estimates that a 10 percent increase in social assistance results in a 0.19 percent increase in childcare utilisation, while a 10 percent decrease in childcare cost leads to a 1.6 percent reduction in social assistance reliance.

Milligan and Stabile (2007) analysed the impact of integrating the National Child Benefit (NCB) with the provincial social assistance program in Manitoba. Their findings reveals a 3.4 to 4.7 percentage point reduction in social assistance participation and a 3.3 to 4.6 percentage point decrease in employment between 1997 and 2000 due to this integration. In a subsequent 2009 study, the authors examined the effects of discontinuing the integration for children aged zero to five and later, extending it to children aged zero to eleven. The results indicate that the policy expansion increased social assistance participation and decreased the percentage of income earners, particularly among less educated families.

3.2.3 Other Child Benefit programs in Canada

Some studies have also evaluated other forms of child benefit programs that exist in Canada using a quasi-experimental method. The major federal programs in recent years have been Universal Canada Child Benefit (UCCB) which was replaced by Canada Child Benefit (CCB) in 2016. Although these programs are not directly related to childcare, they are considered cost-reducing support offered by the federal government. Koebel and Schirle (2016) examine the effect of the 2006 UCCB introduction on women based on their marital status. The findings show a negative income effect on married women, with a decrease of 1.4 percentage points in their participation rate and a reduction of approximately one hour work per week. In contrast, the study finds that the participation of divorced mothers increased by 2.8 percentage points, with no significant effect on hours. Also, no significant effect was found among common-law and never-married mothers. In a similar study, Schirle (2015) analysed the effect of this program on married women with different education levels. In this study, she observes that the negative effect is much more pronounced for low-educated women. Schirle notes a 3.2 percentage point decrease in the participation rate of low-educated women and a reduction of 1.9 hours per week in their weekly hours. In comparison, highly educated mothers experience a reduction of nearly one hour per week in their weekly hours of work.

In Messacar (2021) study, the impact of expanding the Universal Canada Childcare Benefit (UCCB) policy to families with children aged 0 to 17 on earnings and childcare expenditures is investigated. The results indicate that only the earnings of single mothers with above-average incomes decreases. Specifically, a one percent increase in child benefit transfers was associated with a 0.04 percent reduction in their earnings. The effect on the earnings of married women and single mothers with below-average incomes is insignificant. Additionally, the study shows that childcare spending increases by 25 cents for every one-dollar increase in child benefit transfers among married women, while the effect on childcare spending for single mothers is minimal. Baker et al. (2023) study some of the labour market outcomes of the UCCB expansion followed by the introduction of the CCB program. The results show that these programs are successful in reducing poverty by approximately 5 percentage points (12 percent) between 2014 and 2018. They do not find any evidence of a labour supply response among single and married women.

3.3 Institutional Background on Childcare and Child Benefit

In 1997, the Quebec government introduced a new Family policy program where subsidised childcare program was the main part of it. This program aimed to provide subsidised childcare for children aged 0-4. Initially, it covered children aged 4, and gradually expanded and became available for all children under the age of 5 by 2000. Under this program, parents were required to pay a fixed fee of $5 per day, per child. Furthermore, the new family policy also introduced optional full-time kindergarten and subsidised after-school care for children aged 5–12. Initially, the number of subsidised slots were limited, causing long wait times. However, the number of slots increased significantly since the introduction of the program, although the wait time is still considerable.

Before this reform, low-income families were eligible for direct subsidies for childcare. Also, there was a refundable tax credit for childcare expenses, ranging from 75 percent for low-income families to 25 percent for high-income families. Under the new reform, those using the subsidised facilities were no longer eligible for provincial subsidies and tax credits, but they could still claim the expenses towards federal deductions.

In 2004, the fixed contribution increased to $7. In 2015, the fixed flat fee was replaced with a fee proportional to income. The total fee consisted of a fixed amount of $7.30/day plus a varied amount based on gross income and the variable component was paid through tax contributions. Under this system, payments varied between $7.30/day for families earning less than $50,000 up to $20/day for families earning over $150,000. [1]

To benefit from the subsidised system, families must commit to a set, full-time schedule of 5 days a week for a full day, for up to 261 days per year. There is limited flexibility, although part-time options are available, they are limited and less common[2].

The federal and provincial governments have other child benefit programs aimed at providing additional financial support and reducing poverty. These programs are not conditional on formal childcare or employment. In 2015, the main federal child support programs included Canada Child Tax Benefit (CCTB), National Child Benefit supplement (NCB)[3] and Universal Canada Child Benefit (UCCB). All these programs were replaced by Canada Child Benefit (CCB) in 2016. The UCCB is a universal and taxable payment and was determined by the age and number of children. However, single parents were eligible to claim it under an eligible child to reduce the tax payment. The other programs are non-taxable cash transfers and income-tested benefits. NCB was specially targeted to lower-income families. At the provincial level, Quebec's Family Allowance (Retraite Quebec) offers a refundable tax credit where the payments vary by age and number of children, family income, and family status.

For example, consider a single mother with two children, one in preschool and one in primary school in 2015. This family would qualify for a total of $2640 under UCCB. For CCTB, the maximum amount of $2940 is received if the income is approximately less than $45000. The maximum payment from CCB is $11800 which applies to incomes up to $30,000. Additionally, from the provincial government, families are eligible for the Family Allowance with a maximum amount of $4378 for income up to $35000.[4] Before 2016, the maximum benefit for a single mother with an annual income of $30,000 was $9,958. After the introduction of the CCB, which replaced the UCCB and CCTB in 2016, this amount increased to $16,178.

[1] In 2019, the government eliminated the variable component and set the fixed contribution at around $8.25/day. The current daily rate is $9.10.

[2] For more detailed information on childcare policy in Quebec, see Baker et al. (2008), Lefebvre and Merrigan (2008), Lefebvre et al. (2009), and Haeck et al. (2015).

[3] This benefit was available in all provinces except Quebec.

[4] For detailed information on these programs, see CTaCS by Milligan (2016).

3.4 Methodology

Many papers employ a structural approach with discrete choices to model labour supply, as seen in works such as van Soest (1995), van Soest et al. (2002). Structural models offer the advantage of assessing the impact of ex-ante policy and program changes on individuals' choice sets through detailed tax and credit schedules in the budget constraint, something quasi-experimental methods cannot accomplish. Some of these studies expand this approach by including childcare utilisation decisions when analysing a sample of mothers. The idea is that the labour supply decision of individuals with children, particularly women who serve as the primary caregivers, is related to childcare utilisation. Therefore, both aspects must be incorporated into the model. For example, Ribar (1995) estimates a fully structural model on a sample of married women in the US. Similarly, Wrohlich (2011) uses a similar approach for mothers in two-parent households using panel data to account for the highly rationed childcare market in Germany. Wrohlich estimates the expected value of childcare cost by using an estimated probability of subsidised childcare utilisation versus private options as weights. Using a similar model, Haan and Wrohlich (2011) model the labour supply and fertility decision of single mothers to evaluate the impact of subsidised childcare on these decisions. Andrén (2003) estimates the joint decision between labour supply, childcare utilisation and welfare participation for single mothers in Sweden. Additionally, Kornstad and Thoresen (2007) use variations of this structure to estimate the joint decision between labour supply and childcare utilisation, differentiating between types of care such as parental, daycare or other paid options, in order to simulate the effects of homecare cash transfer policy in Norway.

This study uses a similar approach used in these papers. It focuses on a sample of low-educated single mothers who highly rely on social assistance. Therefore, the analysis in this paper will consider the decision about social assistance participation with labour supply and choice of formal childcare utilisation simultaneously. Unlike Wrohlich (2011) and Haan and Wrohlich (2011), we are unable to distinguish between subsidised and private care utilisation as this information is not available in the dataset.

3.4.1 Labour Supply Model

This study employs the standard structural labour supply model based on random utility. This model assumes that individuals get utility from combinations of disposable income (C) and leisure (L). We further assume that the utility function is a log-transformation of the Cobb-Douglas function, which is a common form used in the labour supply literature (among others, van Soest (1995); van Soest et al. (2002); Flood et al. (2004); Haan (2006); Clavet et al. (2013); Hansen and Liu (2015)). Additionally, in accordance with Keane and Moffitt (1998) and Andrén (2003), we include the choice of social assistance participation and paid childcare utilisation as separate additive terms to this utility function. The estimated model is

$$U(C_{jcs}, L_j, SA_s, CC_c) = \beta_c log(C_{jcs}) + \beta_l log(L_j) + \beta_{csq}(logC_{jcs})^2 + \beta_{lsq}(logL_j)^2 + \beta_{cl} log(C_{jcs}) log(L_j) + \beta_{sa} dSA_s + \beta_{cc} dCC_c + \epsilon_j, \quad j = 1, ..., 7; s = 0, 1; c = 0, 1 \tag{21}$$

where C_{jcs} and L_j are respectively, disposable income and leisure time from alternative j within hour categories given social assistance (s) and childcare utilisation (c) status. We assume there are $j = 7$ categories for worked hours ($h = 0, 250, 750, 1250, 1750, 2250, 2750$). dSA and dCC are binary variables that denote the observed state for social assistance participation and paid childcare utilisation. ϵ_j is a random component assumed to follow a type I extreme value distribution, which is a common assumption in discrete choice models, providing an easy and fast estimation process. This accounts for idiosyncratic preferences for certain jobs/alternatives or an error in individuals' evaluation of each alternative.

Individuals face a budget set from which they choose a combination of C, L, SA, CC that maximises their utility. Disposable income (C_{jcs}), or the budget constraint is defined as

$$C_{jcs} = wh_j + y - t(wh_j + y_j^T - D_j) + TC_j + dCC_c \cdot CC_j + dSA_s \cdot SA_j, \tag{22}$$

where,

w: Before-tax hourly wage rate

h_j: Annual hours of work corresponding to alternative j where $L_j = T - h_j$ and $T = 4,000$

y: Annual non-labour income (eg, investment income, child support, employment insurance, etc.)

t: Federal and provincial income tax schedule[5]

y_j^T: Taxable non-labour income D_j and TC_j: Federal and provincial deductibles and tax credits, respectively

CC_j: Daycare expenses if the person uses paid care

SA_j: Social assistance benefits if the person participates in the program. Social assistance benefits are reduced dollar for dollar (a 100% reduction rate) for earnings that exceed the $200 monthly exemption threshold.

The decision to use the social assistance program and paid childcare service enter the model in two ways: as a direct benefit and cost, respectively if the individual opts for either or both options, in which case alters their disposable income and varies with hours of work; and as a binary additive term in the utility function, reflecting the individual's valuation of these two options. This framework enables the model to capture possible trade-offs involved in the choice between work, paid childcare services, and social assistance participation.

Following van Soest (1995), we assume that there is a fixed cost of working. This can be interpreted as transportation or any work-related cost that occurs from working. In empirical studies, the prediction of the proportion of non-workers improves by including this term; otherwise, the proportion would be underpredicted. The fixed cost enters the model as $log(C_{jcs}) - log(fc)$, where $j > 1$.

We further assume that if the mother does not work, she will not use paid childcare either[6]. Additionally, given eligibility, participation in social assistance is only an option if the person does not work full-time hours (ie, available if $h = 0, 250, 750, 1250$). Therefore, there are a total of 20 alternatives: $(dh1 - dh7, dCC = 0, dSA = 0), (dh2 - dh7, dCC = 1, dSA = 0), (dh1 - dh4, dCC = 0, dSA = 1)$, and $(dh2 - dh4, dCC = 1, dSA = 1)$.

The multinomial logistic probability of choosing the alternative (j, c, s) is

[5] Canadian Tax and Credit Simulator (CTaCS) by Milligan (2016)

[6] Only close to 5 percent of the sample uses paid childcare while not working, and they are excluded.

$$p_{jcs} = \frac{exp(V_{jcs})}{\sum_j \sum_q exp(V_{jcs})}, \tag{23}$$

where V_{jcs} is the deterministic part of the utility function in (21). The log-likelihood function is

$$LL = \sum_n \sum_{jcs} \delta_{n,jcs} ln(p_{n,jcs}), \qquad n = 1, \ldots, N; j = 1, \ldots, 7; c = 0, 1; s = 0, 1 \tag{24}$$

where δ_{jcs} is equal to 1 for the observed state of (j, c, s) and 0 otherwise. N is the number of individuals and $p_{n,jcs}$ is the probability of choosing set (j, c, s) by individual n.

The primary limitation of standard conditional logit is the assumption of Independence of Irrelevant Alternative (IIA) which is unrealistic in the context of labour supply decisions. To address this limitation, we relax this assumption by incorporating random preference parameters where an individual's evaluation of each alternative is affected by their unobserved characteristics. These individual-specific preferences are defined for leisure, social assistance participation and paid childcare utilisation, assuming heterogeneity through β_l, β_{cc} and β_{sa} where,

$$\begin{aligned} \beta_l &= b_{0l} + b_l X_l + e_l, \\ \beta_{cc} &= b_{0cc} + b_{cc} X_{cc} + e_{cc}, \\ \beta_{sa} &= b_{0sa} + b_{sa} X_{sa} + e_{sa}. \end{aligned} \tag{25}$$

X_l, X_{cc}, X_{sa} are observable characteristics. $e = (e_l, e_{cc}, e_{sa})$ is individual specific unobserved characteristics and captures differences among individuals.

Moreover, by incorporating individual-specific characteristics, such as age, education level, number and age of children, household characteristics and living arrangements as well as unobserved characteristics, we account for diverse labour market behaviours and provide a more detailed analysis.

3.4.2 Wage Estimation

We estimate the wage equation simultaneously with the structural labour supply model and allow the error term in the wage regression to be correlated with the labour supply model through β_l, β_{cc}, and β_{sa}. The natural log wage is defined as

$$log(w) = b_{0w} + b_w X_w + e_w + \varepsilon_w, \tag{26}$$

where X_w is the individual characteristics. The error term in the wage equation is defined in two parts: first, unobserved heterogeneity among individuals,e_w, which is correlated with unobserved heterogeneity in the preference parameters. Second, there is the measurement error, $\varepsilon_w \sim N(0, \sigma_w^2)$. van Soest et al. (2002) interpret the measurement error as job- or hour-related characteristics that are not in the labour supply part of the model. Blomquist (1996) and van Soest et al. (2002) discuss that ignoring the measurement error can significantly bias the estimates of elasticities. Furthermore, we assume that e_w and ε_w are normally distributed and uncorrelated with each other.

Finally, $\varphi = (e_w, e_l, e_{cc}, e_{sa})$ is i.i.d and follows $MVN(0, \Omega)$.

The choice probability for (j, c, s) set is

$$p_{jcs} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Pr(h_{jcs}|\varepsilon_w, e_w, e_l, e_{cc}, e_{sa}) \cdot f(\varepsilon_w|e_w)$$
$$\varphi(e_w, e_l, e_{cc}, e_{sa}) de_w de_l de_{cc} de_{sa}, \quad j = 1, \ldots, 7, c = 0, 1; s = 0, 1 \tag{27}$$

where $f(\cdot)$ is a wage density given e_w. The choice probabilities are calculated given a set of $(e_w, e_l, e_{sa}, e_{cc})$. Then, to remove the condition, (e_w, e_c, e_l, e_s) are integrated out.

The simulated choice probability for a worker is

$$\hat{p}_{jcs} = \frac{1}{R} \sum_{r=1}^{R} Pr(h_{jcs}|\varepsilon_w, e_w^r, e_l^r, e_{cc}^r, e_{sa}^r) \cdot f(\varepsilon_w|e_w^r), \quad j = 1, ..., 7, c = 0, 1; s = 0, 1 \tag{28}$$

and for a non-worker is

$$\hat{p}_{jcs} = \frac{1}{R}\sum_{r=1}^{R} Pr(h_{jcs}|\varepsilon_w, e_w^r, e_l^r, e_{cc}^r, e_{sa}^r), \quad j = 1, ..., 7, c = 0, 1; s = 0, 1 \tag{29}$$

where $(e_w, e_l, e_{sa}, e_{cc})$ is the r^{th} draw from $MVN(0, \Omega)$. For estimation, we use 100 draws from Halton sequence.

Substitute $\hat{p}_{jcs}$ in (4) to get the simulated log-likelihood function (SLL)

$$SLL = \sum_{n}\sum_{jcs} \delta_{n,jcs} ln(\hat{p}_{n,jcs}), \quad j = 1, ..., 7; c = 0, 1; s = 0, 1; n = 1, ..., N. \tag{30}$$

The parameters to estimate are $\theta = (\beta_c, \beta_{csq}, \beta_{lsq}, \beta_{cl}, b_{0w}, b_{0l}, b_{0cc}, b_{0sa}, b_w, b_l, b_{cc}, b_{sa}, \Omega, \sigma_\varepsilon)$. These parameters are estimated using the Simulated Maximum Likelihood Estimation method.

3.4.3 Estimation of Childcare Expenditure

We need to determine the cost of childcare (CC_j) for each alternative in the budget constraint. This will be based on estimated values derived from a regression for the childcare cost function. The variations in the cost are from hours of work, non-labour income, age and number of children and regional differences. To ensure the reliability of the estimates, the regression uses a sample of mothers with the youngest child under the age of 14 in order to increase the sample size and improve the robustness of the estimates[7]. The childcare cost function is

$$log(childcare\ \ expense) = f(nchild2, nchild2to5, nchild6to14, chpop, bigct, rural, locw, unearned, dh1, dh3, dh4, dh5, dh6, dh7). \tag{31}$$

In this regression, '*childcare expense*' is the annual childcare cost. Three variables are allocated for the number of children in the family who could potentially require care: '*nchild2*' is the number of children younger than two years old, '*nchild2to5*' for the number of children between 2 and 5 years old, '*nchild6to14*' for the number of children between 6 and 14. '*unearned*' is non-labour income, '*dh1*' to '*dh7*' are dummy variables for hour categories where the categories are $dh_j = 0$,

[7]Variations of this larger sample have been tested and the estimated coefficients are close.

250, 750, 1250, 1750, 2250, and 2750, respectively (dh2 = 250 is set as the base category [8]). The same categories are used in the labour supply model. These categories can capture any nonlinear relationship that could exist between hours and childcare costs. '*chpop*' indicates the number of children less than 14 years old relative to the population in the local area where the person resides. This variable captures the interaction of supply and demand for childcare in each neighbourhood. '*bigct*' and '*rural*' are dummy variables for living in big cities and rural areas versus medium-sized cities. *locw* represents the local average wage, serving as a proxy for socio-economic characteristics at the neighbourhood level.

Table 3.1: Childcare Expenses

Variable	Estimate
Dependent Var. *ln (childcare expense)*	
Intercept	6.630^{***} (0.040)
dh1	-0.167^{***} (0.027)
dh3	0.136^{***} (0.021)
dh4	0.227^{***} (0.019)
dh5	0.285^{***} (0.017)
dh6	0.294^{***} (0.017)
dh7	0.315^{***} (0.022)
Nchild2	0.157^{***} (0.009)
Nchild2to5	0.539^{***} (0.006)
Nchild6to14	0.041^{***} (0.005)
Chpop	-0.067^{**} (0.034)
Bigct	0.088^{***} (0.009)
Rural	-0.039^{***} (0.011)
Locw	0.011^{***} (0.001)
Unearned	0.017^{***} (0.001)
R^2	0.183
$Adj\,R^2$	0.183

***, ** significance at 1% and 5%; Values in parentheses are standard errors.

Table 3.1 presents the estimated coefficients. All the estimates are statistically significant.

[8]Childcare utilisation is not considered for non-workers (i.e dh1) in the main sample. That is why dh2 is used as the reference category.

Childcare expense increases with the number of children and is higher for younger kids. Furthermore, costs increase with working hours, city size, and non-work income. In neighbourhoods with a higher average local wage, childcare costs tend to be higher. The estimated coefficient of '*chpop*' is negative suggesting that, after controlling for factors such as age and number of children, regional differences and economic factors, neighbourhoods with a higher density of children, may have a larger supply of, possibly subsidised, childcare centres at a lower cost.

3.4.4 Identification

The identification of the utility parameters relies on variations across individuals and alternatives which is based on multiple categories of hours (leisure) and disposable income. The income variation is also affected by wage, non-labour income, childcare cost as well as hours of work. Furthermore, identification is also based on the highly non-linear nature of tax function. Additionally, the fixed cost is identified by non-workers versus workers.

To further facilitate identification within the main model, β_c is normalised to one. Additionally, we apply the exclusion restriction for the wage, childcare expense and childcare utilisation equations. In the wage equation, variables on language skills are considered to be human capital which affects labour demand. In the childcare expense equation, the proportion of the total number of children in the population is a measure of net childcare supply, which affects childcare costs. For the childcare utilisation equation, the presence of an adult in the household affects the preference for childcare utilisation.[9]

As for the rest of the variables in the preference parameters of the utility function, the mother's education, age and number of children are included in all three preference parameters to reflect the differential preferences and limitations of single mothers. Additionally, preference for leisure and social assistance participation are affected by immigration status and age.

[9] Due to the limited availability of more suitable variables in the dataset, we were unable to use alternative variables. However, we estimated the model with different initial values, and we found that the main estimates of the model remained stable.

3.5 Data and Sample

This study uses the 2016 Canadian Census of Population. Census of population is a mandatory survey with a cross-section design, conducted by Statistics Canada every five years to provide information on the participants' demographic, social and economic characteristics. The 2016 Census microdata file contains 8,651,677 observations that consist of 25 percent of the population. The personal demographic information from the survey is linked to the 2015 personal income tax and benefits record collected by Canada Revenue Agency (CRA) to provide a rich dataset with reliable information on the sampled population's demographic, social and economic characteristics.

To estimate the childcare cost, we use the variable for childcare expense in the census. The variable measures the annual cost of daycares, day camps, babysitters and similar services for all the children in the family, which enables the parent(s) to be employed. However, it does not specify the type of childcare, subsidised, or private use. The outliers, defined as annual childcare costs per child less than \$150 and exceeding \$17,000 are also excluded and represent less than two precent of the sample.

This study focuses on single mothers in Quebec between the ages of 25 and 45 years old with a high school degree or less, whose youngest child less than 14 years old. Specifically, the sample includes those who are the primary financial providers for their households. It excludes individuals receiving retirement income, students, those reporting disabilities, or those receiving disability benefits.

The hourly wage is not directly available in the census data. We estimate this variable by dividing the annual earnings by the number of worked weeks multiplied by hours of work during the reference week. There are limitations to this method. To get a more reliable distribution, we impose some restrictions on the sample: individuals with self-employment earnings are excluded since the reported hours and earnings might be ambiguous. This represents less than 3 percent of the sample. Individuals who worked more than 85 hours weekly are dropped, as are those who worked less than three weeks or less than three hours per week. These account for 0.13 and 0.1 percent of the sample, respectively. We also set the minimum wage at \$10 (Minimum wage in Quebec in 2015) and the maximum at \$100. Consequently, any estimated wage value beyond this range is excluded from the sample. These exclusions represent close to 5 percent and 0.5 percent of the sample, respectively.

The annual hours of work are estimated using weekly work hours and weeks worked annually. Then seven categories are created such that $dh1 - dh7 =$

$$\begin{cases} dh1 = 0 & \text{if hour} = 0, \\ dh2 = 250 & \text{if } 0 < \text{hour} \leq 500, \\ dh3 = 750 & \text{if } 500 < \text{hour} \leq 1000, \\ dh4 = 1250 & \text{if } 1000 < \text{hour} \leq 1500, \\ dh5 = 1750 & \text{if } 1500 < \text{hour} \leq 2000, \\ dh6 = 2250 & \text{if } 2000 < \text{hour} \leq 2500, \\ dh7 = 2750 & \text{if } 2500 < \text{hour}. \end{cases}$$

In addition, individuals who are not employed but have childcare spending are excluded from the main sample, constituting less than 5 percent of the sample. Also, the dummy variable for paid childcare utilisation is defined when there is a reported positive childcare expense in the dataset. Social assistance participation is defined as individuals receiving the equivalent of at least one month of full benefits. Also, '*adult*' is a dummy variable where it is equal to one if there is any child above the age 13, or if there are any relatives or grandparents in the household. Non-labour income is defined as investment income, employment insurance benefits, and other income such as spousal and child support. Since the income information in the census data refers to 2015, this study is based on the 2015 taxes and credits schedule. The tax and credits schedule is based on the parameters in CTaCS by Milligan (2016).

Another variable in this study is the poverty rate, defined as the proportion of the sample living below the poverty line. We use the Low-Income Measure (LIM) in the census, which sets the poverty line at 50% of the median after-tax household income, adjusted for household size using the square root of the household size. This adjustment accounts for the fact that larger households have more needs, but those needs do not increase proportionally with size. [10] LIM is widely used

[10] To calculate the LIM, household income is divided by the square root of the household size to adjust for family needs, and the median of these adjusted incomes is determined. The LIM for each person is then set at 50% of this median, adjusted for household size. See Statistics Canada, Census (2016).

for international poverty comparisons and is increasingly applied within Canada.

3.5.1 Descriptive Statistics

Table 3.2 presents sample means of the variables used in this study, as well as based on the age of the youngest child. The average age in the sample is 36. The overall employment rate in this sample is 68 percent which shows that labour force participation is relatively high but this is driven by those with school-aged children. The average hours worked among workers is approximately 1650 hours, which is equivalent to full-time work and shows that the majority of workers are already working full-time hours. However, the social assistance participation is 31 percent, and more importantly, the poverty rate is also quite high at 46 percent, indicating that a considerable portion of workers earn below the typical income level in society. For reference, in our sample, the average poverty line is $37,812, though this varies by household size.

Additionally, 32 percent of the sample have a preschool-aged child, and 39 percent, which represents over half of the workers, use formal childcare while the rest rely on informal childcare arrangements. Moreover, 33 percent of the sample have another adult living in the household, such as an older sibling, grandparents, or a relative.

Columns 2 and 3 of Table 3.2 present the same statistics for subsamples categorised by the age of the youngest child. As expected, individuals with preschool-aged children show poorer labour market outcomes, with lower labour supply, higher participation in social assistance programs and higher poverty rates. Moreover, a smaller share of this group uses formal childcare, and the average childcare costs are approximately double those of individuals without preschool-aged children.

Table C.1 in the appendix shows similar statistics by childcare utilisation.

3.6 Results

3.6.1 Estimated Parameters of the Model and Model Performance

Table 3.3 shows the model estimates. Unlike the reduced form, most of the estimated parameters of the utility function do not have a direct interpretation. However, it is worth mentioning a few points. With these estimates, the marginal utility of consumption is positive, and utility is concave

Table 3.2: Descriptive Statistics of the Sample

Variable	Full Sample	Preschool-aged	School-aged
Employment Rate	67.93	57.14	72.95
Annual Hours of Work	1649.30 (1035.88)	1630.96 (1068.02)	1656.00 (1023.55)
Hourly Wage	20.10 (20.98)	19.78 (20.89)	20.22 (21.01)
Age	36.18 (10.64)	33.11 (10.12)	37.61 (9.63)
High School Graduates	51.03	48.02	52.43
Having Preschool-aged Child	31.78	100	0.0
Daycare Use	38.80	48.68	34.20
Welfare Participation	30.85	41.40	25.94
Poverty Rate	45.94	51.59	43.31
Immigrant	17.44	18.39	17.01
Bilingual	35.52	36.38	35.12
Anglophone	4.79	5.29	4.56
Adult	33.21	22.09	38.45
Number of Children	1.86 (1.82)	1.99 (1.99)	1.81 (1.71)
Disposable Income	35, 278.45 (27,251.97)	34, 303.40 (25,855.11)	35, 732.40 (27,856.89)
Childcare cost	2973.38 (5518.08)	4422.43 (6705.76)	2012.50 (3399.75)
Average Local Employment	0.77 (0.09)	0.77 (0.09)	0.77 (0.09)
Observations	11,895	3780	8115

Values in parentheses are standard deviations.

The numbers of observations are weighted and rounded.

for all individuals.

The estimated parameter for social assistance participation is the marginal (dis)utility for social assistance. The constant term is negative which shows that after controlling for individual characteristics, there is an average stigma associated with being on social assistance. Also, the estimates of individual characteristics show that the stigma is more pronounced among high school graduates compared to high school dropouts and increases with the mother's age. Additionally, the stigma decreases with the age and number of children, and for immigrants, although these effects are statistically insignificant.

Next, the estimated parameter for childcare utilisation represents the marginal utility of using formal childcare. The estimated parameter shows a small and statistically insignificant disutility from using formal childcare. Also, mothers with high school degrees and those with younger kids derive greater utility from using formal childcare. Moreover, the presence of another adult at home reduces the utility of using formal childcare. The estimations suggest that formal childcare utilisation may not always be the preferred choice for some single mothers. They might instead prefer to use maternal care or unpaid childcare. The disutility of using paid childcare could reflect concerns about daycare quality or long waiting times to secure a spot or limited flexibility in scheduling.

Moreover, the estimated parameter for leisure shows that individuals with high school degrees and those of older age tend to exhibit a greater desire for work. On the other hand, having more children, especially younger ones, and being an immigrant contribute positively to the utility derived from leisure. The effects of all these variables are statistically significant.

Finally, the estimated correlation matrix of unobserved heterogeneity shows the correlation between wage, preference for leisure, formal childcare utilisation and social assistance participation. The direction of estimated correlations is consistent with expectation. I will focus on those with larger correlations. The correlation between wage and childcare utilization is 0.59, indicating that the characteristics that positively influence wages also tend to increase childcare use. Additionally, the correlation between leisure and childcare utilisation and between leisure and social assistance participation are -0.70 and 0.90, respectively, indicating that those who prefer more leisure use formal childcare less and are more likely to receive social assistance. Finally, the correlation between childcare utilization and social assistance participation is -0.94, suggesting that the characteristics that increase formal childcare use tend to reduce participation in social assistance program.

Table 3.3: Estimates for the Utility Function

	Estimate	Standard Error	t value
ln^2C	0.041***	0.006	7.103
$ln(C \cdot L)$	0.363***	0.083	4.377
ln^2L	-17.091***	1.313	-13.022
fc	7.644***	1.685	4.537
Wage Equation			
Age	0.115***	0.017	6.819
Age^2	-0.078***	0.030	-2.611
Immigrant	-0.097***	0.025	-3.906
High school	0.158***	0.020	8.028
Bilingual	0.133***	0.019	7.057
Anglophone	0.152***	0.042	3.651

	Estimate	Standard Error	t value
Local Employment	1.107***	0.124	8.944
Constant	1.906***	0.096	19.906
σ	0.329***	0.013	25.953
Observed Characteristics for Leisure			
High school	-3.885***	0.533	-7.296
Nchild (age under 3)	4.589***	0.944	4.860
Nchild(age 3to5)	2.748***	0.507	5.417
Nchild(age 6to14)	1.674***	0.342	4.887
Immigrant	2.344***	0.550	4.264
Age	-1.340***	0.398	-3.371
Constant	27.971***	2.784	10.046
Observed Characteristics for Childcare Utilisation			
High school	0.537***	0.115	4.688
Nchild(age under 3)	0.991***	0.238	4.165
Nchild(age 3to5)	1.092***	0.122	8.923
Nchild(age 6to14)	-0.099	0.079	-1.264
Adult	-0.839***	0.110	-7.642
Constant	-0.085	0.149	-0.568
Observed Characteristics for Social Assistance Participation			
High school	-5.997*	3.439	-1.744
Immigrant	5.220*	3.121	1.673
Nchild(age under 3)	5.358	3.680	1.456
Nchild(age 3to5)	2.625	1.712	1.534
Nchild(age 6to14)	1.454	1.082	1.343
Age	-2.795*	1.575	-1.775
Constant	-7.422*	4.275	-1.736
Elements of the Cholesky Decomposition Matrix			
W_w	0.218***	0.020	10.861
L_w	1.130**	0.486	2.325
CC_w	0.516***	0.116	4.443
SA_w	-2.908	2.152	-1.351
L_l	6.730***	0.811	8.302
CC_l	-0.705***	0.123	-5.718
SA_l	10.017*	5.809	1.724
CC_{cc}	-0.028	0.259	-0.108
SA_{cc}	-0.469	0.761	-0.616
SA_{sa}	0.191	0.631	0.302
Loglikelihood		-7495.623	

***, ** and * are significance levels at 1%, 5% and 10%.

Table 3.4 compares the predicted and true hour distribution across childcare utilisation and social assistance participation. The predicted values are based on simulations of 40 draws per individual from the estimated multivariate normal distribution, Ω, in Table 3.3.[11] This comparison shows that the generated distribution closely matches the true distribution across the 20 categories.

[11]The number of draws for the policy simulation exercise was limited by the memory limitations of the lab's computer.

Variance-Covariance Matrix for Unobserved Preferences				
	Wage	Leisure	Childcare	SA
Wage	0.048			
Leisure	0.247	46.575		
Childcare	0.113	-4.158	0.764	
SA	-0.635	64.129	-8.545	109.044

Correlation Matrix for Unobserved Preferences				
	Wage	Leisure	Childcare	SA
Wage	1.0			
Leisure	0.166	1.0		
Childcare	0.591	-0.697	1.0	
SA	-0.279	0.900	-0.936	1.0

3.6.2 Policy Simulation

In this section, we explore the impact of various support programs through financial assistance and how they influence individuals' decisions regarding their employment, use of paid childcare, reliance on social assistance, poverty status, and disposable income. The results are also based on 40 simulation draws as described above.

For reference, Table 3.5 displays the average simulated characteristics of the sample based on the model estimates prior to the policy change. The overall estimated employment rate is 66.1 percent, with an average annual working hours of 1639 among workers. Social assistance participation is 28.6 percent, and 37.7 percent of individuals utilise the paid childcare system. The estimated poverty rate within the sample is 46.6 percent, and the average disposable income (after deducting childcare expenses for those using the paid system) is $34,583. These values align closely with the true values reported in Table 3.2.

This table also shows the same estimated averages by the subgroups of single mothers not utilising the paid childcare system ($dCC = 0$) and welfare participants. The first one includes non-employed individuals. These subgroups are identified as having lesser attachment to the labour market and are at higher risk of living in poverty. The effect of simulated policies by the age of the youngest children is reported in Tables C.2 to C.7 in the appendix.

Tables 3.6 to 3.10 present the simulation results following a policy change for the entire sample and the subsamples. The average change in response to each policy change are reported. The outcomes are a change in employment rate (in percentage points), the average percentage change

Table 3.4: Model Fit - Predicted versus True Proportions

	Hour Categories	True	Predicted
CC=0; SA=0	0	3.91	7.14
	250	0.92	0.32
	750	2.27	2.17
	1250	4.41	7.57
	1750	14.67	11.34
	2250	5.13	5.53
	2750	0.29	0.34
CC=1; SA=0	250	1.09	0.34
	750	2.77	2.87
	1250	5.34	9.02
	1750	18.58	14.02
	2250	9.21	9.52
	2750	0.50	1.19
CC=0; SA=1	0	28.16	26.8
	250	0.71	0.4
	750	0.42	0.52
	1250	0.25	0.2
CC=1; SA=1	250	0.59	0.31
	750	0.25	0.32
	1250	0.46	0.08

Predicted values are based on simulations of multiple draws per individual.
The true values are based on weighted and rounded counts.
The unweighted and unrounded frequencies in some categories are closer to the model estimate.

of working hours among workers, a change in paid childcare utilisation (in percentage points), a change in social assistance participation (in percentage points), and the average percentage change in disposable income (after deducting paid childcare cost for those using paid childcare) among those affected by the policy.

Table 3.5: Benchmark

	Employment Rate (pct)	Hours	Welfare Participation (pct)	Childcare (pct)	Poverty rate (pct)	Disposable Income ($)
All	66.06	1639.25	28.64	37.66	46.64	34,583.3
CC=0	45.55	1589.48	44.80	0.00	61.52	31,804.5
Preschool	25.90	1435.54	65.01	0.00	73.28	31,580.4
School	53.28	1618.94	36.85	0.00	56.89	31,892.7
SA=1	6.42	629.40	100.00	2.49	87.90	27,674.1

”pct” stands for percent change.

Free Daycare versus Cash Transfer

The first policy examines the comparison between access to free daycare, which can be viewed as an in-kind transfer, and a dollar equivalent cash transfer. In the first scenario, individuals have access to free daycare if they choose to use the paid childcare system, in which case they have no costs. In the second scenario, everyone is entitled to the same amount of cash as the value of free daycare, regardless of whether they use paid childcare or not. Therefore, in this case, individuals have the choice to allocate the cash towards childcare or other living expenses. The outcomes of this comparison are presented in Tables 3.6 and 3.7.

Table 3.6: Free Childcare

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare (ppt)	Poverty rate (ppt)	Disposable Income (pct)
All	0.459	0.014	-0.081	1.695	0.181	2.653
CC=0	0.737	0.402	-0.128	2.719	-0.167	0.521
Preschool	1.636	1.489	-0.316	4.079	-0.514	0.972
School	0.383	0.194	-0.054	2.184	0.000	0.343
SA=1	0.684	1.150	-0.383	0.757	-0.128	0.364

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

In the case of free daycare, there is a 1.70 and 0.46 percentage points increase in childcare utilisation and employment, as well as a 0.01 percent rise in working hours among workers. Moreover, an analysis of sub-sample responses shows that single mothers with preschool-aged children have a greater labour supply response and uptake of childcare compared to other groups. Specifically, within this demographic, there is a 1.64 and 4.08 percentage points increase in employment and

Table 3.7: Cash Transfer

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare (ppt)	Poverty rate (ppt)	Disposable Income (pct)
All	0.944	-0.261	-0.163	0.587	-5.240	5.230
CC=0	1.514	-0.265	-0.261	1.022	-4.317	4.022
Preschool	2.533	-0.289	-0.476	2.075	-4.447	4.053
School	1.112	-0.260	-0.176	0.607	0.000	4.010
SA=1	1.395	-0.054	-0.635	0.679	-1.392	0.969

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

childcare utilisation, and an average increase of 1.49 percent in hours worked. While the average changes in poverty rates and consumption are positive, they are relatively modest.

The difference between the change in employment rate and childcare utilisation shows that the increase in childcare use is primarily driven by parents who are already working, as they switch from unpaid childcare to the new free-paid system. Of every 10 mothers that start using paid childcare, only 3 have become employed. Moreover, the rise in employment rates is relatively modest (even across the sub-samples), indicating that the cost of daycare may not be the primary barrier to single mothers' employment. Also, the low uptake of paid childcare suggests that single parents with existing childcare arrangements are inclined to continue using them, even when the paid system is available for free.

Next are the results for the second scenario. Table 3.7 shows the impact of an equivalent cash transfer. The simulation results for those with $dCC = 0$ show that when given the choice, only 1.02 percent of the respondents opted for the paid childcare system, in contrast to 2.72 percent of the respondents who chose the free daycare. This suggests that most individuals prefer to use the support to cover other expenses instead of allocating it to childcare. This aligns with Messacar (2021) findings on the expansion of UCCB payments in 2015. He finds that these payments did not change childcare spending among low-income single mothers and reduced working hours for high-income single mothers, confirming that the extra cash was not used to reduce childcare costs for this group.

Furthermore, this policy leads to a broader improvement in employment, poverty rates, and consumption. This is because all working single mothers in the sample benefit from the policy, not just those utilising the paid childcare system. In this study, we observe a minor reduction in welfare participation and a small decrease in work hours (0.26 percent in the sample) as a result of cash transfers. This suggests that on average, the negative income effect of cash transfer is only slightly larger than the substitution effect of increased cash for more worked hours. The modest response at the mean indicates that this type of policy would not significantly distort the overall work incentive for single mothers.

Additionally, among the subgroups, parents with younger children have a larger labour supply response and are more likely to use paid childcare under both policies. This is understandable as they face greater work constraints due to the young age of their children. Therefore, these policies could be more beneficial to them.

In summary, the comparison of these two policies suggests that childcare costs are no longer the primary obstacle to the employment of single mothers. Consequently, a policy such as free daycare does not significantly enhance their poverty or disposable income as the labour supply responses are limited under this policy. The findings show that they prefer to use the support in different ways, either for non-childcare expenses or to achieve a better work-life balance by reducing their work hours. In this scenario, policies such as free daycare may lead to excess demand for daycare services, resulting in rationing and long wait times, potentially depriving those in need of the subsidised system.

Employment Incentive programs

The findings from the previous simulation indicate that childcare costs may not be the primary barrier to the employment of single mothers. Therefore, this section examines alternative factors or support programs to address this issue. The lack of adequate financial incentives or employment support may discourage them from returning to work. Therefore, we will now explore a few employment support policies to assess their impact.

Reducing the Benefit Reduction Rate

Table 3.8 presents the impact of reducing the benefit reduction rate of social assistance from 100 percent to 50 percent.[12] The results demonstrate a 5.2 percentage point increase in the employment rate among welfare participants, nearly doubling the employment rate for this group. Additionally, the hours worked by employees increased by 10.2 percent on average, which is approximately equivalent to 64 hours or two full-time weeks of work. The poverty rate decreased by 3.5 percentage points (4 percent), with an average disposable income increase of 2.6 percent.

Table 3.8: Reducing Benefit Reduction Rate to 50%

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare (ppt)	Poverty rate (ppt)	Disposable Income (pct)
All	1.613	0.289	0.475	0.561	-1.078	0.906
CC=0	2.587	0.351	0.560	1.013	-1.434	1.296
Preschool	3.565	0.903	0.580	2.306	-2.655	1.472
School	2.202	0.245	0.552	0.503	-0.953	1.227
SA=1	5.234	10.208	0.000	1.763	-3.552	2.569

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

The risk of having a flexible social assistance program is that it may attract more participants, increasing welfare dependency. However, social assistance participation increased by close to 0.5 a percentage point. This suggests that this flexible approach to incentivising employment primarily benefits existing recipients rather than drawing more individuals into the program.

Also, among those not using daycare, we see a significant increase of 3.6 percentage points (14 percent) in the employment of mothers with preschool-aged children. The poverty rate also decreases by 2.7 percentage points (4 percent). This is primarily because a large portion of this group is benefiting from social assistance. However, for mothers with school-aged children, these values are 2.2 (4 percent) and -1 percentage point (2 percent), respectively.

Introducing a Generous EITC program

This section evaluates the effect of introducing a generous Earned Income Tax Credit (EITC), similar to the one in the US. However, Canada offers similar, though less generous, programs through the federal government (WITB) and the provincial government of Quebec (Prime au Travail or Work Premium). In this section, these two programs are substituted with the more generous

[12] In the current system, social assistance benefits are reduced dollar for dollar (a 100% reduction rate) for earnings that exceed the $200 monthly exemption threshold.

EITC program. It is worth noting that the estimates here are not the full effects of introduction of such a program since Canada already has a similar policy. So the analysis here could only measure the effects of a more generous program. Otherwise, the full effects are likely to be larger than the estimates here.

Table 3.9 compares the generosity of these programs and Table 3.10 shows the simulated mean estimates. Employment is required for eligibility for these programs. The Canadian programs have a minimum annual earnings requirement of $2400 (the equivalent of 240 annual hours of work at $10 minimum wage). However, since this is very low, it should not affect the outcomes. The main differences between WITB and WP combined and EITC are that the latter offers much more generous benefits to single mothers with two or more children. Also, higher-income families, especially those with more than one child could still be eligible for the benefits under EITC.

Table 3.9: EITC versus WITB and Work Premium for Single Parents in 2015

	WITB*	Work Premium*	EITC		
Number of Children	At least one	At least one	One	Two	Three or more
Minimum Income Requirement	2400	2400		No minimum	
Phase-in Rate	0.12	0.30	0.34	0.40	0.45
Maximum Benefit	956.4	2190	3359	5548	6242
Phase-out Rate	0.20	0.10	-0.1598	0.2106	0.2106
Phase-out Threshold	11,974	9700	18,110	18,110	18,110
Zero Benefit Income	16,756	31,600	39,131	44,454	47,747

In the three programmes employment is required for eligibility.
*These parameters apply specifically to Quebec, as the programme's parameters differ from those in other provinces.

The impact of this policy on labour supply measures is modest. The employment rate increases by nearly 1 percentage point, while the hours worked decrease marginally by -0.08 percent, driven by mothers with school-aged children (See Table C.6). The overall disposable income increases by an average of around 3.8 percent. This is a significant reduction in poverty rate by nearly 7 percentage points (15 percent) due to the scale of the program and the affected population. This is mainly driven by the lower-income individuals who were affected by this policy. This, in fact, shows the efficacy of this program in reducing poverty among those who can benefit from this program.

Analysing the results by the age of the youngest children shows some heterogeneity among families. Labour supply increases more for mothers with preschool-aged children compared to those with school-aged children at both the intensive and extensive margins. Employment rates rise by 1.5 percentage points (6 percent) versus 1.1 percentage points (2.2 percent). Additionally, the average worked hours for mothers with preschool-aged children increase, indicating that the substitution effect of the wage increase through this program outweighs the negative income effect and encourages more work on average among workers in this group. However, the opposite pattern is seen for mothers with school-aged children. Despite this, the poverty rate reduction for mothers with school-aged children is nearly double that of preschool-aged children (11 percent vs 5 percent). This is likely due to their already established attachment to the workforce, making them more eligible for cash transfers under the more generous program, even without significant employment changes.

Table 3.10: Introducing EITC

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare (ppt)	Poverty rate (ppt)	Disposable Income (pct)
All	0.778	-0.076	-0.152	0.406	-6.855	3.815
CC=0	1.248	0.024	-0.234	0.979	-5.587	3.358
Preschool	1.500	0.524	-0.292	1.345	-3.372	2.243
School	1.149	-0.071	-0.212	0.836	-6.459	3.797
SA=1	0.852	1.192	-0.574	0.394	-0.444	0.447

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

Similarly, the analysis of individuals in the social assistance program shows less than one percentage point (13 percent) increase in employment and a 1.2 percent increase in worked hours primarily because their base hours are low. The average growth in disposable income and the improvement in the poverty rate are minimal. Additionally, less than 1 percent of welfare participants exit the social assistance program, indicating a concerning level of welfare dependency.

This program shows that while it can enhance labour force participation to some extent, its effects are limited due to the preexisting similar programs in Canada. However, it has a significant positive effect on poverty reduction, due to the program’s generosity and its broader coverage rather than through stronger employment incentives. This type of policy would also generate higher costs for the government. We will discuss this later in this study.

Comparing the impact of a generous EITC to reducing the benefit reduction rate, the latter has a stronger effect on employment, especially among mothers with preschool-aged children. This is due to the high welfare participation rate in this group and the fact that a lower benefit reduction rate provides a greater net return than the generous EITC for non-workers considering entering the workforce.

Combined Policy: Free Daycare, Generous EITC and Flexible Social Assistance program

Next, we combine the three policies discussed above to evaluate a comprehensive employment incentive program. This includes implementing EITC, providing free childcare, and reducing the benefit reduction rate to 50 percent, providing both employment incentives and poverty reduction approaches. Table 3.11 presents the results.

Table 3.11: Combined Policy

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare (ppt)	Poverty rate (ppt)	Disposable Income (pct)
All	3.007	0.094	0.316	2.714	-8.895	7.800
CC=0	4.823	0.469	0.326	4.542	-8.599	5.858
Preschool	7.140	1.924	0.165	7.843	-8.517	5.851
School	3.911	0.190	0.390	3.242	-8.631	5.861
SA=1	7.417	11.924	-0.684	3.326	-6.881	4.581

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

These changes have resulted in substantial improvements across a range of measures. Among individuals not using the childcare system, there has been nearly a 5 percentage points (11 percent) increase in employment. Additionally, the poverty rate and disposable income have both shown significant improvements of 9 percentage points (14 percent) and 6 percent, respectively. The employment effects are more pronounced for parents with preschool-aged children.

Similarly, among welfare participants, we observe a 7.4 percentage point increase in employment and a 12 percent increase in hours worked. Moreover, the poverty rate and disposable income have improved significantly, by 7 percentage points and an average of 4.5 percent, respectively.

The Effect on Government Budget

Table 3.12 provides an estimate of the net cost of introducing each program for the government. This is the difference between the net government budget before and after the introduction of the program. The net government budget before the policy change is defined as government revenue minus transfers to individuals. The main components of government revenue include net income taxes, CPP (Canada Pension Plan), EI (Employment Insurance), QPIP (Quebec Parental Insurance Plan), and health contributions. The main transfers include refundable tax credits (at the federal level, this includes the Working Income Tax Benefit (WITB) and GST rebates; at the provincial level, this includes Work Premium and solidarity tax credits, such as QST rebate, and housing credit), child benefits (including UCCB (Universal Canada Childcare Benefit), CCTB (Canada Child Tax Benefit), and Family Allowance), and social assistance. To get the net government budget after the policy change, the estimated cost of each new program is incorporated into this calculation.

Offering free daycare costs the government approximately 10 million dollars, though this figure does not account for the cost of expanding (subsidised) childcare centres; it simply reflects the government's payment of private contribution fees. Reducing the benefit reduction rate saves $290 thousand dollars. While social assistance becomes more flexible, the increase in employment could potentially broaden the tax base and improve the government budget. The generous EITC program costs the government approximately $14 million dollars, primarily due to the program's scale and the relatively modest overall positive employment effects. However, it shows strong benefits in reducing poverty. Finally, the combined approach costs the government $25 million dollars. To provide context, we also estimate the cost of replacing the CCB with UCCB, CCTB and NCB which was primarily aimed at reducing poverty. The cost of this change in this population group is approximately $29 million dollars.

Table 3.12: Government Budget: Estimated Collected Net Tax

Policy	Free Childcare	Cash Transfer	50% Benefit Reduction Rate	EITC	Combined Policy	CCB
Net Tax	-10.48	-19.47	0.29	-14.08	-25.25	-28.79

Values are in millions and represent population estimates. Positive values indicate a gain for the government.

For promoting employment, a flexible social assistance program with employment incentivising features seems to be an effective way to boost employment without a significant burden on the government budget.

3.6.3 The Importance of Unobserved Preferences

In this section, we investigate the importance of various factors, such as individual characteristics, financial motives, net tax payments, and preferences, on single mother's decisions about labour market outcomes. To do this, we use the structural model from this study. Given the observed characteristics, we draw 100 samples from the estimated multivariate normal distribution for unobserved heterogeneity to simulate outcomes, including utility, employment, hours of work, childcare utilisation and welfare participation. We then use the simulated sample and run regressions of each outcome variable on these characteristics. Table 3.13 compares the R-squared values from these regressions.

Table 3.13: R^2 from OLS Regressions

Regressors	Utility	Employment	Workers' Hours	Daycare Utilisation	Welfare Participation
Hourly wage	0.0031	0.0080	0.0008	0.0072	0.0106
Observed Characteristics*	0.1756	0.0975	0.0527	0.0676	0.1005
Net tax	0.3560	0.3834	0.2517	0.1321	0.3841
Unobserved Preferences	0.6679	0.4218	0.3273	0.2503	0.4554

*Observed characteristics are age, number and age of children, education, adult, immigrant.

The first column shows the R-squared value for utility. The observed characteristics of the parents such as age, the number and age of children, education level, being an immigrant, and the presence of another adult at home, explain approximately 18 percent of the variations in utility. Net taxes and unobserved preferences account for 36 and 67 percent of the variations in the utility, respectively.

Columns (2)-(5) present a similar analysis for different outcomes. Observed characteristics explain 5 to 10 percent of the variation in these outcome variables, and the hourly gross wage has a small share of the outcomes. Comparing the outcomes, we find that net taxes and unobserved preferences explain a significant portion of the variation in employment and welfare participation, but less so in hours of work and daycare utilisation. This could be because choices regarding work hours and childcare use are less flexible and not entirely dependent on an individual's response to taxes or their observed and unobserved characteristics. Factors such as offered work hours and childcare availability are also important in actual worked hours and childcare utilisation, respectively, in addition to individual's desire to work or use childcare.

This is an important finding as it shows that while financial incentives and tax structures do influence individuals' optimal choices, a significant portion of the choice is driven by unobserved characteristics. This explains why, despite various tax and credit policies and forms of childcare support, improvements are often limited and perhaps less than desired.

3.7 Discussion

How do these results compare to similar studies? Here we mainly focus on the free daycare policy. Refer to Table 3.6 and Table C.3 in the appendix to compare the results by the age of the children. Our results show that the labour supply effects were limited, with an approximately 0.5 percentage point increase in overall employment and close to a one-percentage-point increase for mothers with preschool-aged children who were not using paid childcare before the policy change. On the other hand, we observe nearly a 2-percentage point increase in childcare utilisation in the full sample (Table 3.6) and over a 2-percentage point increase for the subsample of mothers with young children (Table C.3). Our findings align with studies that have found small effects on labour supply but with heterogeneous effects by the age of the youngest children. As explained earlier in Table 3.2, this may be due to the relatively high employment rate in our sample, which is nearly 70 percent, with an average of over 1,600 hours worked, equivalent to full-time employment. However, since the employment rate is lower (close to 57 percent) among mothers of young children, we observe larger responses in this subsample.

It is difficult to numerically compare the studies because the policy environments, types, and magnitude of these policies differ significantly. A common approach in comparing the effects in the literature is to examine the change in employment to the change in childcare use after a childcare reform. Table 3.14 reports some similar studies. Our results from the free childcare policy show an overall ratio of 0.27 meaning that for every 10 mothers who started using childcare after the policy change, approximately 3 became employed. This suggests a substantial crowding-out effect in the informal childcare sector. In other words, 7 mothers switched from informal to formal care without any change at the extensive margin.

Table 3.14: Comparison with Related Studies

Study	Treatment Group by the Age of Children	Country	Policy	Employment to Childcare Take-up Ratio
Nollenberger and Rodriguez-Planas (2015)	3 years old	Spain	Expanding publicly subsidised full-time childcare for preschool-aged children	0.2
Goux and Maurin (2010)	3 years old	France	Universal public pre-elementary school for three-year-olds	0.38
Cascio (2009)	5 years old	US	Expanding public school kindergarten slots	0.4
Baker et al (2008)	Up to 5 years old	Quebec	Introduction of $5-a-day daycare in 1997	0.53
Havnes and Mogstad (2011)	3-6 years old	Norway	Expansion of subsidised childcare beginning in late 1975	0.06
Bauernschuster and Schlotter (2015)	3-4 years old	Germany	Expansion of public childcare in the late 1990s	0.37
Muller and Wrohlich (2020)	Up to 3 years old	Germany	Expansion of public childcare in the mid-2000s	0.2
Givord and Marbot (2015)	1-2 years old	France	Increase in childcare subsidies for preschool-aged children in 2004 reform	0.65
Bettendorf et al (2015) *	Up to 12 years old	Netherlands	Generous increase in childcare subsidies in 2005	Aged 0-3: 0.19 Aged 4-7: Less than 0.29 Aged 8-11: Larger than 0.16
Huebener et al (2020)	Up to age 4	Germany	Abolish private contributions in already highly subsidised daycare	0.5**
This study	Under 14 years old Preschool-aged School-aged	Quebec	Free Daycare	0.27 0.4 0.18

* For children aged 4-7 and 8-11, these values are approximate as separate childcare take-up data for each age category were not available to the authors.
** This ratio is based on hours. It is defined as the increased hours worked to increased hours at daycare per week.

Since most of these studies are on young children, we also report similar results for preschool-aged (age 5 and younger) and school-aged children (ages 6-13), separately. The ratio for mothers with preschool-aged children is 0.4, which is comparable to findings from Bauernschuster and Schlotter (2015) with a ratio of 0.37, Cascio (2009) with a ratio of 0.4, Goux and Maurin (2010) with a ratio of 0.38 and Baker et al. (2008) with a ratio of 0.53, who study similarly aged children. Givord and Marbot (2015) report a larger ratio of 0.65; however, this study focuses on children aged 1 and 2. Given the very limited work ability of mothers with such young children, a larger increase in employment following the provision of subsidised childcare for this age group is expected. The ratio reported in Nollenberger and Rodríguez-Planas (2015) and Havnes and Mogstad (2011) are on the lower end, with values of 0.2 and 0.06, respectively. This discrepancy may be due to the specific environments in which the reforms were implemented. For example, Nollenberger and Rodriguez-Planas point out that the Spanish economy was slow at the time of the policy which could have limited the increase in employment in that country.

For school-aged children, our study reports a ratio of 0.18. Most childcare studies focus on younger children. However, the study by Bettendorf et al. (2015) on Germany, which includes children up to age 12, estimates ratios between 0.16 and 0.29 for children older than 4. The precision of these estimates is limited due to the available data. Nonetheless, our estimates fall within their range and are closer to the lower end, which is consistent with the situation for children aged 6 to 13 in this study.

Table 3.15: Introduction of CCB

	Employment Rate (ppt)	Hours (pct)	Poverty rate (ppt)
All	-0.786	0.004	-4.792
Preschool	-0.577	0.069	-4.762
School	-0.886	-0.024	-4.807

Values represent the mean changes in response to the policy.
”pct” stands for percent change, and ”ppt” stands for percentage point change.

We also simulated the effect of replacing UCCB and CCTB with CCB. The results are in Table 3.15. In our sample, employment decreases by less than one percentage point, and the poverty rate decreases by 4.8 percent. Baker et al. (2023) have evaluated the effect of the expansion of UCCB and the introduction of CCB in 2015 and 2016 on labour supply and poverty focusing mainly on mothers and exploring the effects by their marital status, education level, and age of children using Labour Force Survey (LFS) and Longitudinal Administrative Data (LAD). We focus on their results on single mothers based on LAD as it is annual data and similar to our dataset. Their results show that the impact on labour supply at both the extensive and intensive margins is small and statistically insignificant, with the employment rate changing between -0.2 to +1 percentage point based on the sample used, year of the effect and specification of the model. Additionally, their estimates indicate a reduction in poverty by 5 percentage points, with a range of 2 to 6 percentage points depending on the subsample, year of the effect, and estimation method, and this result is statistically significant. Our estimates are consistent with Baker et al's findings.

3.8 Conclusion

Childcare cost and the provision of affordable childcare have been important policy areas in many countries as it is widely believed that the parents'-particularly mothers' labour market participation, is closely tied to these services. Various programs have been introduced in different countries to reduce these costs through tax credits and subsidised childcare centres. While studies have shown that these policies have improved mothers' labour supply, the effectiveness and efficiency of these policies vary widely and remain a topic of ongoing discussion.

Moreover, as economies rapidly evolve, the employment of parents, especially single parents, has become increasingly crucial. Extended breaks from the labour market due to childcare can lead to skills erosion and make re-entry into the workforce more difficult. It is important to provide the support that helps create and strengthen ties to the labour market. Therefore, this study assesses the importance of childcare costs, evaluates whether alternative forms of financial support can better motivate employment and improve their financial well-being, and finally, examines the role of personal characteristics and financial factors on individual decision-making regarding labour market outcomes.

This study uses a structural model with discrete choices to evaluate the labour supply of low-educated single mothers in Quebec. The model incorporates the labour supply decision along with childcare utilisation and social assistance participation as it is believed that the work decision of mothers, particularly single mothers, is closely linked to childcare. Moreover, since social assistance participation is significantly high among this group, it is important to include the interactions of this program with work decisions.

The findings of this study indicate that childcare cost is no longer the primary barrier to work. While employment does increase following the provision of free childcare, the increase is not significantly large. In fact, when given a choice, mothers often prefer to use financial support to cover other living expenses rather than paying for childcare. It's important to note that the cost function used in this analysis is unable to differentiate between individuals using subsidised care and those utilising the private sector. Therefore, the overall estimates are affected by the proportion of parents using low-cost subsidised care. Hence, it is important to note that while childcare costs may not be a significant factor at this stage, there could be a threshold beyond which expensive childcare costs could serve as a barrier to work. With reduced prices, access to childcare becomes a more important issue than further reductions in prices.

Additionally, we observe that their employment rises when the financial return to work is higher. Generous programs like the EITC increase employment and, more importantly, significantly reduce poverty, largely due to the scale of the program. However, those on social assistance and those with younger children – many of whom are on welfare and have less attachment to the labour market – show a considerable increase in work participation under a more generous social assistance program that promotes employment, such as reducing the benefit reduction rate. In contrast to the EITC, this type of program in the context of this study provides a larger net return to work and greater positive employment effects at both the extensive and intensive margins, with only a marginal negative labour supply response from previously non-welfare participants. This targeted policy generates more employment among the at-risk populations without excess cost to the government by increasing the tax base.

Given the limited response to childcare support and employment support programs, we also investigated the role of personal characteristics, both observed and unobserved, as well as financial incentives, in individuals' decisions regarding work and social assistance participation. The results indicate that unobserved characteristics have a significant influence on decision-making, even more so than observed characteristics and financial incentives. This is important because it explains the limited responses to these programs and highlights where future efforts should focus to achieve larger and more long-term effects in improving people's attitudes and perspectives toward work.

Appendix A

Chapter 1: Tables

Table A.1: Tax and Benefits Schedule in 2015 for Quebec and Ontario

Quebec		Ontario
Federal Progressive Personal Income Tax		
$F1 = inc \cdot 0.15$		$0 \leq F1 \leq 6705.15$
$F2 = (inc - 44701) \cdot 0.22$		$0 \leq F2 \leq 9834$
$F3 = (inc - 89401) \cdot 0.26$		$0 \leq F3 \leq 12788.1$
$F4 = (inc - 138586) \cdot 0.29$		$minF4 = 0$
$FT = F1 + F2 + F3 + F4$		
Federal Non-refundable Tax Credits		
basic amount	11327	
CPP/QPP	$(inc - 3500) \cdot 0.0525$ $0 \leq CPP/QPP \leq 2630.25$	
EI/UI	$inc \cdot 0.0154$ $0 \leq EI/UI \leq 762.30$	
non-refundable tax credit rate		0.15
QC abatement rate		0.165
Net federal = $((FT + CPP/QPP + EI/UI) - creditrate \cdot non-refundable) \cdot (1 - QCabatementrate)$		
Federal Refundable tax Credits = WITB + GST/HST credits		
WITB		
phase-in	$IN_{WITB} = 0.205 \cdot (inc - 2400)$ $0 \leq IN_{WITB} \leq 1633.85$	
phase-out	$OUT_{WITB} = 0.20 \cdot (inc - 11902.80)$ $minOUT_{WITB} = 0$	
Total	$IN_{WITB} - OUT_{WITB}$ $(minTotal = 0)$	
GST/HST Refundable Tax Credit		
basic amount	B_{GST} = 272	
phase-in	$IN_{GST} = 0.02 \cdot (inc - 8833)$ $0 \leq IN_{GST} \leq 143$	
phase-out	$OUT_{GST} = 0.05 \cdot (inc - 35465)$	

	Quebec		Ontario
	$minOUT_{GST} = 0$		
Total	$B_{GST} + IN_{GST} - OUT_{GST}$ $(minTotal = 0)$		
Provincial Progressive Personal Income Tax			
	Quebec		Ontario
$Q1 = inc \cdot 0.16$	$0 \leq Q1 \leq 6709.6$	$O1 = inc \cdot 0.0505$	$0 \leq O1 \leq 2066.6$
$Q2 = (inc - 41935) \cdot 0.2$	$0 \leq Q2 \leq 8386$	$O2 = (inc - 40922) \cdot 0.0915$	$0 \leq O2 \leq 3744.64$
$Q3 = (inc - 83865) \cdot 0.24$	$0 \leq Q3 \leq 4362$	$O3 = (inc - 81847) \cdot 0.1116$	$0 \leq O3 \leq 7605.9$
$Q4 = (inc - 102040) \cdot 0.2575$	$minQ4 = 0$	$O4 = (inc - 150000) \cdot 0.1216$	$0 \leq O4 \leq 8512$
		$O5 = (inc - 220000) \cdot 0.1316$	$minO5 = 0$
$PT = Q1 + Q2 + Q3 + Q4$	$(minPIT = 0)$	$PIT = O1 + O2 + O3 + O4 + O5$	$(minPIT = 0)$
Provincial Non-refundable tax credits			
basic amount	11425		9863
Living Alone			
Basic	1340		
phase-out	$OUT_L = (inc - 33145) \cdot 0.15$ $(minOut_L = 0)$		
Total $(minTotal = 0)$	$Basic - Out_L$ ‘		
Non-refundable Credit Rate	0.2		0.0505
$Ptax = PIT - creditrate \cdot non - refundables$			
Provincial Surtax			
$n.a.$			$S1 = (Ptax - 4418) \cdot 0.2$ $0 \leq S1 \leq 247.2$ $S2 = (Ptax - 5654) \cdot 0.36$ $(minS2 = 0)$
		Total	$S = S1 + S2$
		Ontario Tax Reduction(OTR)	$OTR = 228 \cdot 2 - S$ $(minOTR = 0)$
		ST	$S - OTR$ $(minST = 0)$
Net provincial = Ptax + ST			
Provincial Refundable Tax Credits			
Work Premium			
phase-in	$IN_{WP} = 0.07 \cdot (inc - 2400)$ $0 \leq IN_{WP} \leq 557.9$		

	Quebec		Ontario	
phase-out	$OUT_{WP} = 0.10 \cdot (inc - 10370)$ $minOUT_{WP} = 0$			
Total	$IN_{WP} - OUT_{WP}$ $(minTotal = 0)$			
Quebec Solidarity Refundable Tax Credits			*Ontario Sales Tax Credit*	
QST Credit				
basic amount	B_{QST} = 283+133		basic amount	B_{HST} = 287
phase-out	$OUT_{QST} = 0.03 \cdot (inc - 33145)$ $(minOUT_{QST} = 0)$		phase-out	$OUT_{HST} = 0.04 \cdot (inc - 22057)$ $(minOUT_{HST} = 0)$
Total	$B_{QST} - OUT_{QST}$ $(minTotal = 0)$		Total	$B_{HST} - OUT_{HST}$ $(minTotal = 0)$
Quebec Tax Return On Housing				
basic amount	$B_H = 539$			
phase-out	$OUT_H = 0.03 \cdot (inc - 33145)$ $(minOUT_H = 0)$			
Total	$B_H - OUT_H$ $(minTotal = 0)$			
Residents of Northern Villages				
basic amount	$B_{nv} = 1637$			
phase-out	$OUT_{nv} = 0.15 \cdot (inc - 33145)$ $(minOUT_{nv} = 0)$			
Total	$B_{nv} - OUT_{nv}$ $(minTotal = 0)$			
Provincial Health Contribution				
$h1 = 0$ if $inc < 18370$			$h1 = 0$ if $inc < 20000$	
$h2 = 0.05 \cdot (inc - 18370)$ $0 \leq h2 \leq 100$			$h2 = 0.06 \cdot (inc - 20000)$ $0 \leq h2 \leq 300$	
$h3 = 0.05 \cdot (inc - 40820)$ $0 \leq h3 \leq 100$			$h3 = 0.06 \cdot (inc - 36000)$ $0 \leq h3 \leq 150$	
$h4 = 0.04 \cdot (inc - 132650)$ $0 \leq h4 \leq 800$			$h4 = 0.25 \cdot (inc - 48600)$ $0 \leq h4 \leq 150$	
			$h5 = 0.25 \cdot (inc - 72000)$ $0 \leq h5 \leq 150$	
			$h6 = 0.25 \cdot (inc - 200000)$ $0 \leq h6 \leq 150$	
Total $H = h1 + h2 + h3 + h4$	$(maxH = 1000)$		Total $H = h1 + h2 + h3 + h4 + h5 + h6$	$(maxH = 900)$
Social Assistance $SA = B - BR \cdot (inc - EA)$ where $inc > EA$				
Annual Benefit (B)	$616 \cdot 12$		Annual Benefit (B)	$280 \cdot 9 + 305 \cdot 3$
Benefit Reduction rate (BR)	100%		Benefit Reduction rate (BR)	50%
Earning Allowance (EA)	$200 \cdot 12$		Earning Allowance (EA)	$200 \cdot 12$
Disposable inc				
For Non-welfare Participants: Total inc - Net federal + Net provincial + Federal and provincial refundable credits - H				
For Welfare Participants: Total inc - Net federal + Net provincial + Federal and provincial refundable credits - H + SA				

Note: Common tax and benefits that are applied to single persons. CTaCS by Milligan (2016)

Table A.2: Estimated Parameters Quebec Parameters from the 2006 Census and Different Tax and Transfer Systems

	(1)	(2)
$\log^2 C$	0.59^{***}	0.04
	(0.11)	0.14
$\mathrm{Log}C{\cdot}\log L$	1.97^{***}	0.22
	(0.31)	(0.17)
$\log^2 L$	7.52^{***}	8.98^{***}
	(1.55)	(1.21)
fixed cost of work	2.32^{***}	0.01
	(0.43)	(0.07)
Leisure preference parameters		
constant (type 1)^c	-31.04^{***}	-15.31^{***}
	(5.00)	(2.34)
constant (type 2)^c	-15.86^{***}	
	(1.75)	
age (dummy for age group 25 -29)^a	-1.44^{***}	-1.74^{***}
	(0.51)	(0.38)
age (dummy for age group 30 -34)^a	-1.93^{***}	
	(0.52)	
immigration	2.13^{***}	0.84^{*}
	(0.68)	(0.48)
bilingual	-1.26	-0.11
	(1.1)	(0.24)
English	0.27	0.43
	(0.45)	(0.64)
HS grads or HS dropouts with a certificate	-0.79	
	(0.56)	
HS grads with a certificate	-1.29^{**}	
	(0.53)	
Consumption preference parameters		
constant (type 1)^c	-7.77^{***}	-0.67
	(1.33)	(1.54)
constant (type 2)^c	3.82	
	(2.55)	
age (dummy for age group 25 -29)^a	-0.11	0.01
	(0.18)	(0.08)
age (dummy for age group 30 -34)^a	-0.31^{*}	
	(0.17)	
immigration	-0.04	-0.26
	(0.18)	(0.25)
HS grads or HS dropouts with a certificate	-0.16	
	(0.17)	
HS grads with a certificate	0.11	
	(0.18)	
Welfare preference parameters		
constant (type 1)^c	1.74^{***}	0.87
	(0.45)	(0.69)
constant (type 2)^c	2.71^{***}	
	(0.38)	
age (dummy for age group 25 -29)^a	-1.04^{**}	-1.26^{***}
	(0.44)	(0.35)
age (dummy for age group 30 -34)^a	-1.70^{***}	
	(0.43)	
immigration	0.64	1.44^{**}
	(0.53)	(0.56)
bilingual	-0.80	0.14
	(1.08)	(0.22)
English	0.04	0.64
	(0.48)	(0.43)
HS grads or HS dropouts with a certificate	0.79^{*}	
	(0.42)	
HS grads with a certificate	0.88^{**}	
	(0.39)	

	(1)	(2)
Wage equation		
age (dummy for age group 25 -29)[a]	0.12^{***} (0.03)	0.7^{***} (0.13)
age^2 (dummy for age group 30 -34)[a]	0.17^{***} (0.03)	-0.34^{***} (0.1)
immigration	-0.14^{***} (0.05)	-0.01 (0.04)
bilingual	-0.10 (0.11)	0.01 (0.02)
English	-0.10^{**} (0.05)	0.05 (0.06)
urban	0.03 (0.03)	-0.1^{***} (0.02)
HS grads or HS dropouts with a certificate	0.08^{**} (0.04)	
HS grads with a certificate	0.27^{***} (0.04)	
constant (type 1)[c]	2.52^{***} (0.04)	3.54^{***} (0.09)
constant (type 2)[c]	2.55^{***} (0.04)	
σ	0.41^{***} (0.01)	0.34^{***} (0.01)
Logit type 1 probability parameter[c]	-1.15^{***} (0.07)	
Multinomial logit for probability of education (HS dropout is the base)		
Equation I: HS grads or HS dropouts with a certificate		
age (dummy for age group 25 -29)[a]	0.07 (0.23)	
age (dummy for age group 30 -34)[a]	0.21 (0.23)	
immigration	-0.25 (0.36)	
bilingual	0.31 (0.75)	
English	0.42 (0.35)	
urban	0.48^{***} (0.19)	
constant	0.03 (0.20)	
Equation II: HS grads with a certificate		
age (dummy for age group 25 -29)[a]	0.76^{***} (0.21)	
age (dummy for age group 30 -34)[a]	0.55^{***} (0.21)	
immigration	-0.19 (0.33)	
bilingual	-0.13 (0.70)	
English	-0.09 (0.33)	
urban	0.56^{***} (0.17)	
constant	0.34^{*} (0.19)	

Values in parentheses are standard errors. (1) Census 2006 – Public data – High school dropouts, high school graduates and those with a non-university certificate or diploma. (2) Census 2016 – Master data – High school dropouts. (a) Note: In model (1), age is in three categories. Two dummy variables are defined where the first category (age 20-24) is the base. In model (2), age is continuous, and it is rescaled using this formula age = (Age-mean)/10. (b) Due to confidentiality, the number of observations are weighted and rounded. (c) To allow for unobserved heterogeneity, model (2) uses two types.

Appendix B

Chapter 2: Performance Metrics and Tables

- In Tables (2.2-2.4), bias in percentage, Bayesian Information Criterion (BIC) and means squared error (MSE) are defined as

$$Bias\% = \frac{MeanParameterEstiamte - TrueParameterValue}{TrueParamterValue} \cdot 100,$$

$$MSE = \frac{(ParameterEstimate_d - TrueParameterValue)^2}{ND}, \qquad d = 1, ..., ND$$

$$BIC = -2 \cdot LL + ln(N) \cdot K,$$

where ND: Number of datasets; N: Number of observations; K: Number of estimated parameters; LL: the log likelihood value.

- Table 2.8 shows the number of datasets that pass the chi-squared test for predicted hour distribution. The chi-squared test for each dataset is calculated as

$$\chi^2 = \sum_J \frac{(PredictedFrequency_j - TrueFrequency_j)^2}{PredictedFrequency_j},$$

where $J = 16$ is the number of categories.

- Table 2.9, the model performance is also tested by comparing how accurately it predicts the effects of a policy change. Percentage bias and MSEs are calculated as

$$Bias\% = \frac{MeanPredictedValue - MeanTrueValue}{MeanTrueValue} \cdot 100,$$

$$MSE = \frac{(ParameterEstimate_d - TrueParameterValue)^2}{ND}, \quad d = 1, ..., ND$$

where d refers to individual datasets and $ND = 100$ is the total number of simulated datasets.

- Table B.2.3 shows how well the two methods perform in regenerating the hour distributions. The absolute bias for each alternative is

$$Diff_j = |Predicted\%_j - True\%_j|, \qquad j = 1, ..., J = 16.$$

Table B.1: Sample Statistics

Variables	Mean	Std Dev
Hourly Wage	21.2	10.1
Nonlabour Income	1221.1	3795.4
Average disposable income for each hour category		
h= 0	8292.2	1968.5
h=200	10490.3	2079.9
h=400	11647.9	2976.7
h=600	13560.4	4207.2
h=800	15885.8	5367.9
h=1000	18333.5	6415.2
h=1200	20778.5	7417.7
h=1400	23208.1	8421.8
h=1600	25614.7	9453.9
h=1800	28001.2	10507.8
h=2000	30376.6	11573.3
h=2200	32744.4	12640.9
h=2400	35106.4	13707.5
h=2600	37456.9	14767.2
h=2800	39801.7	15817.6
h=3000	42140.4	16858.8
Observations	2010	

This table presents sample statistics for wage and non-labour income, the two exogenous variables in the model, along with the simulated disposable income based on the tax and credits schedule for given work hours. The simulated sample size consists of 2110 observations.

Table B.2: Parameter Estimates (DGP: (a) Normal - Small Variance)

	True	SCL	MCLN	LC*
Fixed Utility Parameters				
$\beta_{c_{sq}}$	2.87	1.87	2.90	2.93
β_{cl}	0.35	-0.09	0.39	0.44
$\beta_{l_{sq}}$	-2.15	-2.80	-1.89	-1.79
fc	2.25	2.60	2.24	2.24
Parameters of Normal Distribution				
β_{l_1}	22.00	17.25	21.87	27.87
β_{c_1}	15.50	11.96	15.63	13.28
β_{l_2}				19.64
β_{c_2}				18.08
$stddev(\beta_l)$	1.50		1.64	
$stddev(\beta_c)$	1.50		1.69	
$Cov(\beta_c, \beta_l)$	-1.91		-1.09	
$Probability1$				0.11
$Probability2$				0.26
$Probability3$				0.34
Loglikelihood		-3975.62	-3951.99	-3954.21

The table compares the mean estimates across 100 datasets to the true parameters. The simulated sample size consists of 2110 observations.
*Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1})$; $(\beta_{l_1}, \beta_{c_2})$; $(\beta_{l_2}, \beta_{c_1})$; $(\beta_{l_2}, \beta_{c_2})$

Table B.3: Parameter Estimates (DGP: (b) Bimodal – Small Variance)

	True	SCL	MCLN	LC*
Fixed Utility Parameters				
$\beta_{c_{sq}}$	2.60	1.47	2.39	2.67
β_{cl}	0.33	-0.31	0.12	0.40
$\beta_{l_{sq}}$	-2.10	-3.33	-2.52	-1.90
fc	2.40	3.06	2.47	2.38
Parameters of Normal Distribution				
β_{l_1}	17.00			16.41
β_{l_1}	14.00			14.42
β_{l_2}	20.00			22.37
β_{c_2}	11.00			10.89
$stddev(\beta_{l_1})$	0.70			
$stddev(\beta_{c_1})$	0.70			
$Cov(\beta_{c1}, \beta_{l_1})$				
$stddev(\beta_{l_2})$	0.70			
$stddev(\beta_{c_2})$	0.70			
$Cov(\beta_{c_2}, \beta_{l_2})$				
Parameters of the Overall Bimodal Distribution				
β_l	17.90	13.80	18.02	
β_c	13.09	9.78	13.15	
$stddev(\beta_l)$	1.54		1.09	
$stddev(\beta_c)$	1.54		1.98	
$Cov(\beta_c, \beta_l)$	-2		-1.02	
$Probability1$	0.7			0.19
$Probability2$				0.43
$Probability3$				0.16
Loglikelihood		-4361.17	-4332.63	-4328.61

The bimodal distribution is generated by combining two normal distributions, with portions of 70 and 30 percent.
The table compares the mean estimates across 100 datasets to the true parameters. The simulated sample size consists of 2110 observations.
*Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1})$; $(\beta_{l_1}, \beta_{c_2})$; $(\beta_{l_2}, \beta_{c_1})$; $(\beta_{l_2}, \beta_{c_2})$

Table B.4: Parameter Estimates (DGP: (c) Discrete Bimodal - Small Variance)

	True	SCL	MCLN	LC*
Fixed Utility Parameters				
$\beta_{c_{sq}}$	2.60	1.62	2.41	2.71
β_{cl}	0.33	-0.28	0.10	0.42
β_{lsq}	-2.10	-3.46	-2.73	-1.95
fc	2.40	2.95	2.47	2.38
Parameters of Normal Distribution				
β_{l_1}	17.00			16.99
β_{c_1}	14.00			14.40
β_{l_2}	20.00			21.35
β_{c_2}	11.00			11.10
$stddev(\beta_{l_1})$	0.20			
$stddev(\beta_{c_1})$	0.20			
$Cov(\beta_{c_1}, \beta_{l_1})$				
$stddev(\beta_{l_2})$	0.20			
$stddev(\beta_{c_2})$	0.20			
$Cov(\beta_{c_2}, \beta_{l_2})$				
Parameters of the Overall Bimodal Distribution				
β_l	17.90	14.80	18.36	18.21
β_c	13.09	10.36	13.22	13.33
$stddev(\beta_l)$	1.39		0.92	
$stddev(\beta_c)$	1.39		1.83	
$Cov(\beta_c, \beta_l)$	-1.42		-0.70	
$Probability1$	0.7			0.56
$Probability2$				0.13
$Probability3$				0.13
Loglikelihood		-4298.89	-4276.36	-4272.16

The bimodal distribution is generated by combining two normal distributions, with portions of 70 and 30 percent.
The table compares the mean estimates across 100 datasets to the true parameters. The simulated sample size consists of 2110 observations.
*Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1})$; $(\beta_{l_1}, \beta_{c_2})$; $(\beta_{l_2}, \beta_{c_1})$; $(\beta_{l_2}, \beta_{c_2})$

Table B.5: Parameter Estimates (DGP: (aa) Normal - Large Variance)

	True	SCL	MCLN	LC*	LCN
Fixed Utility Parameters					
$\beta_{c_{sq}}$	1.87	-0.67	1.78	0.73	1.77
β_{cl}	0.42	-0.09	0.32	-0.22	0.36
$\beta_{l_{sq}}$	-1.95	-3.41	-1.96	-1.06	-1.43
fc	2.30	5.03	2.31	2.43	2.29
Parameters of Normal Distribution					
β_{l_1}	36.00	9.50	35.90	20.96	34.57
β_{c_1}	29.00	3.44	29.05	13.77	29.43
β_{l_2}				35.96	34.52
β_{c_2}				35.21	28.65
$stddev(\beta_{l_2})$	7.00		7.03		6.56
$stddev(\beta_{c_2})$	7.05		7.04		7.00
$Cov(\beta_{c_2}, \beta_{l_2})$	-42.00		-40.70		-36.93
$stddev(\beta_{l_2})$					5.98
$stddev(\beta_{c_2})$					6.68
$Cov(\beta_{c_2}, \beta_{l_2})$					-31.17
$Probability1$				0.33	0.64
$Probability2$				0.21	
$Probability3$				0.11	
Loglikelihood		-4217.39	-4106.58	-4158.24	-4102.87

The table compares the mean estimates across 100 datasets to the true parameters.
The simulated sample size consists of 2110 observations.
Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1})$; $(\beta_{l_1}, \beta_{c_2})$; $(\beta_{l_2}, \beta_{c_1})$; $(\beta_{l_2}, \beta_{c_2})$

Table B.6: Parameter Estimates (DGP: (bb) Bimodal - Large Variance)

	True	SCL	MCLN	LC*	LCN
Fixed Utility Parameters					
β_{csq}	1.65	-0.19	0.02	1.06	1.58
β_{cl}	0.42	0.06	-1.79	0.22	0.43
$\beta_{c_{sq}}$	-1.95	-6.80	-5.66	-1.97	-1.60
fc	2.00	8.12	2.01	2.09	2.01
Parameters of Normal Distribution					
β_{l_1}	33.00			26.28	32.69
β_{c_1}	29.00			25.61	28.85
β_{l_2}	47.00			41.47	44.93
β_{c_2}	15.00			11.00	16.05
$stddev(\beta_{l_1})$	3.00				3.11
$stddev(\beta_{c_1})$	3.00				5.00
$Cov(\beta_{c_1}, \beta_{l_1})$					0.80
$stddev(\beta_{l_2})$	3.00				6.46
$stddev(\beta_{c_2})$	3.00				3.75
$Cov(\beta_{c_2}, \beta_{l_2})$					-9.55
Parameters of the Overall Bimodal Distribution					
β_l	37.22	14.38	44.41		
β_c	24.77	2.53	31.09		
$stddev(\beta_l)$	7.09		2.01		
$stddev(\beta_c)$	7.08		12.85		
$Cov(\beta_c, \beta_l)$	-42.7		-15.24		
$Probability1$	0.70			0.49	0.68
$Probability2$				0.10	
$Probability3$				0.23	
Loglikelihood		-4365.53	-4239.41	-4202.53	-4188.61

The table compares the mean estimates across 100 datasets to the true parameters.
The simulated sample size consists of 2110 observations.
Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1}); (\beta_{l_1}, \beta_{c_2}); (\beta_{l_2}, \beta_{c_1}); (\beta_{l_2}, \beta_{c_2})$

Table B.7: Parameter Estimates(DGP: (cc) Discrete Bimodal - Large Variance)

	True	SCL	MCLN	LC*	LCN
Fixed Utility Parameters					
$\beta_{c_{sq}}$	1.65	-0.16	-1.07	1.61	1.59
β_{cl}	0.42	-0.07	-3.18	0.44	0.39
$\beta_{l_{sq}}$	-1.95	-8.57	-6.69	-2.41	-1.78
fc	2.00	7.91	1.98	2.03	2.01
Parameters of Normal Distribution					
β_{l_1}	33.00			32.17	33.59
β_{c_1}	29.00			28.02	29.13
β_{l_2}	47.00			46.44	44.57
β_{c_2}	15.00			14.23	14.98
$stddev(\beta_{l_1})$	0.85				2.12
$stddev(\beta_{c_1})$	0.85				4.74
$Cov(\beta_{c_1}, \beta_{l_1})$					11.28
$stddev(\beta_{l_2})$	0.85				6.17
$stddev(\beta_{c_2})$	0.85				1.23
$Cov(\beta_{c_2}, \beta_{l_2})$					3.37
Parameters of the Overall Bimodal Distribution					
β_l	37.22	17.52	48.11		
β_c	24.78	3.25	35.15		
$stddev(\beta_l)$	6.48		2.73		
$stddev(\beta_c)$	6.48		14.78		
$Cov(\beta_c, \beta_l)$	-31.11		35.96		
$Probability1$	0.70			0.61	0.71
$Probability2$				0.06	
$Probability3$				0.10	
Loglikelihood		-4247.20	-4098.91	-3986.09	-3988.13

The table compares the mean estimates across 100 datasets to the true parameters.
The simulated sample size consists of 2110 observations.
Each probability corresponds to these combinations, respectively:$(\beta_{l_1}, \beta_{c_1}); (\beta_{l_1}, \beta_{c_2}); (\beta_{l_2}, \beta_{c_1}); (\beta_{l_2}, \beta_{c_2})$

Table B.8: Model Fit – True and Predicted Hour Distributions (DGP: Normal Distributions, a & aa)

Hour	True	SCL	MCLN	LC	Diff (ppt)			True	SCL	MCLN	LC	LCN	Diff(ppt)			
Category					SCL	MCLN	LC						SCL	MCLN	LC	LCN
	a. Small Variance							aa. Large Variance								
0	25.71	25.46	25.64	25.05	0.25	0.07	0.66	24.95	24.85	24.90	23.77	24.95	0.10	0.05	1.18	0.00
200	2.23	1.69	2.27	2.29	0.54	0.04	0.06	0.12	0.04	0.14	0.20	0.15	0.08	0.02	0.08	0.03
400	1.38	1.34	1.38	1.44	0.04	0.00	0.06	0.05	0.09	0.05	0.12	0.05	0.04	0.00	0.07	0.00
600	0.89	1.12	0.89	0.94	0.23	0.00	0.05	0.11	0.23	0.12	0.18	0.12	0.12	0.01	0.07	0.01
800	1.33	1.67	1.30	1.38	0.34	0.03	0.05	0.54	0.76	0.57	0.76	0.56	0.22	0.03	0.22	0.02
1000	3.11	3.67	3.12	3.19	0.56	0.01	0.08	1.81	2.02	1.85	2.15	1.85	0.21	0.04	0.34	0.04
1200	6.41	6.89	6.43	6.59	0.48	0.02	0.18	4.21	4.23	4.17	4.17	4.17	0.02	0.04	0.04	0.04
1400	10.27	10.40	10.31	10.43	0.13	0.04	0.16	7.55	7.28	7.54	7.29	7.50	0.27	0.01	0.26	0.05
1600	12.88	12.58	12.88	12.85	0.30	0.00	0.03	10.69	10.18	10.62	10.66	10.60	0.51	0.07	0.03	0.09
1800	13.35	12.81	13.33	13.20	0.54	0.02	0.15	12.80	12.29	12.78	13.29	12.71	0.51	0.02	0.49	0.09
2000	10.85	10.42	10.84	10.73	0.43	0.01	0.12	12.50	12.51	12.52	12.99	12.48	0.01	0.02	0.49	0.02
2200	6.98	6.99	6.94	7.10	0.01	0.04	0.12	10.58	10.98	10.57	10.37	10.55	0.40	0.01	0.21	0.03
2400	3.33	3.47	3.35	3.46	0.14	0.02	0.13	7.33	7.85	7.34	7.33	7.38	0.52	0.01	0.00	0.05
2600	1.05	1.21	1.07	1.13	0.16	0.02	0.08	4.14	4.45	4.13	4.48	4.18	0.31	0.01	0.34	0.04
2800	0.20	0.26	0.21	0.22	0.06	0.01	0.02	1.94	1.78	1.98	1.84	2.01	0.16	0.04	0.10	0.07
3000	0.02	0.03	0.03	0.02	0.01	0.01	0.00	0.68	0.47	0.73	0.40	0.74	0.21	0.05	0.28	0.06

These are the averages over 100 datasets. Number of observations in each dataset is 2110.
These two scenarios refer to the two normal distributions which could also be considered as fully overlapping bimodal distributions.

Table B.9: Model Fit – True and Predicted Hour Distributions (DGP: Bimodal Distributions, b & bb)

Hour	True	SCL	MCLN	LC	Diff (ppt)			True	SCL	MCLN	LC	LCN	Diff(ppt)			
Category					SCL	MCLN	LC						SCL	MCLN	LC	LCN
	b. Small Variance							bb. Large Variance								
0	24.01	23.86	23.89	22.69	0.15	0.12	1.32	25.21	25.06	24.47	24.69	24.83	0.15	0.74	0.52	0.38
200	3.02	2.05	3.00	2.93	0.97	0.02	0.09	1.00	0.11	1.39	1.05	0.98	0.89	0.39	0.05	0.02
400	1.92	1.76	1.97	1.87	0.16	0.05	0.05	0.34	0.23	0.37	0.43	0.35	0.11	0.03	0.09	0.01
600	1.44	1.68	1.50	1.41	0.24	0.06	0.03	0.62	0.53	0.50	0.83	0.57	0.09	0.12	0.21	0.05
800	1.77	2.31	1.83	1.77	0.54	0.06	0.00	1.67	1.38	1.22	1.87	1.59	0.29	0.45	0.20	0.08
1000	3.22	3.97	3.23	3.11	0.75	0.01	0.11	3.01	3.10	2.81	2.78	3.05	0.09	0.20	0.23	0.04
1200	5.66	6.44	5.68	5.41	0.78	0.02	0.25	4.53	5.80	5.19	4.04	4.52	1.27	0.66	0.49	0.01
1400	8.64	9.07	8.72	8.38	0.43	0.08	0.26	6.98	9.05	8.45	6.90	6.79	2.07	1.47	0.08	0.19
1600	11.04	10.90	11.09	10.87	0.14	0.05	0.17	10.83	11.96	11.58	10.57	10.82	1.13	0.75	0.26	0.01
1800	12.23	11.57	12.13	12.34	0.66	0.10	0.11	14.37	13.34	13.75	14.00	14.75	1.03	0.62	0.37	0.38
2000	11.21	10.37	11.03	11.72	0.84	0.18	0.51	14.37	12.25	12.97	14.70	14.63	2.12	1.40	0.33	0.26
2200	8.39	8.07	8.40	9.15	0.32	0.01	0.76	10.39	9.30	9.80	11.21	10.43	1.09	0.59	0.82	0.04
2400	4.83	4.90	4.85	5.37	0.07	0.02	0.54	4.93	5.24	5.34	5.28	4.94	0.31	0.41	0.35	0.01
2600	2.03	2.24	2.03	2.29	0.21	0.00	0.26	1.49	2.09	1.82	1.46	1.45	0.60	0.33	0.03	0.04
2800	0.52	0.68	0.55	0.60	0.16	0.03	0.08	0.25	0.50	0.32	0.19	0.26	0.25	0.07	0.06	0.01
3000	0.07	0.12	0.09	0.09	0.05	0.02	0.02	0.01	0.06	0.03	0.01	0.03	0.05	0.02	0.00	0.02

These are the averages over 100 datasets. Number of observations in each dataset is 2110.
These two scenarios refer to bimodal distributions with partial overlap.

Table B.10: Model Fit – True and Predicted Hour Distributions (DGP: Discrete Bimodal Distributions, c & cc)

Hour Category	True	SCL	MCLN	LC	Diff (ppt) SCL	MCLN	LC	True	SCL	MCLN	LC	LCN	Diff(ppt) SCL	MCLN	LC	LCN
	c. Small Variance							cc. Large Variance								
0	23.82	23.65	23.64	22.51	0.17	0.18	1.31	24.47	24.38	23.33	24.33	24.38	0.09	1.14	0.14	0.09
200	2.85	2.05	2.85	2.73	0.80	0.00	0.12	0.95	0.09	1.47	0.93	0.94	0.86	0.52	0.02	0.01
400	1.86	1.74	1.90	1.77	0.12	0.04	0.09	0.31	0.19	0.39	0.32	0.36	0.12	0.08	0.01	0.05
600	1.40	1.61	1.45	1.34	0.21	0.05	0.06	0.62	0.45	0.52	0.66	0.61	0.17	0.10	0.04	0.01
800	1.76	2.19	1.74	1.66	0.43	0.02	0.10	2.06	1.25	1.18	2.22	1.96	0.81	0.88	0.16	0.10
1000	3.22	3.87	3.24	3.03	0.65	0.02	0.19	3.47	3.03	2.71	3.42	3.50	0.44	0.76	0.05	0.03
1200	5.77	6.46	5.84	5.28	0.69	0.07	0.49	4.09	6.00	5.15	3.67	4.37	1.91	1.06	0.42	0.28
1400	8.82	9.27	8.97	8.35	0.45	0.15	0.47	6.24	9.76	8.85	5.27	6.43	3.52	2.61	0.97	0.19
1600	11.37	11.22	11.42	11.09	0.15	0.05	0.28	11.25	13.09	12.58	10.15	11.46	1.84	1.33	1.10	0.21
1800	12.52	11.90	12.48	12.85	0.62	0.04	0.33	16.41	14.46	15.22	16.62	16.63	1.95	1.19	0.21	0.22
2000	11.32	10.57	11.16	12.11	0.75	0.16	0.79	16.24	12.67	14.18	17.31	16.06	3.57	2.06	1.07	0.18
2200	8.33	8.05	8.30	9.30	0.28	0.03	0.97	10.00	8.76	9.59	10.79	9.57	1.24	0.41	0.79	0.43
2400	4.63	4.71	4.64	5.29	0.08	0.01	0.66	3.31	4.26	4.00	3.65	3.15	0.95	0.69	0.34	0.16
2600	1.84	2.05	1.85	2.12	0.21	0.01	0.28	0.56	1.37	0.79	0.63	0.55	0.81	0.23	0.07	0.01
2800	0.43	0.57	0.45	0.51	0.14	0.02	0.08	0.03	0.23	0.05	0.04	0.04	0.20	0.02	0.01	0.01
3000	0.05	0.08	0.06	0.07	0.03	0.01	0.02	0.00	0.02	0.00	0.00	0.00	0.02	0.00	0.00	0.00

These are the averages over 100 datasets. Number of observations in each dataset is 2110.
These two scenarios refer to bimodal distributions with no overlap.

Appendix C

Chapter 3: Tables

Table C.1: Descriptive Statistics by Paid Childcare Utilisation

Variables	CC=0	CC=1
Employment rate	47.60	100.00
High School Graduates	43.61	62.73
Hourly Wage	19.26 (18.93)	20.73 (22.29)
Number of Children	1.97 (1.92)	1.71 (1.58)
Childcare Cost	-	2973.38 (5518.08)
Annual Hours of Work	1616.09 (1078.21)	1674.24 (1001.16)
Immigrants	18.61	15.49
Bilingual	30.08	44.10
Anglophone	6.32	2.49
Poverty Rate	63.46	18.31
Adults	39.49	23.29
Social Assistance Participation	48.28	3.36
Disposable Income	31233.64 (26710.11)	41658.60 (22950.54)
Observations	7280	4615

Values in parentheses are standard deviations. The sample size is weighted and rounded.

Table C.2: Simulated Benchmark by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	59.71	1531.00	36.98	45.63	50.30	34969.49
School	69.09	1684.07	24.65	33.84	44.89	34398.22

Table C.3: Free Childcare by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	0.890	-0.030	-0.177	2.218	0.054	4.543
School	0.253	0.033	-0.056	1.445	0.242	1.748

Values represent the mean changes in response to the policy.

Table C.4: Cash Transfer by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	1.377	-0.369	-0.262	1.090	-6.456	6.836
School	0.736	-0.217	-0.115	0.346	-4.657	4.460

Values represent the mean changes in response to the policy.

Table C.5: 50% Benefit Reduction Rate by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	1.938	0.559	0.521	1.174	-1.949	1.001
School	1.457	0.178	0.453	0.267	-0.661	0.861

Values represent the mean changes in response to the policy.

Table C.6: EITC by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	0.815	0.167	-0.174	0.590	-5.887	3.548
School	0.760	-0.176	-0.141	0.318	-7.319	3.943

Values represent the mean changes in response to the policy.

Table C.7: Combined Policy by the Age of the Youngest Child

	Employment Rate (ppt)	Hours (pct)	Welfare Participation (ppt)	Childcare Use (ppt)	Poverty rate (ppt)	Disposable Income (pct)
Preschool	3.882	0.407	0.267	4.190	-9.097	9.736
School	2.587	-0.036	0.339	2.007	-8.797	6.872

Values represent the mean changes in response to the policy.

References

Allen, D. W. (1993). Welfare and the family: The canadian experience. *Journal of Labor Economics*, 11(1, Part 2):S201–S223.

Andrén, T. (2003). The choice of paid childcare, welfare, and labor supply of single mothers. *Labour Economics*, 10(2):133–147.

Annabi, N., Boudribila, Y., and Harvey, S. (2013). Labour supply and income distribution effects of the working income tax benefit: a general equilibrium microsimulation analysis. *IZA Journal of Labor Policy*, 2:1–33.

Athreya, K., Reilly, D., and Simpson, N. B. (2014). Single mothers and the earned income tax credit: Insurance without disincentives?

Bajari, P., Fox, J. T., and Ryan, S. P. (2007). Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients. *American Economic Review*, 97(2):459–463.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of political Economy*, 116(4):709–745.

Baker, M., Messacar, D., and Stabile, M. (2023). Effects of child tax benefits on poverty and labor supply: Evidence from the canada child benefit and universal child care benefit. *Journal of Labor Economics*, 41(4):1129–1182.

Bansal, P., Daziano, R. A., and Achtnicht, M. (2018a). Comparison of parametric and semiparametric representations of unobserved preference heterogeneity in logit models. *Journal of choice modelling*, 27:97–113.

Bansal, P., Daziano, R. A., and Achtnicht, M. (2018b). Extending the logit-mixed logit model for a combination of random and fixed parameters. *Journal of choice modelling*, 27:88–96.

Bargain, O. and Doorley, K. (2011). Caught in the trap? welfare's disincentive and the labor supply of single men. *Journal of Public Economics*, 95(9-10):1096–1110.

Bargain, O. and Doorley, K. (2017). The effect of social benefits on youth employment: Combining regression discontinuity and a behavioral model. *Journal of Human Resources*, 52(4):1032–1059.

Barrett, G. F. (2000). The effect of educational attainment on welfare dependence: Evidence from canada. *Journal of Public Economics*, 77(2):209–232.

Bastin, F., Cirillo, C., and Toint, P. L. (2010). Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations. *Transportation science*, 44(4):537–549.

Bauernschuster, S. and Schlotter, M. (2015). Public child care and mothers' labor supply—evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16.

Bettendorf, L. J., Jongen, E. L., and Muller, P. (2015). Childcare subsidies and labour supply—evidence from a large dutch reform. *Labour Economics*, 36:112–123.

Bitler, M. P., Gelbach, J. B., and Hoynes, H. W. (2008). Distributional impacts of the self-sufficiency project. *Journal of Public Economics*, 92(3-4):748–765.

Blau, D. and Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education*, 2:1163–1278.

Blomquist, S. (1996). Estimation methods for male labor supply functions how to take account of nonlinear taxes. *Journal of Econometrics*, 70(2):383–405.

Blundell, R., Duncan, A., McCrae, J., and Meghir, C. (2000). The labour market impact of the working families' tax credit. *Fiscal studies*, 21(1):75–104.

Blundell, R. and Hoynes, H. W. (2004). Has' in-work'benefit reform helped the labor market? In *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, pages 411–460. University of Chicago Press.

Brewer, M., Cattan, S., Crawford, C., and Rabe, B. (2016). Free childcare and parents' labour supply: Is more better?

Brewer, M., Duncan, A., Shephard, A., and Suarez, M. J. (2006). Did working families' tax credit work? the impact of in-work support on labour supply in great britain. *Labour economics*, 13(6):699–720.

Brouillette, D. and Fortin, B. (2008). *Impact de la Prime au travail sur l'effort au travail: une approche expérimentale*. CIRANO.

Busse, A. and Gathmann, C. (2020). Free daycare policies, family choices and child development. *Journal of Economic Behavior & Organization*, 179:240–260.

Cameron, S. V. and Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political economy*, 106(2):262–333.

Card, D. and Robins, P. K. (1996). Do financial incentives encourage welfare recipients to work? evidence from a randomized evaluation of the self-sufficiency project.

Card, D. and Robins, P. K. (2005). How important are "entry effects" in financial incentive programs for welfare recipients? experimental evidence from the self-sufficiency project. *Journal of Econometrics*, 125(1-2):113–139.

Card, D. E. and Hyslop, D. R. (2006). *Understanding the Dynamic Effects of the Self-Sufficiency Project Applicant Study*. Social Research and Demonstration Corporation.

Carta, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children's outcomes: evidence from italy. *Journal of Public Economics*, 158:79–102.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into american public schools. *Journal of Human resources*, 44(1):140–170.

Choi, E. J. (2018). Evaluating a structural model of labor supply and welfare participation: Evidence from state welfare reform experiments. *Available at SSRN 3289937*.

Chyi, H. (2012). The 1993 eitc expansion and low-skilled single mothers' welfare use decision. *Applied Economics*, 44(13):1717–1736.

Clavet, N.-J., Duclos, J.-Y., and Lacroix, G. (2013). Fighting poverty: Assessing the effect of guaranteed minimum income proposals in quebec. *Canadian Public Policy*, 39(4):491–516.

Collischon, M., Kühnle, D., and Oberfichtner, M. (2020). Cash-for-care, or caring for cash? the effects of a home care subsidy on maternal employment, childcare choices, and children's development.

Company, M. . and Manyika, J. (2017). *Technology, jobs, and the future of work*. McKinsey Insights.

Connelly, R. and Kimmel, J. (2003). The effect of child care costs on the employment and welfare recipiency of single mothers. *Southern Economic Journal*, 69(3):498–519.

Del Boca, D. et al. (2015). The impact of child care costs and availability on mothers' labor supply. *Erişim adresi: https://www. carloalberto. org/wp-content/uploads/2018/11*, (399).

Eckstein, Z. and Wolpin, K. I. (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, 67(6):1295–1339.

Eissa, N. and Hoynes, H. W. (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of public Economics*, 88(9-10):1931–1958.

Eissa, N. and Hoynes, H. W. (2006). Behavioral responses to taxes: Lessons from the eitc and labor supply. *Tax policy and the economy*, 20:73–110.

Fitzpatrick, M. D. (2012). Revising our thinking about the relationship between maternal labor supply and preschool. *Journal of Human Resources*, 47(3):583–612.

Flood, L., Hansen, J., and Wahlberg, R. (2004). Household labor supply and welfare participation in sweden. *Journal of Human Resources*, 39(4):1008–1032.

Fortin, B., Fougère, D., and Lacroix, G. (1999). The impact of government-sponsored training programs on labour market transition. Technical report, CIRANO.

Fosgerau, M. and Bierlaire, M. (2007). A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research Part B: Methodological*, 41(7):784–794.

Fosgerau, M. and Mabit, S. L. (2013). Easy and flexible mixture distributions. *Economics Letters*, 120(2):206–210.

Fox, J. T., Kim, K. I., Ryan, S. P., and Bajari, P. (2011). A simple estimator for the distribution of random coefficients. *Quantitative Economics*, 2(3):381–418.

Francesconi, M. and Van der Klaauw, W. (2007). The socioeconomic consequences of "in-work" benefit reform for british lone mothers. *Journal of Human Resources*, 42(1):1–31.

Gathmann, C. and Sass, B. (2018). Taxing childcare: Effects on childcare choices, family labor supply, and children. *Journal of labor Economics*, 36(3):665–709.

Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *American*

Economic Review, 92(1):307–322.

Gilbert, L., Kamionka, T., and Lacroix, G. (2011). The impact of government-sponsored training programs on the labor market transitions of disadvantaged men. *ADVANCES IN ECONOMETRICS-THEORY AND APPLICATIONS*, page 47.

Givord, P. and Marbot, C. (2015). Does the cost of child care affect female labor market participation? an evaluation of a french reform of childcare subsidies. *Labour Economics*, 36:99–111.

Goux, D. and Maurin, E. (2010). Public school availability for two-year olds and mothers' labour supply. *Labour Economics*, 17(6):951–962.

Gregg, P., Harkness, S., and Smith, S. (2009). Welfare reform and lone parents in the uk. *The economic journal*, 119(535):F38–F65.

Haan, P. (2006). Much ado about nothing: Conditional logit vs. random coefficient models for estimating labour supply elasticities. *Applied Economics Letters*, 13(4):251–256.

Haan, P. and Wrohlich, K. (2011). Can child care policy encourage employment and fertility?: Evidence from a structural model. *Labour Economics*, 18(4):498–512.

Haeck, C., Lefebvre, P., and Merrigan, P. (2015). Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics*, 36:137–157.

Hansen, J., Kristensen, N., and Andersen, H. L. (2020). The bottom 20%: Early career paths of adolescents with low gpa.

Hansen, J. and Liu, X. (2015). Estimating labour supply responses and welfare participation: Using a natural experiment to validate a structural labour supply model. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(5):1831–1854.

Havnes, T. and Mogstad, M. (2011). Money for nothing? universal child care and maternal employment. *Journal of Public Economics*, 95(11-12):1455–1465.

Hotz, V. J. (2003). The earned income tax credit. In *Means-tested transfer programs in the United States*, pages 141–198. University of Chicago press.

Hoynes, H. W. (1996). Work, welfare, and family structure: what have we learned?

Huebener, M., Pape, A., and Spiess, C. K. (2020). Parental labour supply responses to the abolition of day care fees. *Journal of Economic Behavior & Organization*, 180:510–543.

Huffman*, S. K. and Jensen, H. H. (2005). Linkages among welfare, food assistance programmes and labour supply: evidence from the survey of programme dynamics. *Applied economics*, 37(10):1099–1113.

Ito, H. and Yamamoto, I. (2022). Do comprehensive and diverse childcare services affect women's labour supply and well-being? *Applied Economics Letters*, 29(2):173–178.

Jones, S. R., Lange, F., Riddell, W. C., and Warman, C. (2020). Waiting for recovery: The canadian labour market in june 2020. *Canadian Public Policy*, 46(S2):S102–S118.

Keane, M. and Moffitt, R. (1998). A structural model of multiple welfare program participation and labor supply. *International economic review*, pages 553–589.

Keane, M. and Wasi, N. (2013). Comparing alternative models of heterogeneity in consumer choice behavior. *Journal of Applied Econometrics*, 28(6):1018–1045.

Koebel, K. and Schirle, T. (2016). The differential impact of universal child benefits on the labour supply of married and single mothers. *Canadian Public Policy*, 42(1):49–64.

Kornstad, T. and Thoresen, T. O. (2007). A discrete choice model for labor supply and childcare. *Journal of Population economics*, 20:781–803.

Kosonen, T. (2014). To work or not to work? the effect of childcare subsidies on the labour supply of parents. *The BE Journal of Economic Analysis & Policy*, 14(3):817–848.

Lacroix, G. and Brouillette, D. (2011). Assessing the impact of a wage subsidy for single parents on social assistance. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4):1195–1221.

Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from canada. *Journal of Labor Economics*, 26(3):519–548.

Lefebvre, P., Merrigan, P., and Verstraete, M. (2009). Dynamic labour supply effects of childcare subsidies: Evidence from a canadian natural experiment on low-fee universal child care. *Labour Economics*, 16(5):490–502.

Lemieux, T. and Milligan, K. (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics*, 142(2):807–828.

Lemieux, T., Milligan, K., Schirle, T., and Skuterud, M. (2020). Initial impacts of the covid-19 pandemic on the canadian labour market. *Canadian Public Policy*, 46(S1):S55–S65.

Löffler, M., Peichl, A., and Siegloch, S. (2018). The sensitivity of structural labor supply estimations to modeling assumptions.

Lundin, D., Mörk, E., and Öckert, B. (2008). How far can reduced childcare prices push female labour supply? *Labour Economics*, 15(4):647–659.

McFadden, D. and Train, K. (2000). Mixed mnl models for discrete response. *Journal of applied Econometrics*, 15(5):447–470.

Messacar, D. (2021). Employment, child care spending, and child tax benefits. *National Tax Journal*, 74(2):553–575.

Milligan, K. (2016). Canadian tax and credit simulator. *Database, software and documentation, Version*, 2.

Milligan, K. and Stabile, M. (2007). The integration of child tax credits and welfare: Evidence from the canadian national child benefit program. *Journal of public Economics*, 91(1-2):305–326.

Moffette, F., Boccanfuso, D., Richard, P., and Savard, L. (2013). Estimating the impact of the québec's work incentive program on labour supply: An ex post microsimulation analysis. Technical report.

Moffitt, R. (1983). An economic model of welfare stigma. *The American economic review*, 73(5):1023–1035.

Moffitt, R. (1986). Work incentives in the afdc system: An analysis of the 1981 reforms. *The American Economic Review*, 76(2):219–223.

Moffitt, R. A. (2003). The negative income tax and the evolution of us welfare policy. *Journal of economic perspectives*, 17(3):119–140.

Naz, G. (2004). The impact of cash-benefit reform on parents' labour force participation. *Journal of Population Economics*, 17:369–383.

Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from spain. *Labour Economics*, 36:124–136.

Pacifico, D. (2013). On the role of unobserved preference heterogeneity in discrete choice models of labour supply. *Empirical Economics*, 45:929–963.

Ribar, D. C. (1995). A structural model of child care and the labor supply of married women. *Journal of labor Economics*, 13(3):558–597.

Riddell, C. and Riddell, W. C. (2014). The pitfalls of work requirements in welfare-to-work policies: Experimental evidence on human capital accumulation in the self-sufficiency project. *Journal of Public Economics*, 117:39–49.

Schirle, T. (2015). The effect of universal child benefits on labour supply. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(2):437–463.

Schøne, P. (2004). Labour supply effects of a cash-for-care subsidy. *Journal of Population Economics*, 17:703–727.

Soest, A. v., Das, M., and Gong, X. (2002). A structural labour supply model with flexible preferences. *Journal of Econometrics*, 107(1–2):345–374.

Statistics Canada (2019). Table 37-10-0130-01 educational attainment of the population aged 25 to 64, by age group and sex, organisation for economic co-operation and development (OECD), Canada, provinces and territories. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3710013001.

Statistics Canada, Census (2016). Dictionary, census of population, 2016 low-income measure. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/fam021-eng.cfm.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Train, K. (2016). Mixed logit with a flexible mixing distribution. *Journal of choice modelling*, 19:40–53.

Train, K. E. (2008). Em algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1):40–69.

van Soest, A. (1995). Structural models of family labor supply: a discrete choice approach. *Journal of human Resources*, pages 63–88.

Van Soest, A. and Das, M. (2001). Family labor supply and proposed tax reforms in the netherlands. *De Economist*, 149(2):191–218.

van Soest, A., Das, M., and Gong, X. (2002). A structural labour supply model with flexible preferences. *Journal of econometrics*, 107(1-2):345–374.

Wrohlich, K. (2011). Labor supply and child care choices in a rationed child care market.

Yang, T.-T. (2018). Family labor supply and the timing of cash transfers: Evidence from the earned income tax credit. *Journal of Human Resources*, 53(2):445–473.