# Empirical essays on human capital: determinants, returns and components

Nagham Sayour

Department of Economics

McGill University Montreal, Quebec

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Ph.D. in Economics.

© Nagham Sayour 2016

## Dedication

I dedicate this thesis to my parents, husband and daughter.

## Abstract

This thesis is comprised of three empirical essays on the theme of human capital. The essays use natural and laboratory experiments to study the determinants, returns and components of human capital. We first consider the determinants of human capital by studying the effects of maternal care as a determinant of children's human capital. Then we investigate the returns to human capital by studying the effects of immigration policies on immigrants' characteristics and labour market outcomes. Lastly, we examine specific components of human capital through an experiment on non-cognitive skills and preferences.

The first essay estimates the causal impact of maternal care on the developmental outcomes of children aged 2-3 years using a parental leave reform implemented in Canada at the end of 2000 as an exogenous variation to maternal care. The reform increased the time mothers spend with their newborns by 3 months without affecting their income net of taxes, transfers and child care costs. Using the Canadian National Longitudinal Survey of Children and Youth, we employ a difference-in-differences methodology to compare children with a sibling born after the reform to those with a sibling born before the reform, relative to children of the same birth cohorts who did not have a younger sibling in the period surrounding the reform. We find that treated children enjoy a 16 percent increase in the time they spend with their mothers, with maternal care crowding out informal care. The increase in maternal care does not translate into better cognitive, non-cognitive or health outcomes in the short-run or the medium-run.

The second essay uses a natural experiment to study the effects of a change in the point system, a system that selects immigrants based on specific observable characteristics, on immigrants' characteristics and labor market outcomes. Specifically, in 2001, Quebec changed its point system, by increasing the points for education and French language and decreasing the points for a subjective category "adaptability". The objective of the reform was to increase the number of French-speaking immigrants without deteriorating their labor market performance. Using a difference-in-differences and triple differences methodology, we show that, compared to immigrants to the Rest of Canada, immigrants to Quebec after the reform hold more bachelor's degrees and know more French than immigrants to Quebec before the reform. However, this does not translate into better labor market outcomes. This essay shows how point systems can be used to shape the immigrant workforce according to policy goals.

Non-cognitive skills are a recently incorporated component of human capital in the economics literature. In the third essay, we contribute to this literature through a laboratory experiment on personality traits and risk and ambiguity preferences. We also study the effects of personality traits prevalence in a group on the decision making of each group member. In the experiment, subjects reveal their risk and ambiguity preferences through lottery choices. They then participate in an unstructured group chat. Afterwards, they are given the chance to revise their initial lottery choices. Results show that personality traits affect risk and ambiguity preferences before the chat. Specifically, conscientiousness is negatively related to risk and ambiguity aversion and agreeableness is negatively related to ambiguity aversion. We also show that the probability of changing decisions after the chat is affected by the individual's non-cognitive traits but not by the traits of the other group members.

## Abrégé

Cette thèse est composée de trois essais empiriques sur le thème du capital humain. Les essais utilisent des expériences naturelles et en laboratoire pour étudier les déterminants, les rendements et les composantes du capital humain. Nous considérons d'abord les déterminants du capital humain en étudiant les effets du temps maternel en tant que déterminant du capital humain des enfants. Ensuite, nous étudions les rendements du capital humain en étudiant les effets des politiques d'immigration sur les caractéristiques des immigrants et leur intégration dans le marché du travail. Enfin, nous examinons des composantes spécifiques du capital humain à l'aide d'une expérience sur les compétences non-cognitives et les préférences.

Le premier essai évalue l'impact causal du temps maternel sur le développement des enfants âgés de 2-3 ans en utilisant une réforme du congé parental mis en œuvre au Canada à la fin de 2000, comme une variation exogène aux temps maternel. La réforme a augmenté le temps que les mères passaient avec leurs nouveau-nés de 3 mois sans affecter leur revenu net des impôts, les transferts et les frais de garde d'enfants. En utilisant l'Enquête Longitudinal Nationale sur les Enfants et les Jeunes, nous employons une méthodologie de différence dans les différences pour comparer les enfants avec un frère ou une sœur né(e) après la réforme à ceux avec un frère ou une sœur né(e) avant la réforme, par rapport aux enfants de même cohortes de naissance n'ayant pas un frère ou une sœur né(e) dans la période entourant la réforme. Nous constatons que les enfants traités bénéficient d'une augmentation de 16 pourcent du temps passé avec leurs mères. Le temps maternel supplémentaire n'améliore pas les résultats cognitifs, non cognitifs ou de santé des enfants à court terme ni à moyen terme.

Le deuxième essai utilise une expérience naturelle pour étudier les effets d'un changement dans le système de points, un système qui sélectionne les immigrants en fonction de caractéristiques spécifiques observables, sur les caractéristiques des immigrants et sur leur intégration au marché du travail. Plus précisément, en 2001, le Québec a changé son système de points, en augmentant les points pour l'éducation et la langue française et en diminuant les points pour une catégorie subjective d'«adaptation». L'objectif de la réforme était d'augmenter le nombre d'immigrants maitrisant la langue française sans détériorer leur performance sur le marché du travail. En utilisant la méthodologie de la différence dans les différences et des différences triples, nous montrons que la réforme a attiré des immigrants françophones plus scolarisés mais que la réforme n'a eu aucun

effet sur leur intégration au marché du travail. Cet essai montre comment les systèmes de points peuvent être utilisés pour sélectionner les immigrants qualifiés en fonction des objectifs politiques.

Les compétences non-cognitives ont été récemment intégrées dans la littérature économique sur le capital humain. Dans le troisième essai, nous contribuons à cette littérature grâce à une expérience en laboratoire sur les traits de personnalité et les préférences au risque et à l'ambiguïté. Nous étudions également les effets de la prévalence des traits dans un groupe sur la prise de décision de chaque membre du groupe. Dans l'expérience, les sujets révèlent leurs préférences au risque et à l'ambiguïté par leurs choix face à une loterie. Ils participent ensuite à une discussion de groupe en ligne, non structurée. Ensuite, on leur donne une chance de changer leurs choix initiaux de loterie. Les résultats montrent que les traits de personnalité affectent les préférences au risque et à l'ambiguïté avant la discussion. Plus précisément, la conscience est négativement liée à l'aversion au risque et l'agréabilité est négativement liée à l'aversion à l'ambiguïté. Nous montrons aussi que la probabilité de changer les décisions après la discussion est affectée par les traits noncognitifs de l'individu, mais pas par les traits des autres membres du groupe.

## Acknowledgments

I am grateful to my supervisors Professor Matthieu Chemin and Professor Sonia Laszlo for their guidance, advice, encouragement and understanding throughout all the stages of my PhD studies. I especially thank them for helping me shape my ideas to develop interesting research questions.

Special thanks to my thesis committee members Professors Catherine Haeck, Fabian Lange and Markus Poschke for their valuable comments and suggestions. Many thanks to Professors Saraswata Chaudhuri for writing my recommendation letters and for Professors Theodore Papageorgiou and Francesco Amodio for their guidance and feedback during the job market. I thank as well all the Professors who taught me during my doctoral studies: Christopher Ragan, Ngo Van Long, John Galbraith, Licun Xue, Jean Marie-Dufour, Franque Grimard, Hassan Benchekroun, Francisco Ruge-Murcia and Damba Lkhagvasuren. I thank Judy Dear, Lisa Stevenson and Angela Fotopoulos for their administrative assistance.

For the financial support, I thank the following: CIREQ-McGill Doctoral Fellowship, McCall MacBain Fellowship, McGill University through the Graduate Excellence Award, Principal Graduate Fellowship and Provost Graduate Fellowship.

I thank all my friends at the Economics department, specifically, Irakli Japaridze, Jean-Francois Mercier and Douglas Barthold for their support and helpful discussions about my research.

I am indebted to my parents, Fouad and Sana, for their unconditional love and support. Heartfelt thanks to my husband, Wissam, for his understanding, patience and belief in me. Lastly, a big thank you to my source of inspiration, my daughter, Yasmeen.

## **Acknowledgements of Chapters**

The analysis presented in the first two essays was conducted at the Quebec Inter-University Centre for Social Statistics (QICSS) which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS 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 - Société et culture (FRQSC), the Fonds de recherche du Québec - Santé (FRQS) and the Quebec universities. The views expressed in the papers are those of the authors, and not necessarily those of the CRDCN or its partners.

For the first essay, I thank the participants of the McGill-CIREQ seminar.

For the second essay, the authors thank two anonymous referees for their valuable comments and suggestions as well as the participants of the McGill-CIREQ seminar.

For the third essay, the authors thank Juan Robledo, Kaywana Raeburn, Anastasia Mshvidobadze, Dina Tasneem and Naureen Fatema for research assistance. We acknowledge the Centre for Interuniversity Research and Analysis on Organizations and the Fonds de la Recherche sur la Société et la Culture for funding.

## **Contribution of Authors**

This thesis contains three original empirical essays. Two of them are co-authored. The work in this thesis is primarily my own.

I claim the sole-authorship of the first essay.

The second essay is co-authored with Professor Matthieu Chemin. I conducted all the empirical analysis at the Quebec Inter-University Center for Social Sciences. I reviewed the literature and wrote the preliminary draft. The content was finalized with a close guidance from Professor Chemin who also framed the research question. The paper is published in the *Journal of Population Economics* (2016), DOI: 10.1007/s00148-016-0594-z.

The third essay is co-authored with Professor Jim Engle-Warnick and Professor Sonia Laszlo. The data used originates from an experiment for Engle, Engle-Warnick and Laszlo (2011). I conducted the literature review, the empirical work and did the writing, with guidance and input from both Professors Engle-Warnick and Laszlo.

# **Table of Contents**

Dedication	ii
Abstract	iii
Abrégé	v
Acknowledgments	vii
Acknowledgements of Chapters	viii
Contribution of Authors	ix
1. Introduction	1
2. The Impact of Maternal Care on Child Development: Evidence from Sibling Effects of a Parental Leave Expansion	s <b>Spillover</b> 6
2.1 Introduction	6
2.2 The Reform	
2.3 Identification Strategy	
2.4 Data	
2.4.1. Sample	
2.4.2 Variables Description	
2.4.3 Descriptive Statistics	
2.5 Methodology	
2.5.1 Wild-Cluster Bootstrap	
2.5.2. Multiple Outcomes Test	
2.6 Results	
2.6.1 Changes in Non-Maternal Care	
2.6.2 Changes in Child Non-Cognitive and Pre-Cognitive Outcomes	
2.6.3 Changes in Child Health Outcomes	
2.6.4 Changes in Parent and Family Outcomes	
2.7 Robustness Checks	
2.7.1 Baseline Results	
2.7.2 Common Time Trend Test	
2.8 Medium-Run Results	
2.9 Heterogeneous Effects	
2.9.1 Child Gender	
2.9.2 Mother's Education	

2.9.3 Household Income	26
2.9.4 Intensity of Early Non-Maternal Care	27
2.10 Moving Beyond the Intention to Treat Estimates	27
2.11 Conclusion	28
Tables	30
Figures	45
Appendix: NLSCY Scores	47
Connecting Text	51
3. The Effects of a Change in the Point System on Immigration: Evidence from the 200 Quebec Reform	1 52
3.1 Introduction	52
3.2 The Point System	55
3.2.1. Description	55
3.2.2 Existing Literature	56
3.3 Identification Strategy	57
3.3.1 The 2001 QC Reform	57
3.3.2 Exogeneity of the Reform	60
3.4 Data	60
3.5 Methodology	63
3.6 Results	64
3.6.1 Immigrants' Characteristics	64
3.6.2 Labour Market Outcomes	65
3.7 Triple Differences	66
3.8 Discussion	68
3.9 Conclusion	69
Tables	71
Figures	77
Appendix 1	78
Appendix 2	79
Appendix 3	82
Appendix 4	86
Appendix 5	88
Connecting Text	89
4. Experimental Evidence on Personality Traits and Preferences	90
4.1 Introduction	90

4.2 Literature Review	
4.3 Experimental Design	
4.3.1 Overview	
4.3.2 Stage 1	
4.3.3 Stage 2	
4.3.4 Stage 3	
4.4 Experimental Procedure	
4.5 Data Description	
4.5.1 Descriptive Statistics	
4.5.2 Choices in the Risk and Ambiguity Instruments	
4.5.3 The Big-Five Personality Traits	
4.6 Results	
4.6.1 Pre-Chat Analysis	
4.6.2 Post-Chat Analysis	
4.7 Conclusion	
Tables	
Figures	
Appendix 1: The Big Five Personality Traits	
Appendix 2: Example Chat	
5. Conclusion	
References	

## **Chapter 1**

## Introduction

Human capital is a central topic in economics since it plays an imperative role in economic growth (Nelson and Phelps 1966; Mankiw, Romer and Weil 1992). This thesis uses natural and laboratory experiments to contribute to the literature on the determinants, returns and components of human capital. We first study maternal care as a determinant of children's human capital, then examine the returns to immigrants' human capital. Lastly, we delve into a specific component of human capital that has recently been incorporated in economic studies and provide experimental evidence on non-cognitive skills and preferences.

An essential theme in the human capital literature is the acquisition, investment and determinants of human capital at different stages in the life cycle. Many studies have shown that the developmental outcomes of children between the ages of 0 and 5 predict economic success in adulthood (Thomas and Strauss 1997; Bloom, Canning and Sevilla 2004; Heckman, Stixrud, and Urzua 2006; Mueller and Plug 2006; Almond and Currie 2011; Lindqvist and Vestman 2011). Parental involvement is considered a crucial determinant of children's human capital. The first essay, **"The Impact of Maternal Care on Child Development: Evidence from Sibling Spillover Effects of a Parental Leave Expansion"**, contributes to this literature by studying the causal effects of maternal care (one component of parental involvement) on the human capital of children aged 2-3 years old.

The rise of maternal-employment over the last half-century has had important policy implications for early child-care arrangements. Advocates of child-care argue that good quality child-care is beneficial for child development whereas opponents claim that children's cognitive, non-cognitive and health outcomes are best fostered if care is provided by mothers. In addition, the 2-3 years old age group is of particular importance since parents exert high degree of control on the decision of using non-maternal care. The decision to use non-maternal care for younger and older children is usually affected by parental leave policies and compulsory education laws, respectively. Identifying the effects of maternal care on the developmental outcomes of this age group is crucial due to a lack of empirical consensus, which further fuels the debate. Good quality

child-care has been shown to positively affect the development of at-risk children (Karoly et al. 1998), however, less is known for the general population (Baker and Milligan 2010).

Identifying the causal impact of maternal care on the 2-3 years old age group is challenging given that most policies target infants rather than the age group of interest. We overcome this challenge by employing a novel identification strategy. We use a parental leave reform implemented in Canada at the end of 2000 as an exogeneous variation to the time mothers spend with their older children. The reform granted eligible mothers giving birth on December 31, 2000 onwards, up to 12 months of partially paid parental leave. Mothers delivering before the reform were only eligible for 6 months. Baker and Milligan (2010) show that eligible mothers increased the time they spend with their newborns after the reform by 3 months. They also show that the reform did not negatively impact mothers' income net of taxes, benefits and child-care costs. Using the National Longitudinal Survey of Children and Youth, we employ a difference-in-differences methodology to compare children aged 2-3 years old with a sibling born after the reform against those with a sibling born before the reform, relative to children of the same birth cohorts who did not have a sibling in the period surrounding the reform.

We find that treated children enjoy a 16 percent increase in the time spent with their mothers, mainly due to a decrease in the time spent in informal care (care provided by a relative or non-relative in the child's house or in the other person's house) rather than center-based care. The increase in maternal care does not translate into better cognitive, non-cognitive or health outcomes for the children in the short-run or the medium-run.

This essay contributes to the literature on the determinants of children's human capital by showing that an increase in maternal care induced by a parental leave expansion does not have a significant effect on the developmental outcomes of children aged 2-3 years old.

A second fundamental theme in the human capital literature is the formation and returns of immigrants' human capital. Immigrants are a key component of labour markets in several countries such as Canada, the United States and Australia. For example, in Canada, immigrants account for 19 percent of the population. Research in this field of economics has tackled many interesting topics such as the economic assimilation of immigrants into the host country's labour markets (Baker and Benjamin 1994, Grant 1999), the returns to human capital acquired in the host and home countries (Bratsberg and Ragan 2002, Friedberg 2000) and the role of immigration policies (Green and Green 1995; Beach, Worswick and Green 2011). The second essay, **"The Effects of a** 

Change in the Point System on Immigration: Evidence from the 2001 Quebec Reform", contributes to this literature by studying the effect of immigration policies, specifically the point system, on the stock of immigrants' human capital (education and language) and its return through employment and earnings.

The point system is a color-blind policy used to evaluate applicants for immigration by assigning points to specific observable characteristics such as age, education and language. It was initiated in Canada and is used in many other countries such as Australia, New Zealand and the United Kingdom. The United States and the European Union are actively considering introducing the point system (Aydemir 2011). A major advantage of the point system is that governments can set the points in a way that attracts immigrants with specific characteristics. However, skeptics argue that some fundamental characteristics, such as geographic proximity, returns to skills and ability, determine immigrants' composition, rather than a change in the point system (Borjas 1991; Jasso and Rosenzweig 2008). Given this controversy, it is crucial to determine whether changes in the point system affect immigrants' composition and integration into the labour market.

Changes in the point system usually occur at the national level which makes identifying their causal impact on immigrants' composition challenging. In our essay, we use the fact that Quebec (QC) has a different point system than the one used in the Rest of Canada (ROC). In 2001, QC drastically changed its point system by allocating more points for the French language and education, specifically bachelors' degree, and less points for a subjective characteristic "adaptability" that proxies immigrants' personal qualities and motivation. The objective of the reform was to increase the number of French-speaking immigrants without deteriorating their labor market outcomes (MICC 2000). Meanwhile, the point system in ROC was unchanged. We use this intra-national change in the point system and apply a difference-in-differences and a triple differences analysis on the 2006 Canadian Census microdataset to compare immigrants who immigrated to QC before and after 2001, relative to those who immigrated to the ROC.

We find that, relative to immigrants in the ROC, immigrants to QC after 2001 are 6 percentage points more likely to hold a Bachelor's degree and 4 percentage points more likely to speak French only (as opposed to speaking both French and English) than immigrants to QC before the reform. In line with the reform objective, we find no deterioration in immigrants' labor market outcomes measured in terms of employment and earnings. Our results show that the point system can be changed to attract the intended pool of immigrants and thus can be used to shape immigration.

In the last essay, we delve into a recently incorporated component of human capital in the economics literature. The early literature on human capital concentrates on the role of acquiring and investing in education and cognitive skills in human capital formation and labour market returns (ter Weel 2008). However, recently, a growing line of literature is stressing the importance of non-cognitive skills. Heckman, Stixrud and Urzua (2006), Heckman, Humphries and Mader (2011) and Heckman and Kautz (2012) show that non-cognitive skills affect behavioral and labour market outcomes such as teenage pregnancy and marriage, schooling and employment. They argue that some interventions such as the Perry Preschool Program have a long lasting effect on labour market outcomes even though they do not have an effect on children's IQ. They show that the channel through which those programs are affecting labour market outcomes is a permanent change in children's non-cognitive skills, mainly an improvement in conscientiousness and agreeableness traits. Other studies document the crucial effect of non-cognitive skills on several economic outcomes such as health behavior (Chiteji 2010), gender pay gap (Grove, Hussey and Jetter 2011) and gender gap in high school education (Jacob 2002).

The last essay, **"Experimental Evidence on Personality Traits and Preferences"**, uses a laboratory experiment to study the relationship between personality traits, measured by the Big-Five attributes, and individual risk and ambiguity preferences. We also study the effects of personality traits prevalence in a group on the decision making of each group member. Studying the relationship between personality traits and preferences is critical, since both influence the same economic outcomes such as educational attainment, wages and employment (Guiso and Paiella 2005, Eckel et al 2005, Heckman et al 2006, Muller and Plug 2006, Lindqvist and Vestman 2011). However, until now, this relationship is still understudied (Borghans et al 2008, Preevo and ter Weel 2015). Our essay contributes to filling this gap in the literature since it considers for the first time, to our knowledge, the relationship between individual decision making under uncertainty and the personality traits of other group members. The latter contributes to the peer effects literature.

In the experiment, subjects reveal their risk and ambiguity preferences through lottery choices. They then participate in an unstructured internet-based chat in small groups consisting of three individuals. Afterwards, they are given the chance to revise their decisions. We use chat content analysis to measure subjects' personality traits. Results show that personality traits are correlated with risk and ambiguity preferences before the chat. Specifically, conscientiousness is negatively related to risk and ambiguity aversion and agreeableness is negatively related to ambiguity aversion. We also show that the probability of changing decisions after the chat is affected by the individual's personality traits but not by the traits of other group members. The latter only affects the direction of the change. We find an asymmetry in the effect of the traits of other group member, since they only affect the probability of becoming more risk or ambiguity averse but not the probability of becoming less risk or ambiguity averse.

## Chapter 2

## The Impact of Maternal Care on Child Development: Evidence from Sibling Spillover Effects of a Parental Leave Expansion

## **2.1 Introduction**

The rise of maternal employment over the last half-century has had important policy implications for early child-care arrangements. Child-care advocates propose that the lack of affordable child-care can hinder a woman's ability to re-enter the job market, whereby constraining her long term earnings potential. They also argue that non-maternal child-care positively influences child development. Meanwhile, opponents suggest that cognitive, non-cognitive and health growth of children is best fostered when care is provided by mothers. Parental, and especially maternal, leave as well as child-care policies are in the cross-hairs of this debate.

This paper contributes to the debate by examining the effects of maternal child-care on the developmental outcomes of children who are not young enough to be directly affected by parental leave policies but not old enough to attend school<sup>1</sup>. Identifying the effects of maternal care on the developmental outcomes of this age group is crucial since early developmental outcomes have been proven to predict economic success in adulthood (Thomas and Strauss 1997; Bloom, Canning and Sevilla 2004; Heckman, Stixrud, and Urzua 2006; Mueller and Plug 2006; Almond and Currie 2011; Lindqvist and Vestman 2011). A lack of empirical consensus on the effects of maternal care and maternal employment on child development for this age group further fuels the debate. Evidence shows that high quality child-care is beneficial for at risk children (Karoly et al. 1998), but much less is known for the general population (Baker and Milligan 2010). Similarly, most studies show that maternal employment in the first year of child's life is detrimental for child development (Baum 2003, 2004; Ruhm 2004), however, the evidence for older children is mixed

<sup>&</sup>lt;sup>1</sup> The parental decision of using childcare arrangements for this age group is usually not influenced by governmental policies. Parental leaves and compulsory schooling laws play an important role in parents' decision of using non-maternal care for younger and older children, respectively.

(Waldfogel, Han and Brooks-Gunn 2002; Rhum 2008). This lack of consensus manifests itself in striking variations in family policies for these in-between years across OECD countries<sup>2</sup>.

Given that most policies target infants, identifying the causal effects of maternal care on older children is challenging. Typical studies in the literature compare children whose parents have chosen to use non-maternal care with those whose parents have not, while controlling for a wide range of child and parents' observable characteristics. To the extent that confounding factors are not controlled for, this results in biased estimates (Rhum, 2004). In this study, we exploit a reform to the parental leave system implemented in Canada at the end of 2000 which provided eligible mothers delivering on December 31, 2000 onwards, up to 12 months of partially-paid parental leave compared to 6 months for mothers giving birth before that date. Empirically, eligible mothers spent 3 additional months at home with their newborns after the reform. The reform did not negatively affect their income net of taxes, transfers, and child-care costs (Baker and Milligan 2010). Using a difference-in-differences analysis, we exploit the variation induced by this reform to study the effects of maternal care on cognitive, non-cognitive and health outcomes of children aged 2-3 years old. The treatment group consists of children with siblings born after 2000 whose parents were eligible for the extended parental leave. By contrast, the control group consists of children with siblings born before 2001 whose parents were not eligible for the extended parental leave. We examine the impact of the reform on outcomes across these treatment and control groups and also against children of the same birth cohorts who did not have a younger sibling in the period surrounding the reform. The first difference controls for siblings' effects and the second one controls for cohorts' effects and any time trend in children's developmental outcomes.

Our results show that the reform reduced the time the newborns' older siblings spend in nonmaternal care by 4.5 hours per week. This means that time spent in non-maternal care declined by 16 percent among the treated children due to the reform, showing that parental leave might have spillover effects on the newborns' older siblings that are not factored in the design of such policies. The decrease is mainly coming from a 27 percent decrease in the time spent by children in informal care (care provided by a relative or non-relative in the child's house or in the other's person house)

 $<sup>^2</sup>$  For example, on the one hand, France and Belgium provide free pre-kindergarten for children starting at the ages of 2 and 2.5 years respectively (Caille 2001). Quebec offers heavily subsidized daycares starting at the age of 0 and up to the age of 4 (Baker, Grueber and Milligan 2008). On the other hand, Norway provides an incentive for parents to take their children aged less than 3 years old out of daycares (Bettinger, Haegeland, and Rege 2014), and many European countries, such as Spain, Germany and Sweden, offer parental leaves that extend beyond the second year of child's life (Ray 2008).

rather than a change in the time spent in center based care (daycare centers or preschools). However, the increase in maternal care does not translate into better child developmental outcomes in the short-run or the medium-run. The results hold after correcting for the small number of clusters using a wild-cluster bootstrap methodology and after testing for multiple outcomes. We provide evidence that the results are not driven by a violation of the difference-in-differences common time trend assumption. We also test for the presence of heterogeneous effects by child's gender, mother's education level, household's income and the intensity of non-maternal care use when the child was infant.

This paper contributes to the literature on the effects of non-maternal care on child development. High-quality non-maternal care has been shown to positively influence the cognitive outcomes of at risk children (Karoly et al. 1998). However, less evidence exists for the population at large, which might be of high interest to policy makers. Magnuson, Ruhm and Waldfogel (2004) and Loeb et al. (2007) show that non-maternal care is coupled with positive effects on cognitive ability but negative effects on behavior. Baker, Gruber and Milligan (2008, 2015) show that the introduction of the subsidized daycare fees in Quebec, while increasing maternal employment and the use of daycares, led to a worsening of non-cognitive outcomes of 0 to 5 years old children in the short-run as well as in the long-run.

Kottelenberg and Lehrer (2013) argue that what is captured in Baker et al. is the LATE effect, i.e. the effect on families that change their behavior due to the introduction of the policy. They estimate that, the average causal effect of attending child-care is an increase in cognitive outcomes and no worsening in non-cognitive outcomes. Kottelenberg and Lehrer (2014) show differential impacts of subsidized child-care by age groups. The negative effect on non-cognitive skills is mainly coming from children aged 0-2. Older children actually benefit from the introduction of subsidized daycare fees. Although our results are not directly comparable to the earlier studies where maternal employment status is changing coupled with a shift from maternal care to center based (daycare) care, our results are mainly in line with Kottelenberg and Lehrer (2013) showing that the introduction of universal child-care does not negatively affect the non-cognitive outcomes of non-infant children.

Our results contribute to a small but growing number of papers in this literature that compares the different types of non-maternal care as well as the non-maternal care intensity. In our paper, maternal care is crowding out informal care and the intensity of care is changing from 28 hours per week to 22 hours. Gupta and Simonsen (2010) show that informal care does not lead to significant differences in child outcomes relative to the use of formal care (daycares and pre-schools). The only negative effect is on the behavioral outcomes of boys with low-educated mothers. Our results extend those of Baker and Milligan (2008a, 2010) to older children. Baker and Milligan (2008a, 2010) show that an increase in maternal care, while crowding out the use of informal care, does not significantly improve the developmental outcomes of infants aged 6 to 24 months.

This paper also contributes to the literature on intensity of non-maternal care. Many studies have looked at the effect of part-time versus full-time maternal employment. For example, Gregg et al. (2005) show that part time work is not harmful for children aged less than 18 months. Other studies, have measured the intensity by the number of hours children spend in non-maternal care. Loeb et al (2007) argue that the intensity of non-maternal care matters: additional hours positively affect cognitive skills but negatively impact behavioural outcomes. Similarly, the NICHD Early Child Care Research Network (2003) show that the more hours children aged 2-3 years old spend in center-based care, the better are their language skills. Gupta and Simonson (2010) using the Danish Longitudinal Survey of Children show that an increase in the time children spend in nonmaternal care (center based or family day care) above 30 hours per week is associated with worse behavioural outcomes measured by the strengths and difficulties questionnaire index. This 30 hours threshold is also documented in the NICHD Early Child Care Research Network (2003) who argue that children spending more than 30 hours per week in center-based care exhibit worse behavioral outcomes. In our paper, the number of weekly hours spent in non-maternal care decreases from 28 hours to 22 hours. In other words, the reform is moving the time treated children spend in non-maternal care away from the high intensity threshold.

Finally, by looking at the unintended effects of the maternal/parental leave policies whose main target is newborns, this paper studies the spillover effects of these policies on older siblings. This topic, to our knowledge, has not been studied before. Therefore, parental leaves might have additional benefits and/or costs that are not factored into their designs.

The paper continues as follows. Section 2.2 summarizes the 2001 parental leave reform and Section 2.3 discusses the identification strategy. We describe the data and methodology in Sections 2.4 and 2.5. We present the results in Section 2.6. Section 2.7 provides robustness checks and Section 2.8 extends the results to the medium-run. In Sections 2.9 and 2.10, we present

heterogeneous effects and the treatment on the treatment estimates. Section 2.11 summarizes the findings and provides policy recommendations.

#### 2.2 The Reform

Parental leave in Canada is mandated by both provincial and federal governments. Uncompensated job protection is provided by provincial governments and gives mothers the right to return to their job after a specific time period following child's birth. The federal government provides benefits to parents faced by a decrease in income due to the birth of a child. The benefits are part of the Employment Insurance (EI) program and are calculated as 55 percent of the average insurable earnings during the six months preceding the claim. The insurable earnings' cap was \$39,000 per year at the time of the reform.

The reform used in this paper is a federal one that increased the period over which benefits are paid for mothers delivering after December 31, 2000. The reform was introduced to the Parliament on April 2000 and became a law on June 2000. Prior to December 31, 2000, benefits were paid for up to 25 weeks, with 15 weeks exclusive to the mother and 10 weeks that can be split among the parents. Eligibility was determined by working 700 hours in the 12 months preceding the claim. After the reform, mothers were eligible for up to 50 weeks of benefits as the period split among the parents increased from 10 to 35 weeks<sup>3</sup>. In addition, the requirement for eligibility was decreased to 600 hours.

To match the federal reform, most provinces<sup>4</sup> increased their uncompensated job protection. As of June 2001, all provinces offered job protection of 50 weeks or more. Figure 1 shows the changes to the job protection period in the different Canadian provinces.

## 2.3 Identification Strategy

In this paper, we exploit the fact that the parental leave reform provides an exogenous change to the time mothers spend outside the labor market. Baker and Milligan (2008) and Hanratti and Trzcinski (2009) show that the increase in parental leave had large impacts on maternal care in the

<sup>&</sup>lt;sup>3</sup> Marshall (2008) shows that only 10% of fathers claimed parental benefits in 2001.

<sup>&</sup>lt;sup>4</sup> Except Quebec, since it already provided 70 weeks of job protection.

child's first year. They estimate the effect to be 2 months for all mothers and 3 months for eligible mothers which is roughly equal to a 50 percent increase in the time spent with the newborn compared to pre-reform. In addition, Baker and Milligan (2008) show, using simulations, that mothers' income after tax, transfers, and child-care costs is not significantly different before and after the reform. Thus, the main channel through which the parental leave reform can affect children's outcomes is through the increase in maternal care<sup>5</sup>.

To study the effect of maternal care on 2-3 years old children, one can compare children with a sibling born before the reform to those with a sibling born after the reform. Mothers who delivered after December 31, 2000 were eligible for up to 50 weeks of parental leave, whereas those who delivered before are only eligible for 25 weeks. However, a before/after identification might lead to biased estimates, since children's developmental outcomes might be changing over time for reasons other than the parental leave reform. For this reason, we perform a difference-in-differences analysis that compares children who had siblings before and after the reform with those who did not have younger siblings in this period. The first difference controls for siblings' effects and the second difference controls for cohort effects, time trends and macroeconomic conditions that might be affecting child outcomes.

This identification strategy assumes that the reform did not affect fertility behavior. The reform could encourage families who were not planning on having more children to have additional children, hence, mothers who had a child after the reform might be systematically different than the ones who had a child before the reform which in turn might translate into different child developmental outcomes. However, Phipps (2001) show that mothers do not modify their labor market behavior or their fertility decisions based on changes in parental leaves. Moreover, Baker and Milligan (2015) show that the 2000 parental leave reform did not have a significant effect on the fertility rate of mothers with children aged 13 to 71 months.

Another threat to identification is that the reform might have affected mothers' work behavior. For example, it could be the case that mothers used to quit their jobs before the reform in order to spend more time with their newborns, or that women decided to start working prior to having a child to benefit from the extended parental leave. Lefebvre and Merrigan (2008) show that the maternal employment trends for mothers with children aged less than 5 are in general stable around

<sup>&</sup>lt;sup>5</sup> It can also increase breastfeeding duration. However, this is not a concern in this paper, since the sample consists of 2-3 years old children.

the 2000 parental reform<sup>67</sup>. Similarly, Baker and Milligan (2008b) and Schönberg and Ludsteck (2014) show that post-birth employment is slightly affected by expansions in maternity leave. In the next section, we provide descriptive statistics showing that, in our sample, mothers who had a child before 2001 are not significantly different than the ones who had a child after 2001, thus providing additional support for our identification strategy.

#### **2.4 Data**

#### 2.4.1. Sample

To study the effects of maternal care on child development, we use the Canadian National Longitudinal Survey of Children and Youth (NLSCY). The NLSCY is confidential and access is provided through the Research Data Centres (RDC). The NLSCY started in 1994/1995 (Cycle 1) by collecting data on a nationally representative sample of children aged 0-11, and followed these children biannually until 2007/2008 (Cycle 8). In every cycle, around 2000 children aged 0-1 year old were added and followed until they became 4-5 years old. Our sample of interest consists of children aged 2-3 years old in cycles 3 (1998/1999) and 5 (2002/2003). Given that the NLSCY does not provide the year of birth of the siblings, we compare the number of younger siblings from one cycle to the next to determine whether a younger sibling was born before or after the reform. An increase in the number of younger siblings between cycles 2 (1996/1997) and 3 (1998/1999) identifies siblings born before 2000 and between cycles 4 (2000/2001) and 5 (2002/2003) identifies siblings born after 2000. We do not use Cycle 4 since it is not possible to determine if the younger siblings born between Cycles 3 and 4 are born before or after the reform. We restrict our sample to Cycles 3 and 5 to ensure that all the children studied in this paper are born before December 2000 so that they were all subject to the shorter parental leave<sup>8</sup>.

In addition, we restrict our sample to two parents' families living outside Quebec. The first restriction is due to a change in the National Child Benefit<sup>9</sup> affecting the labor force participation

<sup>&</sup>lt;sup>6</sup> Except in Quebec due to the introduction of the Quebec's Family Policy (with highly subsidized child care) around the same time period.

<sup>&</sup>lt;sup>7</sup> For the US, Waldfogel (1999) demonstrates that the Family and Medical Leave Act (FMLA) did not have a significant effect on the employment of mothers with a child less than 1-year-old.

<sup>&</sup>lt;sup>8</sup> We use Cycle 2 to test for the common time trend assumption. We refrain from using Cycle 1 due to major differences in the available variables.

<sup>&</sup>lt;sup>9</sup> The National Child Benefit is a benefit paid to low income families that increased from \$605 in 1998 to \$1293 in 2002. Some provinces impose some labour market attachment to receive the benefit.

of single mothers (Milligan and Stabile, 2007). The second restriction stems from the fact that Quebec introduced a universal child-care policy in 1997 that made subsidized child-care services available to 4-year-old children. The policy was introduced in phases and in 2000 all children aged between 0 and 4 had the right to access subsidized child-care. Baker, Gruber and Milligan (2008), Lefebvre, Merrigan, and Verstraete (2009) and Haeck, Lefebvre and Merrigan (2015) show that this universal child-care policy has an effect on maternal labor supply and on child development. After these restrictions, the sample consists of 4,985 observations. In all our analysis, we use the weights provided by the NLSCY. These weights are adjusted for non-response and post-stratified on province, age and gender to make the survey nationally representative.

The NLSCY does not specify whether the mother is eligible to take a maternity leave or not. For this reason, our estimates are the intention to treat estimates rather than the treatment on the treated. This practice is common in the literature. We try to reconcile the treatment on the treated estimates in the last section of this paper by restricting our sample to mothers most likely eligible for parental leaves.

#### 2.4.2 Variables Description

The NLSCY interviews the person most knowledgeable (PMK) of the child. It provides socioeconomic information about the PMK and his/her spouse. It also offers a large range of parent-reported pre-cognitive, non-cognitive and health variables and scales for children aged 2-3 years old. Direct assessments of cognitive abilities start at age 4.

The first set of variables we are interested in, consists of measures on the use of non-maternal care as well as the intensity of usage measured in hours per week. Non-maternal care is divided into informal care (care provided by a relative or non-relative inside or outside the child's house) and center based care (professional daycare centers).

The second set of variables describes child's pre-cognitive and cognitive outcomes. The Motor and Development Section of the questionnaire is used as a measure of pre-cognitive outcomes for children aged up to 3 years old. It measures dimensions of the children's motor, social and cognitive development, and consists of a set of 15 questions that are answered by the PMK about the child's physical ability (such as washing and drying hands alone, doing a somersault without help) as well as his/her cognitive and social abilities (such as naming at least 4 colors, counting up to 10). We use the standardized version of the score that is set to a mean of 100 and a standard deviation of 15. The Motor and Development Scale has been used in other surveys such as the National Longitudinal Survey of Youth in the USA. The NLSCY provides several measures of cognitive outcomes for children aged 4-5 years old such as the Peabody Picture Vocabulary Test-revised (PPVT), the "Who Am I?" test and the Number knowledge test. We use these tests to study medium-run effects of maternal care. The PPVT measures receptive vocabulary. It consists of 175 stimulus words. The child has to choose the image that best describes the word from a set of 4 black and white drawings. The PPVT is extensively used in many disciplines. The "Who Am I?" test consists of 10 questions divided into two parts, the copying and writing tasks. Each question is awarded a score of 1 to 4 with a higher score implying better performance. In total, the "Who Am I" score ranges from 1 to 40. The main objective of the test is to measure the child's ability to understand symbols including geometry, letters and numbers. The Number Knowledge test assesses the child's understanding of numbers through 30 questions used to place the child on a four-point scale, where a higher point on the scale implies that the child has reached a higher level of understanding (Statistics Canada 2003).

The NLSCY provides four parent reported behavioral/non-cognitive indices for children aged above 2 years old. The scales are computed from answers to questions asked to the PMK about the child's behavior in different situations. The answer can take the value 0 (not true), 1 (sometimes true) or 2 (often true). The hyperactivity score consists of 5 questions (child is restless, is easily distracted, cannot concentrate, ...) and ranges from 0 to  $10^{10}$ . The emotional disorder score ranges from 0 to 12 on the basis of answers to 6 questions such as whether the child seems unhappy, is worried, is nervous. The physical aggression score ranges from 0 to 16 and asks questions about behavior associated with opposition such as whether the child gets into many fights, has hot temper. The last behavioral scale is the separation anxiety score that ranges from 0 to 10 on the basis of the answers to 5 questions such as whether the child cries a lot, constantly seeks help. In all the non-cognitive scales, a higher value indicates a more problematic behavior. All of these measures are commonly used in the literature and are shown to reliably predict underlying outcomes (Charach, Lin and To 2010).

The health variables in the NLSCY consist of measures of the current height and weight of the child that we use to create different categories of BMI levels: underweight, overweight and obese;

<sup>&</sup>lt;sup>10</sup> Compared to Cycle 3, two questions were removed from the hyperactivity score and one question was added in Cycle 5. We calculate a new hyperactivity score by keeping only the questions asked in both cycles.

based on the CDC child BMI-for-age growth charts. In addition, the NLSCY offers measures provided by the PMK on whether the child is in excellent health and whether the child has been injured, had an ear infection, and had a throat/nose infection in the last 12 months.

The last set of variables used in this paper are the parent and family outcomes. Mother is in excellent health is a dichotomous variable equal to 1 if the mother self-reports being in excellent health. A similar variable is used for the father being in excellent health. The mother depression score is a set of 12 questions asked to the PMK (how often, during the last week, the PMK felt lonely, felt depressed, ...). The answers can take the value 0 (rarely or none of the times, less than 1 day), 1 (some or little of the time, 1-2 days), 2 (Occasionally or a moderate amount of time, 2-3 days) and 3 (Most or all of the time, 5-7 days). The score ranges from 0 to 32 with a higher number implying the presence of depression symptoms. The family dysfunction score ranges from 0 to 36 and is based on 12 questions about the quality of the relationship between the parents. The answers vary from strongly agree (0) to strongly disagree (4). A higher value implies more family dysfunction.

The Appendix reports the exact questions asked to calculate the motor and development score, the non-cognitive scores, the mother depression score and the family dysfunction score.

Most of the NLSCY measures used in this study are parent-reported and thus, might suffer from systematic biases. However, De Los Rayes and Kazdin (2005) show that parental measures are usually very informative of the underlying measure and are highly correlated with the professional assessment.

#### **2.4.3 Descriptive Statistics**

A summary of descriptive statistics for all the outcomes variables is provided in Table 1. Fiftyfour percent of the children in the sample attend some type of non-maternal care for an average of 28 hours per week. The majority of the children attend informal care<sup>11</sup>. On average, the noncognitive scores are concentrated in the lower portion of their range. The proportions of overweight and obese children are 12 and 35 percent respectively. The proportion of obesity is consistent with Kottelenberg (2015) who uses a sample of 2-4 years old children from the NLSCY and shows the

<sup>&</sup>lt;sup>11</sup> Note that informal care and centre-based care are not mutually exclusive. A child could be using both services at the same time. The percentages of children attending informal care and centre-based care are reported conditional on the child using some type of non-maternal care.

proportion of obesity in his sample to be 30 percent. However, the proportion of overweight is much smaller in our sample compared to the 40 percent present in Kottelenberg (2015). Sixty-four percent of the parents report that their kids are in excellent health. Sixty-two percent of the children had an ear infection in the last year and a similar proportion had a throat/nose infection. Thirty-eight percent of mothers and fathers report being in excellent health.

Table 2 compares mothers, fathers and family characteristics between children in the treatment and control groups. Columns 1 and 2 report the mean and standard deviation (in parentheses) for children with sibling born before and after the reform respectively. Similarly, parent and family characteristics are reported for children of the same birth cohorts who did not have a sibling in the period surrounding the reform in column 3 (for children observed before the reform) and 4 (for children observed after the reform). Column 5 reports the difference-in-differences estimates for the control variables along with a t-test [in brackets]. In general, parents and family characteristics are balanced between the different groups with the exception of the number of younger or same age children and the family residing in rural areas. In subsequent analysis, we report results with and without controlling for these variables. The last column of Table 2 reports the difference in means between children with a sibling born before the reform and those with a sibling born after the change (columns 1 and 2). This comparison is important to show that there was no selfselection due to the reform, at least on the observable characteristics. As can be seen, none of the differences is significant between the two groups giving additional support for the identification strategy.

#### 2.5 Methodology

Our empirical analysis relies on the change in the parental leave policy that granted mothers delivering a child after December 31, 2000 up to 50 weeks of leave compared to 25 weeks for those delivering before the reform. In order to determine the effects of the increase in mother's time at home on cognitive, non-cognitive and health outcomes of older children, we perform the following difference-in-difference:

$$y_{it} = \beta_0 + \beta_1 sibling_i + \beta_2 post_t + \beta_3 sibling x post_{it} + \theta X_{it} + u_{it}$$

where *i* corresponds to individual *i*, and *t* to year *t*.  $y_{it}$  is the variable of interest (child's cognitive, non-cognitive and health variables). *sibling<sub>i</sub>* is a dichotomous variable equal to 1 if the child has a younger sibling, *post<sub>t</sub>* is a dichotomous variable equal to 1 if the child is observed after December 31, 2000. *siblingxpost<sub>it</sub>* is the variable of interest. It isolates children with a sibling born after the reform. Thus, the coefficient of interest is  $\beta_3$ . It measures the intention to treat. The variable  $X_{it}$  is a set of control variables including child's gender, age of child in months at the time of survey, month of birth fixed effects, a dummy indicating whether the child is bilingual, dummies for the number of younger and older siblings, province fixed effects, city size, household income as well as age, education, and immigration status of each of the parents<sup>12</sup>.  $u_{it}$  is the error term.

In all the analysis, we correct for the small number of clusters and test for multiple outcomes.

#### 2.5.1 Wild-Cluster Bootstrap

To deal with the issue of serial correlations that Bertrand, Duflo and Mullainathan (2004) warn about when using the difference-in-differences methodology, we cluster the standard errors by province, the level at which the job protection reform took place (Moulton 1990). Failing to cluster leads to great over-rejection if the errors are correlated within province. However, clustering alone is not sufficient in this study, since our data consists of a small number of clusters (9 provinces). Cameron, Gelbach and Miller (2008) show that, even after clustering, the asymptotically normal critical values provide a poor approximation of the Wald test statistic which might lead to overrejection. Using Monte Carlo simulations with 10 clusters and different cluster sizes and error structures, they show that the asymptotically normal critical values lead to great over-rejection in the range of 10.6 percent to 77 percent. Even after clustering, the rejection rate ranges from 8.2 percent to 18.3 percent. They recommend using a wild-cluster bootstrap methodology to deal with the issue of small number of clusters. The wild-cluster bootstrap provides new critical values for the Wald test statistic by re-sampling the residuals and randomly assigning weights that keep the correlation between the errors in the same cluster intact. They show that when using wild-cluster bootstrap, the rejection rate falls to a range of 4.5 percent to 6.4 percent which is not significantly different that 5 percent.

<sup>&</sup>lt;sup>12</sup> The control variables are discretized. For example, the mother's age is divided into 5 categories. This is done to avoid mis-specifying the functional form of those variables in the regression.

The exact procedure of the wild-cluster bootstrap is as follows. First, from the original sample, we calculate the Wald test statistic while imposing the null hypothesis. Second, we form a sample of nine clusters by re-sampling the residuals and randomly assigning weights  $a_g$  satisfying the following criteria:  $E[a_g]=0$ ,  $E[a_g^2]=1$  and  $E[a_g^3]=0$ . Third, we calculate a new Wald test statistic using the new sample. Fourth, we repeat the second and third steps 999 times. These 999 samples approximate the distribution of the Wald test statistic. Finally, we rank the newly computed Wald test statistics and calculate a wild-cluster bootstrap p-value by ranking the original Wald test among the 999 constructed ones. Several weights  $a_g$  have been proposed for the wild-cluster bootstrap. We use the 6-point distribution proposed by Webb (2014).

#### 2.5.2. Multiple Outcomes Test

In our analysis, we test the effects of the parental leave reform on a host of variables. Failure to account for multiple outcomes might lead to false discoveries, Type I error. For example, the probability of finding at least one significant result when one is testing five outcomes at the 5 percent significance level is 22.6 percent. This means that there is a 22.6 percent probability of finding a significant result due to chance. To this end, we control for the false discovery rate (FDR), i.e. the expected proportion of incorrectly rejected null hypothesis. Multiple outcome tests are widely used in psychology and biostatistics (Benjamini and Yekutieli 2001; Hochberg 1998), but much less used in the economics discipline (Kling, Liebman and Katz 2005; Anderson 2008; Milligan and Stabile 2011; Kottelenberg and Lehrer 2013, 2014).

We use the two-stage linear step-up model proposed in Benjamini, Krieger and Yekutieli (2006) to control for the FDR. The model works as follows: Suppose we are testing *m* hypotheses at significance level *q*. First, sort the hypotheses,  $H_1,...,H_m$  such that  $p_1 < p_2 < ... < p_m$ . Define q'=q/(1+q). Second, let *c* be the largest *r* for which  $p_r < qr/m$ . If such a *c* exists, reject hypotheses  $H_1,...,H_c$ , and continue to the third step otherwise stop and do not reject any hypotheses. Third, define  $m_0=m-c$ . Redo the second step at a level  $q^*=q'm/m_0$ . This procedure controls for FDR at level *q*.

The two-stage linear step model reports whether a hypothesis is rejected or not at a given level of q, but does not report the smallest q for which the hypothesis is rejected. This number is the equivalent of the p-value after controlling for FDR. We use the implementation proposed by Anderson (2008) to estimate this q-value by running the two-stage linear step-up procedure for

each hypothesis for every possible q and recording the q-value for which the hypothesis fails to be rejected. In all subsequent analysis, we report both the wild-cluster bootstrap p-value and the corresponding q-value.

## 2.6 Results

For every outcome, we run three different regressions. The first specification does not include any control variables. The second regression replicates the specification usually used in the literature studying the effects of maternal care on child development, by adding all the control variables except household income. All controls are used in the last specification. We present the results in four steps. First, we study the effect of increased maternal time at home on the use of child-care, both on the extensive and the intensive margins, as well as on the type of non-maternal care used. Afterwards, we evaluate the impact of the change on child's non-cognitive and pre-cognitive outcomes, health outcomes, and parents and family outcomes.

#### 2.6.1 Changes in Non-Maternal Care

To study the effect of the increase in mother's time at home on non-maternal care, we perform the difference-in-differences analysis on a general measure of child-care usage (a dichotomous variable equal to 1 if the child receives any type of non-maternal care), the type of child-care (informal care and centre based care) as well as the hours of care. Table 3 reports the results. Each row corresponds to a regression for a specific outcome variable. Columns 1, 3 and 5 present the estimates of our coefficient of interest,  $\beta_3$ , for the three specifications explained above. The wild-cluster bootstrap p-values are reported in parentheses under each coefficient. Columns 2, 4 and 6 report the corresponding q-values in brackets for each outcome variable.

Even though mothers spend more time away from the labor market, there is no significant change in non-maternal care usage and in the type of care used. However, children spend fewer hours of non-maternal care. Specifically, compared to children without a younger sibling, those with a sibling born after the reform spend 4.5 hours/week less in non-maternal care than those with a sibling born before the reform. This is equivalent to a 16 percent decrease in non-maternal care per week and the result is significant at the 5 percent level (10 percent after controlling for the FDR). This estimate provides a lower bound of the decrease in the time spent by children in non-

maternal care, since some mothers might have already finished their parental leave and returned to work at the time of the interview. In the last two rows of Table 3, we show that the decrease in non-maternal care hours is mainly due to a 27 percent decrease in the hours spent by children in informal care. There is no significant change in the time spent at daycare centers. Maternal care is crowding out informal care for older children.

#### 2.6.2 Changes in Child Non-Cognitive and Pre-Cognitive Outcomes

In Table 4, we analyze the effect of increased maternal care on children's non-cognitive outcomes: hyperactivity score, emotional disorder score, physical aggression score, and separation anxiety score. We standardize these scales so that the mean is 0 and the standard deviation is 1. The coefficients can then be interpreted as a percentage change of a standard deviation. Maternal care is positively associated with emotional disorder and physical aggression, and negatively associated with hyperactivity and separation anxiety. However, we fail to reject that the effect is significantly different from zero even before controlling for multiple outcomes.

The last row of Table 4 reports the intention to treat coefficient for the pre-cognitive outcome, the motor and development standardized scale. The negative coefficient supports the evidence from the literature about the positive effects of non-maternal care on cognitive skills but this estimate is insignificant in all three specifications.

#### 2.6.3 Changes in Child Health Outcomes

Table 5 reports the difference-in-differences results for the health outcomes: dichotomous variables of whether the child is underweight, overweight or obese (with normal weight being the omitted category), a dichotomous variable excellent health equal to 1 if the parent reported measure of the child's health is excellent, dichotomous variables indicating whether the child was injured during the last 12 months, had an ear infection in the last 12 months, and had a throat/nose infection in the last 12 months. Maternal care does not affect any of the health variables significantly, as can be seen in Table 5, which extends the results found in Baker at al. (2008) regarding the effects of maternal care and breast-feeding on infant health outcomes.

#### 2.6.4 Changes in Parent and Family Outcomes

The effects of maternal care on parent and family outcomes are studied in Table 6. For this category, we define two dichotomous variables, mother/father in excellent health to be equal to 1 if the mother/father reports being in excellent health, 0 otherwise. We also use the scales provided by the NLSCY on mother's depression score and family dysfunction score. We report the results for standardized measures of the last two variables. Less mothers report being in excellent health. This measure is only significant in the first specification that does not include any control. After correcting for the FDR, the significance disappears. In general, it seems that an increase in maternal care affects negatively mother's reported health and increases their depression score, although, none of the coefficients is significant. Baker et al. (2010) report that increases in parental leave do not positively influence maternal health and depression status.

#### **2.7 Robustness Checks**

One main assumption of the difference-in-differences methodology is the common time trend. It assumes that if the treated group was not treated, it would have followed the same trend as the control group. In our analysis, this implies that if the children with a sibling born after the reform did not spend longer time with their mothers than the children with a sibling born before the reform, they would have evolved similarly to the children without a younger sibling in the period surrounding the reform. In this section, we test provide baseline results to test whether the children time in our sample were different when they were infants. We also provide a test for the common time trend assumption.

#### 2.7.1 Baseline Results

In this subsection, we test whether there were any significant differences between the treated and control children in our sample during their first year of life, as early differences might persist into older ages. We use the data available in the NLSCY to look at the children in our sample one cycle earlier, i.e. in cycles 2 and 4, when the children were 0-1 year old. All these children were born before the parental leave reform, thus their mothers should have spent a comparable time with them after their birth. We apply the same difference-in-differences methodology described in Section 2.4 to a set of birth variables, mother's work after birth and non-maternal care, health, parent and family, and child temperament variables<sup>13</sup>. For the common time trend assumption to hold, there should not be any significant differences between the children in their early years of life. We only report the results of the second specification that is usually used in the literature, but similar results arise when the other specifications are considered.

Table 7 reports the results. The first panel includes birth and after birth variables. Pre-mature child, breast-fed and normal birth weight are dichotomous variables equal to 1 if the child was born before the due date, if the child was breast-fed and if the child was born with a normal birth weight respectively. These variables can directly affect child development at older ages as well as economic success in the long-run (Quinn et al. 2001; Behrman and Rosenzweig 2004; Kramer et al. 2007, 2008). None of these variables is significantly different between the treated and control groups in their first year of life. The second set of variables is related to mother's work after birth and the use of non-maternal care. The variable "worked after birth" is a dichotomous variable equal to 1 if the mother has worked at any time between the child's birth and the date of the interview, 0 otherwise. We also report the child's age in months when the mother returned to work and how many hours per week the mother was working. The second panel of Table 7 shows that the children in our sample had, on average, similar maternal care experience when they were infant. These results give support to our identification strategy by showing that the difference in the total hours spent in non-maternal care between the treated and control groups in our sample was not significantly different, when those children were still infants. The third to fifth panels include precognitive, health, and parent and family variables similar to the ones used in the original regressions. The last set of variables are child temperament measures used as proxies for the behavioral and non-cognitive outcomes, since the non-cognitive scales used earlier in this paper are only reported for children aged more than 2 years. We present the same variables used in Baker et al. (2010). Each variable is a seven scale measure with a lower number implying better results. For example, for the variable difficulty to calm, a lower number means that the child is easy to calm. All the results are reported in a standardized form.

From table 7, we can safely conclude that there are no baseline difference between the children with a sibling born around the reform and those who do not have a sibling are parallel. As a matter

<sup>&</sup>lt;sup>13</sup> We do not include the same set of variables used in the original analysis, since the NLSCY questions are age specific. The questions asked about the 0-1 year old children are different than the ones asked about the 2-3 years old children.

of fact, the majority of the results are insignificant except for the child being in excellent health and how changeable the child's mood is, which show that treated children were more likely to be reported as being in excellent health and less likely to be reported as having mood swings when they were 0-1 year old. This means that, if anything, treated children were better off when they were still infant. The lack of significant results at ages 2-3 implies that increase in maternal care does not have positive effects on child development at that age. In any case, this significance might be due to pure chance since it vanishes when controlling for the FDR.

#### 2.7.2 Common Time Trend Test

Figure 2 presents the trends of children's cognitive, non-cognitive and health outcomes as well as the probability of using non-maternal care and the weekly number of hours in all non-maternal care from 1996-1997 until 2002-2003 for the children with younger siblings compared to children without a younger sibling. For the common time trend assumption to hold the trend before the reform, i.e. in 1996-1997 and 1998-1999, between the treatment and control groups should be parallel<sup>14</sup>. This parallel trend can be clearly seen in Figure 2.

To formally test for any violations of the common time trend, we assume that the reform was implemented in 1996-1997 instead of 2000. We compare children aged 2-3 years in 1996-1997 with a sibling born before the hypothetical reform to children aged 2-3 years in 1998-1999 with a sibling born after the hypothetical reform. We also compare them to children of the same birth cohorts who did not have a younger sibling in that period of time. Since the reform is hypothetical, we should not find any significant differences in the outcomes of the children, as finding otherwise implies a violation of the common time trend between the children with and without younger siblings.

The results are reported in Table 8 for the cognitive, pre-cognitive and health outcomes of the children as well as the parents and family outcomes. These results show that the pre-reform trends are parallel as none of the results is significant expect for the emotional disorder score and mother in excellent health at the 10 percent level. This implies that relative to children who do not have a

<sup>&</sup>lt;sup>14</sup> Note that children with a younger sibling in 2000-2001 might be affected by the reform if the younger sibling is born after December 31, 2000. However, we do not know the sibling's exact day of the birth, thus the trend might start changing in 2000-2001. For the common time trend assumption to hold, we just need that parallel trends in 1996-1997 and 1998-1999.

sibling, children with a sibling born after 1996-1997 are less likely to exhibit emotional disorder than children aged 2-3 years old with a sibling born before 1996-1997. In addition, their mothers are more likely to self-report in excellent health. Those significant results do not pass the multiple outcomes test as can be seen in the last column of Table 8. In any case, these results signify that children with a sibling were faring better over time in the absence of the reform. The lack of significant results in Section 2.5 puts further doubts on the positive effects of the extra maternal care induced by the parental leave expansion.

#### 2.8 Medium-Run Results

The insignificant effects found in Section 2.5 measure the short-run impact of the increase in maternal care on child development. It might be the case that the increase in maternal care takes a longer time to manifest. For this reason, we follow the children in our sample one cycle in the NLSCY, i.e., when they become 4-5 years old. The advantage is not only studying medium-run effects, but also having direct assessment of cognitive skills through the use of the PPVT, the "Who Am I?" and the number knowledge scores<sup>15</sup>. Table 9 reports the results of the difference-indifferences analysis for four different sets of variables: cognitive outcomes, non-cognitive outcomes, health outcomes and parent and family outcomes. None of the outcomes is significant except for underweight and excellent health that are significant at the 10 percent level. Relative to children without younger sibling in the period surrounding the reform, children with a sibling born after the reform are 2 percentage points more likely to be reported as underweight and 10 percentage points less likely to be reported in excellent health compared to children with a sibling born before the reform. These results cease to be significant when considering the q-values. This shows again that additional maternal care induced by the reform does not have a positive effect on children's cognitive, non-cognitive and health outcomes, either in the short-run or the mediumrun.

The last panel of Table 9 reports the results for non-maternal care use when the sample of interest is aged 4-5 years old. The insignificant coefficient of the weekly number of hours spent in non-maternal care supports our identification strategy. It shows that the mothers of treated children

<sup>&</sup>lt;sup>15</sup> The NLSCY provides direct assessment of the children's cognitive skills starting at age 4.
decided to spend more time with their children, mainly because of the parental leave reform, rather than for other reasons correlated with unobserved characteristics of the mothers and/or children.

# 2.9 Heterogeneous Effects

In the previous sections, we assumed that the effects of maternal care on child development is homogeneous. However, heterogeneous effects might arise in various settings. In this section, we test for the presence of heterogeneous effects by child gender, mother's education, household income and the intensity of early non-maternal care.

## 2.9.1 Child Gender

The effects of non-maternal care might differ greatly depending on the child's gender. Several studies have shown that girls benefit more than boys from the use of non-maternal care (Belfield et al. 2006; Anderson 2008). Baker et al. (2015) show that the negative effects of universal child-care on non-cognitive outcomes are mainly driven by boys. Table 10 reports the results of the difference-in-differences methodology of the effects of maternal care on pre-cognitive, non-cognitive, health, and parent and family outcomes by gender. The first two columns report the results for boys and the last two columns report the results for girls. The coefficient of the motor and development score is negative for girls and positive for boys, in line with the literature showing a positive effect of child-care on girls' cognitive outcomes, however, those coefficients are not significantly different from zero.

The main difference arises in a worse separation anxiety score for boys, which is in line with Kottelenberg and Lehrer (2012) showing that boys reveal more separation anxiety than girls. In addition, boys have less prevalence of throat/nose infections. Mothers of boys report less being in excellent health and report having better family functioning. However, none of these results is significant after correcting for the FDR.

### 2.9.2 Mother's Education

The effects of maternal care might have differential impact on child development depending on the quality of the time spent with the child. Maternal education can be used as a proxy for the quality of maternal care. Children of low educated mothers have been shown to benefit mostly from child-care (Magnuson et al. 2004; Havnes and Mogstad 2011). Table 11 replicates the analysis for two sub-samples: mother with a university degree and mothers without a university degree. We expect that children of mothers with a university degree will benefit more from an increase in maternal care than those of mothers without a university degree. The results in Table 11 do not support this hypothesis. In general, there is no difference in the developmental outcomes of the children in the two groups, except for an increase in the behavior related to hyperactivity and the percentage of children reported as overweight in the sample of holders of university degrees. Keeping in mind that all the measures used in this table are parent reported, when studying heterogeneity due to maternal education, a systematic difference in reporting between high and low educated mothers might affect the results. More educated mothers might be more likely to report worrying child behavior. In any case, none of the differences between the two groups remain when correcting for the FDR.

#### 2.9.3 Household Income

Children from low income households benefit more from the presence of high quality child-care than children of high income households (Karoly et al. 2008; Felfe, Nollenberger and Rodriguez-Planas 2015). In this sub-section, we study the differential impact of maternal care by households' income. These results should be interpreted with caution, since household income might be endogenous to the reform. However, Baker and Milligan (2010) show that mother's income after tax, transfers and child-care costs is not significantly affected by the reform. Given that only a small percentage of fathers take parental leave (Marshall 2008), there is no reason to expect that the reform affects fathers' income. In Table 12, we divide our samples into families with household income below versus above the median household income of \$60,000. The motor development score is negative for families with income below the median and positive for families above the median, in line with the research showing a larger effect of child-care on the cognitive outcomes of children in low income families; however, those results are not statistically significant. More mothers with an income above the median report being in excellent health due to the reform. This might be due to the fact that mothers in families with income above the median might be more likely to be eligible and to take up the parental leave; however, the significance of these results vanishes when considering the q-values.

### 2.9.4 Intensity of Early Non-Maternal Care

The intensity of non-maternal care usage might have differential impact on child development. Loeb et al. (2007) show that high intensity usage leads to greater benefits in cognitive outcomes but larger problems in non-cognitive outcomes. In Table 13, we study the differential effects of an increase in maternal care by the intensity of use of non-maternal care when the child was an infant. Heavy users are defined as children using non-maternal care for more than 20 hours per week<sup>16</sup> when they were aged 0-1 years old. The decrease in the motor development standardized score is larger, in absolute values, for the heavy users than the non-heavy users in line with Loeb et al. (2007), however, the difference is not statistically significant. In general, non-cognitive outcomes are positively affected by maternal care for heavy-users except for the separation anxiety score, but the effects are very small and insignificant. The only result that is significant after correcting for the small number of clusters and the multiple outcomes is the child's health. For the heavy-users, children with a sibling born after the reform are 11 percentage points more likely to be reported by their mothers to be in excellent health compared to children with a sibling born before the reform, relative to children of the same birth cohort who did not have a younger sibling in the period surrounding the reform.

## 2.10 Moving Beyond the Intention to Treat Estimates

Given that the NLSCY does not report whether, in our sample, a mother is eligible for parental leave or not, our earlier estimates were the intention to treat rather than the treatment on the treated. The practice of reporting the intention to treat is common in the literature evaluating maternal care and parental leaves. The intention to treat is informative about the effects of an increase in maternal care for the general population. The treatment on the treated estimates the effect of maternal care on children whose mothers actually took a longer parental leave.

We restrict our sample in a way that captures children of the mothers who are most likely eligible for the parental leave. We consider the mothers who were working in the cycle before delivering their newborns (Cycles 2 and 4 in the NLSCY) as eligible for parental leave, and those who were

<sup>&</sup>lt;sup>16</sup> 20 hours per week correspond to a part-time use of non-maternal care.

not working as non-eligible. Note that eligibility to parental leave does not necessarily imply takeup. Data from the Survey of Employment Insurance Coverage shows that around 85 percent of eligible mothers claimed benefits around the reform time (Baker et al. 2010). Even though this analysis is informative, caution should be taken when interpreting the results.

Table 14 reports the difference-in-differences estimates for the children of the mothers who are most likely eligible for the parental leave and the children of mothers who are most likely noneligible. We should expect to find an increase in maternal care for the children of eligible mothers only. This is confirmed in the first panel of Table 14 showing that the weekly hours in non-maternal care only decreased for children whose mothers were working in the previous cycle. The precognitive and non-cognitive outcomes, health outcomes, and parent and family outcomes are reported in the second to fourth panel. Relative to children with eligible mothers who did not have a younger sibling in the period surrounding the reform, children of eligible mothers with sibling born after the reform are 8 percentage points more likely to be reported in excellent health and 4 percentage points less likely to have encountered a nose and throat infection in the last 12 months compared to children of eligible mothers with siblings born before the reform. Likewise, we find a 12 percentage points increase in the share of mothers reporting to be in excellent health. However, the increase in maternal care does not translate into better cognitive or non-cognitive outcomes. Even for the health outcomes that are significant, the significance disappears once controlling for the FDR.

In summary, both the intention to treat and the treatment on the treated estimates do not provide clear evidence that the increase in maternal care affects children aged 2-3 years old positively.

## 2.11 Conclusion

Using a parental leave reform implemented in Canada on December 31, 2000, we study the effects of an increase in maternal care on cognitive, non-cognitive and health outcomes of children aged 2-3 years old. The reform increased the time eligible mothers spend at home with their newborns by 3 months without affecting their after-tax income, net of transfers and child-care costs. We perform a difference-in-differences analysis using data from the NLSCY, to compare children with a sibling born after the reform to children with a sibling born before the reform relative to children of the same age group who did not have a younger sibling in the period surrounding the reform.

Results show that mothers giving birth after the reform spend more time with their older children, with maternal care crowding out informal care. Specifically, treated children spend 16 percent less time in non-maternal care moving away for the high-intensity non-maternal of 30 hours per week. However, the increase in maternal care does not translate into better child outcomes in the short-run or medium-run.

This paper shows that the parental leave expansion has spillover effects on older siblings that should be taken into account when devising such policies.

We have to keep in mind that most of the outcomes studied in this paper are parent reported. Future research should consider using datasets that provide professional assessment of child outcomes and measures about the quality of maternal care and non-maternal care. In addition, the results of this paper can be extended to include other age groups and children of single mothers.

# Tables

Table 1: ]	Descriptive S	tatistics for	the Outcom	ie Variables

	(1)	(2)	(3)
	Nbr of obs	Mean	Std Dev
Non-maternal care			
Use of care	4985	0.54	0.50
Informal care	4985	0.82	0.39
Center based care	4985	0.33	0.47
Weekly hours in all non-maternal care	2226	28.46	15.62
Weekly hours in informal care	2235	24.86	15.19
Weekly hours in center based care	557	25.87	14.42
Non-cognitive and pre-cognitive scores			
Hyperactivity score	4879	2.79	2.04
Emotional disorder score	4866	1.30	1.54
Physical aggression score	4723	4.90	2.61
Separation anxiety score	4885	2.76	2.03
Motor and Development score	4787	100.49	15.12
Chid health outcomes			
Underweight	3766	0.12	0.32
Overweight	3766	0.12	0.33
Obese	3766	0.35	0.48
Excellent health	4985	0.64	0.48
Injured in the last 12 months	4941	0.11	0.31
Ear infection in the last 12 months	4932	0.62	0.49
Throat/nose infection in the last 12 months	4938	0.63	0.48
Parent and family outcomes			
Mother in excellent health	4818	0.38	0.48
Father in excellent health	4759	0.38	0.49
Mother's depression score	4530	4.16	4.72
Family dysfunction score	4696	8.64	5.06

**Notes:** Descriptive statistics for the sample of interest. The first column reports the number of observations. The mean and standard deviation are reported in columns 2 and 3 respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Sib	ling	No S	ibling	DID	Difference
	Pre	Post	Pre	Post		(2) - (1)
Mother's characteristics						
Age	31.26	31.79	33.46	33.72	0.26	0.53
	(4.28)	(5.91)	(5.08)	(5.69)	[0.56]	[1.34]
Immigration status	0.14	0.19	0.22	0.24	0.02	0.05
	(0.34)	(0.42)	(0.40)	(0.43)	[0.59]	[0.42]
Did not complete	0.08	0.08	0.09	0.09	0.001	0.001
high school	(0.27)	(0.30)	(0.28)	(0.29)	[0.03]	[0.02]
University graduate	0.52	0.56	0.51	0.54	0.01	0.04
	(0.49)	(0.54)	(0.49)	(0.50)	[0.18]	[0.95]
Father's characteristics						
Age	33.46	33.43	35.68	35.55	0.21	-0.03
	(4.79)	(6.36)	(6.30)	(6.16)	[0.55]	[-0.07]
Immigration status	0.14	0.19	0.18	0.23	0.002	0.05
	(0.34)	(0.44)	(0.02)	(0.02)	[0.06]	[1.52]
Did not complete	0.13	0.12	0.14	0.09	0.02	-0.02
high school	(0.34)	(0.35)	(0.34)	(0.29)	[0.78]	[-0.85]
University graduate	0.55	0.55	0.51	0.53	-0.02	-0.001
	(0.49)	(0.51)	(0.49)	(0.51)	[-0.49]	[-0.02]
Family characteristics						
Married parents	0.92	0.92	0.91	0.90	0.02	0.001
	(0.27)	(0.29)	(0.28)	(0.30)	[0.69]	[0.32]
Child is a boy	0.51	0.52	0.53	0.51	0.03	0.01
	(0.49)	(0.54)	(0.49)	(0.50)	[0.58]	[0.29]
Number of older	0.55	0.50	1.09	0.94	0.10	-0.05
siblings	(0.95)	(1.14)	(0.98)	(0.97)	[1.14]	[0.54]
Number of younger	1.07	1.05	0.02	0.08	-0.08	-0.02
or same age siblings	(0.26)	(0.26)	(0.14)	(0.28)	[-3.04]***	[-0.71]
Resides in rural area	0.12	0.14	0.12	0.09	0.05	0.02
	(0.32)	(0.37)	(0.32)	(0.29)	[2.00]**	[0.88]
Resides in a large	0.43	0.43	0.47	0.45	0.02	0.004
city (>500K)	(0.49)	(0.54)	(0.49)	(0.50)	[0.42]	[0.12]
Household income	68037	67525	71869	71773	-416	-512
	(42566)	(47446)	(40325)	(48223)	[-0.11]	[-0.16]

# Table 2: Parents and Family Characteristics

**Notes:** Each row corresponds to an independent variable. The first two columns provide the mean and standard deviation (in parentheses) for the children with a sibling born before 2001 and the after 2000, respectively. Column 3 and 4 provide the mean and standard errors for the control groups observed before and after the reform. Column 5 reports the difference-in-differences in the means with the t-test [in brackets]. Column 6 reports the difference in means between children with a sibling born after the reform and children with a sibling born before the reform. T-test are reported in brackets.

	(1)	(2)	(3)	(4)	(5)	(9)
	(p-value)	[q-value]	(p-value)	[q-value]	(p-value)	[q-value]
Use of non-maternal care	0.04		0.03		0.02	
	(0.15)	[0.225]	(0.32)	[0.64]	(0.50)	[0.93]
Informal care	-0.01		-0.02		-0.02	
	(0.78)	[0.78]	(0.83)	[0.92]	(0.84)	[0.93]
Center based care	-0.04		0.01		0.01	
	$(0.03)^{**}$	$[0.08]^{*}$	(0.92)	[0.92]	(0.93)	[0.93]
Weekly hours in all non-maternal care	-3.52		-4.41		-4.47	
	$(0.02)^{**}$	$[0.08]^{*}$	$(0.03)^{**}$	$[0.09]^{*}$	$(0.03)^{**}$	$[0.08]^{*}$
Weekly hours in informal care	-4.45		-6.59		-6.63	
	$(0.04)^{**}$	$[0.08]^{*}$	$(0.02)^{**}$	$[0.09]^{*}$	$(0.02)^{**}$	$[0.08]^{*}$
Weekly hours in center based care	1.86		1.19		1.32	
	(0.67)	[0.78]	(0.83)	[0.92]	(0.83)	[0.93]
Controls						
Child age, gender, language, nbr siblings	n	0	y.	SS	y.	SS
Provincial dummies and city size	n	0	y.	es	y.	SS
Parents' age, education, immigration	n	0	y.	SS	y.	SS
Household income	n	0	n	0	y(	SS

U.
Non-Maternal
for
Results
: Difference-in-differences
3
-

the set of control variables, respectively. Wild-cluster bootstrap p-values are reported in parentheses. The corresponding q-values are reported in columns 2, 4 and 6 [in brackets]. \*\*\*p < 0.01, \*\* p < 0.05, \* p < 0.1

	and Pre-Co	gnitive Out	comes			
	(1)	(2)	(3)	(4)	(5)	(9)
	(p-value)	[q-value]	(p-value)	[q-value]	(p-value)	[q-value]
Hyperactivity score	0.03		0.03		0.03	
	(0.47)	[0.71]	(0.64)	[0.92]	(0.61)	[0.94]
Emotional disorder score	-0.17		-0.16		-0.16	
	(0.33)	[0.71]	(0.35)	[0.875]	(0.35)	[0.87]
Physical aggression score	-0.02		-0.01		-0.01	
	(0.94)	[0.94]	(0.83)	[0.92]	(0.84)	[0.94]
Separation anxiety score	0.07		0.10		0.10	
	(0.32)	[0.71]	(0.11)	[0.55]	(0.11)	[0.55]
Motor and development score	-0.08		-0.05		-0.05	
	(0.57)	[0.71]	(0.92)	[0.92]	(0.94)	[0.94]
Controls						
Child age, gender, language, nbr siblings	n	0	ý	es	y.	es
Provincial dummies and city size	n	0	y	es	y,	es
Parents' age, education, immigration	n	0	y	es	, Ac	es
Household income	n	0	u	10	ye	es

..... Č Ż 4 14.0 Ē J:tt .; Table 4. Diff.

Notes: Each row corresponds to an outcome variable. Columns 1, 3 and 5 report the intention to treat estimate for the specification without any controls, the one with all the controls except household income and the one with all the set of control variables, respectively. Wild-cluster bootstrap p-values are reported in parentheses. The corresponding q-values are reported in columns 2, 4 and 6 [in brackets]. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 5: Differenc	e-in-differer	nces Results	for Health	Outcomes		
	(1)	(2)	(3)	(4)	(5)	(9)
	(p-value)	[q-value]	(p-value)	[q-value]	(p-value)	[q-value]
Underweight	-0.02		-0.02		-0.03	
	(0.53)	[0.87]	(0.41)	[0.93]	(0.40)	[0.93]
Overweight	0.03		0.04		0.04	
	(0.22)	[0.87]	(0.21)	[0.93]	(0.24)	[0.93]
Obese	0.01		-0.01		-0.01	
	(0.73)	[0.87]	(0.89)	[0.93]	(06.0)	[0.93]
Excellent health	-0.02		-0.01		-0.01	
	(0.48)	[0.87]	(0.75)	[0.93]	(0.70)	[0.93]
Injured in the last 12 months	0.01		0.02		0.02	
	(0.80)	[0.87]	(0.69)	[0.93]	(0.69)	[0.93]
Ear infection in the last 12 months	0.02		0.01		0.01	
	(0.36)	[0.87]	(0.71)	[0.93]	(0.70)	[0.93]
Throat/nose infection in the last 12 months	-0.01		0.02		0.02	
months	(0.87)	[0.87]	(0.93)	[0.93]	(0.93)	[0.93]
Controls						
Child age, gender, language, nbr siblings	Ū	0	y.	es	A6	S
Provincial dummies and city size	ũ	0	y.	es	, AG	S
Parents' age, education, immigration	ũ	0	y.	es	A6	S
Household income	n	0	n	0	y(	S
<b>Notes:</b> Each row corresponds to an outcome specification without any controls, the one w control variables, respectively. Wild-cluster l are reported in columns 2, 4 and 6 [in bracke	variable. Co ith all the con pootstrap p-v. ts]. *** p<0.6	olumns 1, 3 $i$ ntrols except alues are rep 01, ** p<0.0	and 5 report t household j oorted in parv 5, * p<0.1.	the intentior income and 1 entheses. Th	1 to treat esti- the one with e correspond	mate for the all the set of ing q-values

_ <b>_</b> _
2
- <b>H</b>
E
5
$\odot$
- <b>-</b>
<u>+</u>
اتة
H
<u> </u>
5
لک ا
2
=
<u></u>
မ၂
~
<u></u>
୍
୍ଧା
ာ
- <u>F</u>
<u>_</u>
•=
느느
·7
ادە
ົວ
Ē
ച
- <b>2</b>
ି ପ୍ରତି ।
<u>ا ا ا</u>
<u></u>
••
5
ച
<b>–</b>

	(1)	(2)	(3)	(4)	(5)	(9)
	(p-value)	[q-value]	(p-value)	[q-value]	(p-value)	[q-value]
Mother in excellent health	-0.04		-0.01		-0.01	
	$(0.05)^{*}$	[0.20]	(0.83)	[0.94]	(0.79)	[0.94]
Father in excellent health	0.0003		0.04		0.03	
	(0.98)	[86.0]	(0.94)	[0.94]	(0.94)	[0.94]
Mother depression score	0.47		0.62		0.63	
	(0.37)	[0.49]	(0.29)	[0.64]	(0.28)	[0.92]
Family dysfunction score	-0.06		-0.03		-0.02	
	(0.13)	[0.26]	(0.32)	[0.64]	(0.46)	[0.92]
Controls						
Child age, gender, language, nbr siblings	n	0	ý	es	y.	es
Provincial dummies and city size	n	0	y	es	y	es
Parents' age, education, immigration	n	0	y	es	y.	es
Household income	u	0	U	0	ý	Se

Table 6: Difference-in-differences Results for Parent and Family Outcomes

ē specification without any controls, the one with all the controls except household income and the one with all the set of control variables, respectively. Wild cluster bootstrap p-values are reported in parentheses. The corresponding q-values are reported in columns 2, 4 and 6 [in brackets]. \*\*\* p<0.01, \*\* p<0.05 and \* p<0.1.

	siblingxpost	(p-value)	[q-value]
Birth variables			
Pre-mature child	0.07	(0.24)	[0.72]
Breast-fed	-0.01	(0.90)	[0.90]
Normal birth weight	-0.004	(0.87)	[0.90]
Mother's work after birth and non-maternal care variables			
Worked after birth	-0.03	(0.78)	[0.84]
Child's age in month when the mother returned to work	0.47	(0.46)	[0.84]
Hours of work	-0.32	(0.54)	[0.84]
Use of non-maternal care	0.01	(0.84)	[0.84]
Weekly hours in all non-maternal care	-1.43	(0.28)	[0.84]
Pre-cognitive outcomes			
Motor and development score	-0.06	(0.73)	[0.73]
Health variables			
Has excellent health	0.03	(0.08)*	[0.32]
Had an injury in the last 12 months	0.03	(0.26)	[0.37]
Had an ear infection in the last 12 months	-0.06	(0.28)	[0.37]
Had a throat/nose infection in the last 12 months	-0.10	(0.42)	[0.42]
Parent and family outcomes			
Mother in excellent health	0.04	(0.49)	[0.85]
Father in excellent health	0.01	(0.64)	[0.85]
Mother depression score	-0.04	(0.93)	[0.93]
Family dysfunction score	-0.16	(0.35)	[0.85]
Child temperament			
How easily upset	-0.04	(0.72)	[0.91]
How loud crying when upset	0.02	(0.78)	[0.91]
How much crying	-0.03	(0.91)	[0.91]
Difficulty to calm	-0.12	(0.17)	[0.38]
How often irritable	-0.06	(0.49)	[0.78]
How much smile/laugh	-0.19	(0.19)	[0.38]
Average mood	-0.17	(0.12)	[0.38]
How changeable is mood	-0.14	(0.08)*	[0.38]

Table 7: Test for the Common Trend Assumption (Children aged 0-1 year)

**Notes:** Each row corresponds to an outcome variable. The first column reports the intention to treat estimate for the sample of interest were 0-1 year old. Control variables include child, parent and family variables (except household income) and provincial fixed effect. The second and third columns report the wild- cluster bootstrap p-value and the q-values respectively. \*\*\* p<0.01, \*\* p < 0.05, \*p<0.1.

	siblingxpost	(p-value)	[q-value]
Non-Cognitive and Pre-Cognitive Outcomes			
Hyperactivity score	0.02	(0.75)	[0.75]
Emotional disorder score	-0.08	(0.09)*	[0.27]
Physical aggression score	-0.12	(0.16)	[0.27]
Separation anxiety score	-0.15	(0.14)	[0.27]
Motor and Development Score	0.04	(0.37)	[0.46]
Health outcomes			
Underweight	-0.01	(0.84)	[0.84]
Overweight	0.02	(0.26)	[0.49]
Obese	0.04	(0.27)	[0.49]
Excellent health	0.08	(0.41)	[0.57]
Injured in the last 12 months	-0.01	(0.55)	[0.64]
Ear Infection in the last 12 months	-0.05	(0.13)	[0.49]
Throat/nose infection in the last 12 months	-0.07	(0.28)	[0.49]
Parent and family outcomes			
Mother in excellent health	0.09	(0.08)*	[0.28]
Father in excellent health	0.06	(0.46)	[0.46]
Mother depression score	0.03	(0.22)	[0.29]
Family dysfunction score	0.08	(0.14)	[0.28]
Non-maternal care			
Use of non-maternal care	0.01	(0.51)	[0.72]
Weekly hours in all non-maternal care	-1.67	(0.72)	[0.72]

## Table 8: Falsification Test- Children aged 2-3 years old in 1996 to 1999

Notes: Each row corresponds to an outcomes variable. The first column reports the intent to treat estimate for children aged 2-3 years old in 1996 until 1999. This provide a falsification test. Control variables include child, parents and family variables as well as provincial fixed effects. The second and third columns report the wild cluster bootstrap p-values and the q-values respectively. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1. The sample consists of 4826 observations.

	siblingxpost	(p-value)	[q-value]
Cognitive outcomes			
PPVT standardized score	0.05	(0.67)	[0.67]
Who Am I?	0.13	(0.28)	[0.42]
Number knowledge	0.07	(0.14)	[0.42]
Behavioural outcomes			
Hyperactivity score	-0.08	(0.88)	[0.91]
Emotional disorder score	0.01	(0.91)	[0.91]
Physical aggression score	-0.04	(0.76)	[0.91]
Indirect aggression score	-0.17	(0.31)	[0.91]
Health outcomes			
Underweight	0.02	(0.09)*	[0.23]
Overweight	0.04	(0.39)	[0.49]
Obese	-0.09	(0.73)	[0.73]
Excellent health	-0.1	(0.09)*	[0.23]
Had an injury in the last 12 months	0.03	(0.32)	[0.49]
Parent and family outcomes			
Mother in excellent health	-0.002	(0.98)	[0.98]
Father in excellent health	-0.01	(0.79)	[0.98]
Mother depression score	-0.16	(0.55)	[0.98]
Family dysfunction score	-0.02	(0.82)	[0.98]
Non-maternal care			
Use of non-maternal care	-0.0002	(0.99)	[0.99]
Weekly hours in all non-maternal care	-0.51	(0.83)	[0.99]

### **Table 9: Medium Run Results**

**Notes:** Each row corresponds to an outcome variable. The first column reports the intention to treat estimate after following the sample of interest one cycle until they are aged 4-5 years old. Control variables include child, parent and family variables (except household income) and provincial fixed effects. The second and third columns report the wild-cluster bootstrap p-values and the q-values. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 10: Heterogeneous Effects by Child Gender					
	(1)	(2)	(3)	(4)	
	Во	oys	Gi	irls	
	(p-value)	[q-value]	(p-value)	[q-value]	
Non-cognitive and pre-cognitive outcomes					
Hyperactivity score	0.08		-0.08		
	(0.15)	[0.38]	(0.41)	[0.67]	
Emotional disorder score	-0.19		-0.09		
	(0.35)	[0.59]	(0.62)	[0.78]	
Physical aggression score	-0.02		-0.01		
	(0.80)	[0.80]	(0.80)	[0.80]	
Separation anxiety score	0.08		0.14		
	(0.04)**	[0.20]	(0.14)	[0.67]	
Motor and development score	0.06		-0.15		
	(0.73)	[0.80]	(0.28)	[0.67]	
Child health outcomes					
Underweight	-0.03		-0.04		
	(0.37)	[0.87]	(0.65)	[0.86]	
Overweight	0.01		0.07		
	(0.40)	[0.87]	(0.52)	[0.86]	
Obese	0.04		-0.04		
	(0.58)	[0.87]	(0.41)	[0.86]	
Excellent health	-0.03		0.01		
	(0.74)	[0.87]	(0.65)	[0.86]	
Injured in the last 12 months	0.02		0.01		
	(0.62)	[0.87]	(0.84)	[0.86]	
Ear infection in the last 12 months	-0.004		0.04		
	(0.91)	[0.91]	(0.53)	[0.86]	
Throat/nose infection in the last 12 months	-0.08		0.11		
	(0.04)**	[0.28]	(0.86)	[0.86]	
Parent and family outcomes					
Mother in excellent health	-0.11		0.11		
	(0.05)**	[0.10]	(0.72)	[0.96]	
Father in excellent health	-0.002		0.06		
	(0.95)	[0.98]	(0.96)	[0.96]	
Mother depression score	-0.004		0.23		
	(0.98)	[0.98]	(0.03)**	[0.12]	
Family dysfunction score	-0.07	50 107	0.03	F.0. 0. (7)	
	(0.03)**	[0.10]	(0.50)	[0.96]	

Each row corresponds to an outcome variable. Columns 1 and 2 report the intention to treat estimates with the wild-cluster bootstrap p-value (in parentheses) and the corresponding q-value [in brackets] for the children whose mothers have a university degree (2544 observations). Columns 3 and 4 report the results for the children whose mothers do not have a university (2340 observations). \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

	(1)	(2)	(3)	(4)	
				Non-university	
	Universit	University degree		gree	
	(p-value)	[q-value]	(p-value)	[q-value]	
Non-cognitive and pre-cognitive outcomes					
Hyperactivity score	0.09		-0.02		
	(0.03)**	[0.14]	(0.93)	[0.96]	
Emotional disorder score	-0.07		-0.25		
	(0.63)	[0.63]	(0.67)	[0.96]	
Physical aggression score	-0.10		0.11		
	(0.20)	[0.43]	(0.53)	[0.96]	
Separation anxiety score	0.10		0.10		
	(0.26)	[0.43]	(0.20)	[0.96]	
Motor and development score	-0.03		-0.02		
-	(0.58)	[0.63]	(0.96)	[0.96]	
Child health outcomes	\$ <u>}</u>		. ,		
Underweight	-0.07		0.02		
C	(0.17)	[0.52]	(0.82)	[0.82]	
Overweight	0.05		0.04		
	(0.08)*	[0.52]	(0.26)	[0.82]	
Obese	0.02		-0.04		
	(0.66)	[0.66]	(0.72)	[0.82]	
Excellent health	-0.03	[]	0.01	[]	
	(0.37)	[0.52]	(0.56)	[0.82]	
Injured in the last 12 months	0.02	[]	0.02	[]	
	(0.48)	[0.56]	(0.77)	[0.82]	
Ear infection in the last 12 months	0.06	[]	-0.03	[]	
	(0.35)	[0 52]	(0.54)	[0.82]	
Throat/nose infection in the last 12 months	0.05	[0.0-]	0.11	[0:0-]	
	(0.34)	[0 52]	(0.61)	[0.82]	
Parent and family outcomes	(0.5 1)	[0.02]	(0.01)	[0:02]	
Mother in excellent health	-0.05		0.05		
Wother in excenent nearth	(0.22)	[0 44]	(0.83)	[0.83]	
Father in excellent health	-0.03	[ידי.ט]	0.11	[0.05]	
Tatlet in excellent health	-0.05	[0.66]	(0.70)	[0.83]	
Mother depression score	0.00)	[0.00]	0.15	[0.05]	
	(0.15)	[0 44]	(0.13)	[0 82]	
Family dysfunction score	0.13)	[0.44]	0.04	[0.03]	
ranning dystanction score	-0.00	[0 55]	-0.04	[0 02]	
	(0.41)	10.331	(0.33)	10.83	

Table 11: Heterogeneous effects by mother education

Each row corresponds to an outcome variable. Columns 1 and 2 report the intention to treat estimates with the wild-cluster bootstrap p-value (in parentheses) and the corresponding q-value [in brackets] for the children whose mothers have a university degree (2544 observations). Columns 3 and 4 report the results for the children whose mothers do not have a university (2340 observations). \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

	(1)	(2)	(3)	(4)
	Income be	Income below median		ove median
	(p-value)	[q-value]	(p-value)	[q-value]
Non-cognitive and pre-cognitive outcomes				
Hyperactivity score	0.002		0.02	
	(0.97)	[0.97]	(0.83)	[0.91]
Emotional disorder score	-0.26		-0.05	
	(0.34)	[0.57]	(0.91)	[0.91]
Physical aggression score	-0.12		0.05	
	(0.20)	[0.57]	(0.56)	[0.91]
Separation anxiety score	0.11		0.11	
	(0.54)	[0.68]	(0.61)	[0.91]
Motor and development score	-0.19		0.09	
	(0.27)	[0.57]	(0.73)	[0.91]
Child health outcomes				
Underweight	0.01		-0.05	
	(0.69)	[0.96]	(0.18)	[0.63]
Overweight	0.04		0.04	
	(0.20)	[0.96]	(0.45)	[0.76]
Obese	0.03		-0.05	
	(0.62)	[0.96]	(0.52)	[0.76]
Excellent health	-0.07		0.05	
	(0.44)	[0.96]	(0.13)	[0.63]
Injured in the last 12 months	0.03		0.01	
	(0.37)	[0.96]	(0.85)	[0.90]
Ear infection in the last 12 months	-0.002		0.04	
	(0.96)	[0.96]	(0.54)	[0.76]
Throat/nose infection in the last 12 months	0.02		0.01	
	(0.83)	[0.96]	(0.90)	[0.90]
Parent and family outcomes				
Mother in excellent health	-0.07		0.06	
	(0.29)	[0.58]	(0.07)**	[0.28]
Father in excellent health	0.02		0.05	
	(0.91)	[0.91]	(0.91)	[0.91]
Mother depression score	0.15		0.09	
	(0.14)	[0.56]	(0.41)	[0.82]
Family dysfunction score	-0.01		-0.05	
	(0.89)	[0.91]	(0.67)	[0.89]

Table 12: Heterogeneous effects by household income

Each row corresponds to an outcome variable. Columns 1 and 2 report the intention to treat estimates with the wild-cluster bootstrap p-value (in parentheses) and the corresponding q-value [in brackets] for children of families below the median income of \$60,000 (2487 observations). Columns 3 and 4 report the results for children of families above the median income (2485 observations). \*\*\*p<0.01, \*\*p<0.05, \*\*p<0.1.

	(1)	(2)	(3)	(4)
	Non hea	Non heavy users		users
	(p-value)	[q-value]	(p-value)	[q-value]
Non-cognitive and pre-cognitive outcomes				
Hyperactivity score	0.12		-0.02	
	(0.04)**	[0.18]	(0.72)	[0.75]
Emotional disorder score	-0.002		-0.24	
	(0.99)	[0.98]	(0.46)	[0.75]
Physical aggression score	0.13		-0.05	
	(0.82)	[0.98]	(0.75)	[0.75]
Separation anxiety score	0.17		0.07	
	(0.53)	[0.98]	(0.28)	[0.75]
Motor and development score	-0.03		-0.10	
	(0.98)	[0.98]	(0.55)	[0.75]
Child health outcomes				
Underweight	-0.03		0.003	
	(0.41)	[0.72]	(0.79)	[0.89]
Overweight	0.05		0.01	
	(0.14)	[0.72]	(0.78)	[0.89]
Obese	-0.02		-0.01	
	(0.79)	[0.89]	(0.89)	[0.89]
Excellent health	-0.06		0.11	
	(0.23)	[0.72]	(0.01)**	[0.06]*
Injured in the last 12 months	0.03		-0.05	
	(0.57)	[0.80]	(0.25)	[0.89
Ear infection in the last 12 months	0.03		-0.02	
	(0.40)	[0.72]	(0.54)	[0.89]
Throat/nose infection in the last 12 months	-0.02		-0.03	
	(0.89)	[0.89]	(0.66)	[0.89]
Parent and family outcomes				
Mother in excellent health	-0.06		0.10	
	(0.18)	[0.50]	(0.17)	[0.68]
Father in excellent health	0.03		0.07	
	(0.88)	[0.88]	(0.55)	[0.88]
Mother depression score	0.17		0.03	
	(0.25)	[0.50]	(0.83)	[0.88]
Family dysfunction score	-0.02		-0.01	
	(0.47)	[0.63]	(0.88)	[0.88]

Table 13: Heterogeneous effects by non-maternal care usage when the child was an infant

Each row corresponds to an outcome variable. Columns 1 and 2 report the intention to treat estimates with the wild-cluster bootstrap p-value (in parentheses) and the corresponding q-value [in brackets] for non-heavy users of child-care when infant defined as a spending less than 20 hours per week in non-maternal care when the child was aged 0-1 years old (3641 observations). Columns 3 and 4 report the results for heavy users of child-care when infant defined as spending more than 20 hours per week in non-maternal care when the child was aged 0-1 years old (1519 observations). \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

	(1)	(2)	(3)	(4)
	Children whose		All other	
	mothers were working		Children	
	in the previous cycle			
	(p-value)	[q-value]	(p-value)	[q-value]
Non-maternal care				
Use of care	-0.004		0.03	
	(0.92)	[0.92]	(0.42)	[0.57]
Weekly hours in non-maternal care	-4.90		-1.93	
	(0.03)**	[0.06]*	(0.57)	[0.57]
Non-cognitive and pre-cognitive outcomes				
Hyperactivity score	0.03		-0.05	
	(0.84)	[0.84]	(0.68)	[0.92]
Emotional disorder score	-0.15		-0.18	
	(0.33)	[0.84]	(0.68)	[0.92]
Physical aggression score	0.06		0.07	
	(0.60)	[0.84]	(0.74)	[0.92]
Separation anxiety score	0.05		0.06	
	(0.68)	[0.84]	(0.65)	[0.92]
Motor and development score	-0.04		-0.01	
	(0.79)	[0.84]	(0.92)	[0.92]
Child health outcomes				
Underweight	-0.04		-0.08	
	(0.34)	[0.48]	(0.64)	[0.81]
Overweight	0.04		-0.06	
	(0.36)	[0.48]	(0.68)	[0.81]
Obese	-0.06		0.06	
	(0.24)	[0.48]	(0.40)	[0.81]
Excellent health	0.08		-0.07	
	(0.06)*	[0.20]	(0.20)	[0.81]
Injured in the last 12 months	0.02		0.02	
	(0.63)	[0.74]	(0.73)	[0.81]
Ear infection in the last 12 months	-0.02		0.01	
	(0.80)	[0.80]	(0.81)	[0.81]
Throat/nose infection in the last 12 months	-0.04		0.08	
	(0.04)**	[0.20]	(0.50)	[0.81]
Parents and Family Outcomes				
Mother in excellent health	0.12		-0.15	
	(0.04)**	[0.16]	(0.03)**	[0.12]
Father in excellent health	0.09		0.03	
	(0.72)	[0.72]	(0.63)	[0.82]
Mother's depression score	0.20		0.02	
	(0.45)	[0.72]	(0.82)	[0.82]
Family dysfunction score	0.11		-0.18	
	(0.58)	[0.72]	(0.06)*	[0.12]

#### **Table 14: Treatment on the treated estimates**

**Notes:** Each row corresponds to an outcome variable. Columns 1 and 2 report the results of the differencein-differences regression with the wild-cluster bootstrap p-value (in parentheses) and the corresponding qvalue [in brackets] for children whose mothers were working in the previous cycle (2639 observations). Columns 3 and 4 report the results for all other children (1688 observations). \*\*\* p<0.01, \*\*p<0.05, \*p<0.1.

# **Figures**



Figure 1: Job Protection by Province from 1996 to 2002



Figure 2: Trends from 1996-1997 until 2002-2003 for children with and those without younger siblings

# **Appendix: NLSCY Scores**

This appendix provides the exact questions used in the calculation of NLSCY scores used in this paper. The questions are taken from the NLSCY questionnaires for cycles 3 and 5.

- Motor and Development Score: ranges from 0 to 15 based on 15 questions, taking the value 1 if the answer is Yes, 0 if No. Questions asked on whether the child had ever:
  - Let someone know, without crying, that wearing wet (soiled) pants or diapers bothered him
  - Spoken a partial sentence of 3 words or more
  - Walked up stairs by himself without holding on to a rail
  - Washed and dried his hands without any help except for turning the water on and off
  - Counted 3 objects correctly
  - Gone to the toilet alone
  - Walked upstairs by himself with no help, stepping on each step with only one foot
  - Known his own age and sex
  - Said the names of at least 4 colors
  - Pedalled a tricycle at least 10 feet
  - Done a somersault without help from anybody
  - Dressed himself/herself without any help except for tying shoes (and buttoning the backs of outfits)
  - o Said his/her first and last name together without someone's help
  - Counted out loud up to 10
  - Drawn a picture of a man or woman with at least 2 parts of the body other than a head
- Hyperactivity/Inattention score: ranges from 0 to 10 based on 5 questions with a higher score indicating the behaviors associated with hyperactivity/inattention. The answer can take the value 0 (not true), 1 (sometimes true) or 2 (often true). How often would you say the this child:
  - Can not sit still or restless
  - Is easily distracted
  - Can not concentrate, can not pay attention for long

- Can not settle on anything for more than a few moments
- Is inattentive
- Emotional disorder score: ranges from 0 to 12 based on 6 questions with a higher score indicating behaviors associated with emotional disorder. The answer can take the value 0 (not true), 1 (sometimes true) or 2 (often true) How often would you say the this child
  - Seems to be unhappy or sad
  - Is not as happy as other children
  - Is too fearful or nervous
  - Is worried
  - Is nervous, high strung or tense
  - Has trouble enjoying himself/herself
- Physical aggression score: ranges from 0 to 16 based on 8 questions with a higher score indicating behaviors associated with physical aggression and opposition. The answer can take the value 0 (not true), 1 (sometimes true) or 2 (often true). How often would you say the this child:
  - Is defiant
  - Gets into many fights
  - o Does not change his/her behavior after punishment
  - Has temper tantrums or hot temper
  - Does not seem to feel guilty after misbehaving
  - o When somebody accidently hurts him, he/she reacts with anger and fighting
  - Has angry moods
  - Kicks, bites or hits other children
- Separation anxiety score: ranges from 0 to 10 based on 5 questions with a higher score indicating behaviors associated with separation anxiety. The answer can take the value 0 (not true), 1 (sometimes true) or 2 (often true). How often would you say the this child:
  - Cries a lot
  - Clings to adults or is too dependent
  - o Constantly seeks help
  - Gets too upset when separated from parents
  - Does not want to sleep alone

- Depression score: ranges from 0 to 36 based on 12 questions with a higher score indicating the presence of depression symptom. The answers can take the value 0 (rarely or none of the times, less than 1 day), 1 (some or little of the time, 1-2 days), 2 (Occasionally or a moderate amount of time, 2-3 days) and 3 (Most or all of the time, 5-7 days). How often have you felt or behaved this way during the last week:
  - I do not feel like eating
  - I felt that I could not shake off the blues even with help from families or friends
  - o I had trouble keeping my mind on what I was doing
  - I felt depressed
  - I felt that everything I did was an effort
  - I felt hopeful about the future (reversed)
  - My sleep was restless
  - I was happy (reversed)
  - I felt lonely
  - I enjoyed life (reversed)
  - I had crying spells
  - I felt that people disliked me
- Family dysfunction score: ranges from 0 to 36 based on 12 questions with a higher score indicating more family dysfunction. Points are given from 0 (Strongly Agree) to 3 (Strongly Disagree)
  - Planning family activities is difficult because we misunderstand each other (reversed)
  - In times of crisis we can turn to each other for support
  - We cannot talk to each other about sadness we feel(reversed)
  - o Individuals (in the family) are accepted for what they are
  - We avoid discussing our fears or concerns (reversed)
  - We express feelings to each other
  - There are lots of bad feelings in our family (reversed)
  - We feel accepted for what we are
  - Making decisions is a problem for our family (reversed)
  - We are able to make decisions about how to solve problems

- We don't get along well together (reversed)
- We confide in each other

## **Connecting Text**

The first essay of this thesis contributes to the literature on the determinants of human capital by studying the impact of one component of parental involvement, maternal care, on child development. The essay employs a difference-in-differences analysis while controlling for the small number of clusters to estimate the causal impact of maternal care on the cognitive, non-cognitive and health outcomes of children aged 2-3 years. The second essay uses a similar methodology and contributes to the literature on the returns to adult's human capital. Specifically, we consider the human capital of immigrants. We use difference-in-differences and triple differences analysis to study the effects of a change in the point system on immigrants' stock of human capital and its return in the labour market.

Both essays use novel identification strategies through natural experiments to estimate the causal impact of maternal care and immigration policies, respectively. In a world with limited resources, disentangling causation from correlation is crucial for impact evaluation and for devising proper policy implications. In addition, the two essays contribute to the policy evaluation literature. In the first, we study the spillover effects of parental leaves on older siblings. The second essay evaluates immigration policies, specifically, the point system.

# Chapter 3

# The Effects of a Change in the Point System on Immigration: Evidence from the 2001 Quebec Reform

# **3.1 Introduction**

Many developed countries have chosen, or are actively considering, a point system that selects immigrants based on specific characteristics (education, language skills, age, experience, occupation, or motivation). This system has been implemented in Canada, Australia (already accounting for 67 and 63 percent of immigrants to Canada and Australia<sup>17</sup>), New Zealand and the UK. Other countries choose some immigrants based on their skills (the US with the H1B visa, and the EU with the Blue Card program), and are actively considering adopting a more formal point system (Aydemir, 2011). A major advantage of a point system over other immigration policies, such as family reunification or refugees, is that the points can be changed by governments to shape immigration depending on their economic, political, or cultural needs. A point system thus provides a useful policy lever to shape immigration.

Skeptics argue that tweaking the point system will not affect the composition of immigration: more fundamental forces, such as returns to skills and geographic proximity, dominate the nuances of selection systems (Jasso and Rosenzweig, 2008). For example, if Canada and Australia had the same point system, it is unlikely that they will have the same immigration, considering the different neighboring countries. Skeptics further argue that changes in points for observables will not affect unobservables, such as motivation or ability, and the integration or economic performance of immigrants (Borjas, 1991).

Considering the wide use of the point system despite these criticisms, it is vital to know whether changing the points can affect the immigrant's composition and performance. Yet, empirically, it is very difficult to answer this question. Comparing countries with different point systems is unlikely to deliver the causal impact of selection systems since countries differ on many levels, such as the returns to skills and geographic proximity of host and source countries. Immigrants to different countries might be systematically different for reasons other than a difference in the point

<sup>&</sup>lt;sup>17</sup> Citizenship and Immigration Canada (CIC) 2013; Australian Bureau of Statistics 2013.

system (Borjas, 1991; Antecol, Cobb-Clark and Trejo 2003). Within-country analyses hold more promise since factors such as geography and history are held constant (Green and Green 1995). However, when the points change, they often do so at a national level (Beach, Worswick and Green, 2011). The comparison of immigrants selected before and after changes may be confounded by business cycle or cohort effects: more recent immigrants may perform worse simply because of deteriorating macroeconomic conditions and fewer years to integrate, not because of a different point system. A "before/after" comparison is unlikely to deliver the causal impact of a change in the point system.

This paper is the first to answer these important questions using a large change in the point system that occurred in only one of the ten provinces of Canada, which allows us to use a difference-in-differences and triple difference analysis. In 2001, a new immigration policy was implemented in Quebec (subsequently QC): points for education (specifically to bachelor's degree) and the French language increased, while points for a subjective assessment of the immigrants' "adaptability", i.e., personal qualities and motivation to integrate in society, decreased.

Relative to the Rest of Canada (subsequently ROC), in which there was no change in the point system over the same period, we find that immigrants selected in QC after the reform are 6 percentage points more likely to hold a bachelor's degree than immigrants selected before, and 4 percentage points more likely to speak French only (as opposed to speaking both French and English). This indicates that immigrants' characteristics respond to changes in the point system. The point system thus represents an effective policy lever to shape immigration.

Despite being more educated and speaking more the local language, the new immigrants did not perform better on the labour market after the reform (in terms of employment and earnings). One explanation is that foreign bachelor's degree holders have difficulties integrating in the QC labour market: all else equal, we find that they were on average 12 percentage points less employed than those not holding a foreign bachelor's degree. Other studies have shown that they take twice as much time as high school graduates to find a job (Godin and Pinsonneault, 2004; Renaud and Cayn, 2006). Immigrants speaking only French also have difficulties integrating: we find that they were 8 percentage points less employed than bilingual immigrants. These difficulties to integrate may have been exacerbated by the policy shift. If immigrants with a foreign bachelor's degree and speaking French only are not highly substitutable for other workers and hence are in a very narrow labour market, we would expect an increase in their number to lead to worse labour market

outcomes, all else equal. Finally, an alternative explanation for the poor performance of new immigrants is that the sharp drop in points for "adaptability" was detrimental to the selection process.

In any case, labour market performance was not the primary goal of the reform. The stated objective of the reform was to "increase the French-speaking immigration, while maintaining the socioeconomic requirements"<sup>18</sup>. This new immigration policy was decided by the governing party, the Parti Québécois, whose two first priorities are independence from Canada and protection of the French language (programme du Parti Québécois, 2001). The control of immigration was seen as a key issue to realize these goals. After the 1995 referendum of independence was lost by a narrow margin, but in which Francophones massively voted for independence, party officials stated that the next referendum would be successful with only a few percentage more of French speakers (Cardinal, 2005). The underlying motivations for the 2001 reform were thus more political and cultural than economic. Relative to this goal of increasing the number of French-speaking immigrants with no deterioration in labour market outcomes, the reform was an unequivocal success. These results thus show that the points can be changed to shape immigration according to policy goals.

A common concern when comparing QC to ROC is that QC has very different cultural, political and economic conditions from ROC. In other words, QC may be on different trends from ROC. To address this criticism, we employ a triple differences methodology. To find additional control groups within QC, we use other immigrants who would fall short of the passing grade based on their observable characteristics but who immigrated through other programs such as family-reunification or refugees. We also use natives as an additional control group. These control groups reside in QC, and are therefore influenced by the same cultural, political and economic conditions as our treatment group. But they did not go through the point system, and should thus be unaffected by the 2001 QC reform. Using these additional control groups, we find no support for differing trends between QC and ROC over the period, and our triple differences confirm our findings.

This paper is the first to use a change in points in a sub-unit of a country to analyze its impacts on immigrants in a difference-and-differences analysis. The existing literature on the point system has compared immigrants coming through this system, to immigrants entering through other systems, such as family-reunification immigrants and refugees, in Canada and Australia (Borjas

<sup>&</sup>lt;sup>18</sup> Plan Stratégique 2001-2004, Ministère des Relations avec les Citoyens et de l'Immigration.

1993; Miller 1999; Cobb-Clark 2000, 2003; Abott and Beach 2011; Aydemir 2011) or to immigrants coming before the implementation of the points system (Green and Green 1995). Despite the importance of this literature, it does not provide an answer to the question we ask, i.e., will a change in the point system affect the immigrants' composition. To answer this particular question, other papers more directly related to ours have looked at national changes in points (Beach et al., 2011), or cross country differences between Canada, Australia and the US (Borjas, 1991; Antecol et al., 2003). Our paper is the first to use a within country difference-in-differences and triple differences methodology, which controls for cohort, business cycle, cultural, political, and economic effects. We show that more points on French attracted more French speaking immigrants. More points on bachelor's degrees attracted more bachelor's degrees' holders. In this particular case, this was not accompanied with better labour market outcomes since the return to foreign bachelors' degrees and speaking French only are negative on the QC job market. One implication is that assigning more points to characteristics that fetch a higher return (such as bilingualism, Master and Ph.D. in the case of QC) could be associated with better labour market outcomes.

The paper proceeds as follows: Section 3.2 provides background on the point system and on the related impact evaluation literature. Section 3.3 presents our identification strategy. Section 3.4 describes the data, and Section 3.5 the methodology used. Section 3.6 presents the results. Section 3.7 provides a discussion of these results, and Section 3.8 concludes.

## **3.2 The Point System**

#### **3.2.1.** Description

In 1967, Canada became the first country in the world to initiate a point system. Immigrants to Canada are classified into three categories: family class, humanitarian or refugee class, and the economic class. Only the last class is assessed through the point system. Figure 1 shows the number of Canadian immigrants by category from 1999 till 2003. Since 1999, more than 55 percent of immigrants are admitted under the economic class each year<sup>19</sup> (CIC 2007).

<sup>&</sup>lt;sup>19</sup> Note that the economic class comprises different subcategories such as skilled worker class, business class and investor class. In this paper, we study the skilled worker point system. 89 percent of the economic class' principle applicants apply under the skilled workers program (CIC 2007).

The point system is a color-blind system that allocates a number of points to some observable skills of immigrants. The main categories are education, training, experience, occupational sector in demand, arranged employment, regional and labour markets' needs, age, language abilities, adaptability, and characteristics of the spouse and children if applicable (McWhinney 1998, see Table 1 for the exact points on each category in ROC over the period 1999-2002). If the applicant garners more points than a specified threshold<sup>20</sup>, he/she is admitted into the country.

#### **3.2.2 Existing Literature**

To evaluate how a change in points affects immigrant composition, the existing literature has followed three strategies. First, some studies have used time variation in the point system (Beach et al., 2011). There is considerable time variation in the points system. The Canadian government alternated between "labour market specific" models, "human capital" models and a combination of both (see O'Shea (2009) for a description of the changes that occurred in the Federal Skilled Workers grids from 1967 until 2008). The "labour market specific" models assign points based on current market needs, and are short-term in nature. The "human capital" models are more long-term, and assign points on characteristics thought to help immigrants integrate in Canada, such as education, age, language and experience. These large and frequent changes all serve to illustrate that there is no consensus on how the points should be set, and how they affect the immigrants' composition. On an econometric level, one issue with comparing individuals who immigrated before and after changes is that their performance on the labour market may be affected not only by changes in the points system, but also by cohort effects or coincidental macroeconomic shocks.

A second strand of the literature has used cross-country comparisons. There are wide differences between point systems implemented in different countries. For example, Australia assigns 46 percent of the passing grade for applicants aged between 25 and 32 years old, whereas Canada assigns only 15 percent for 21-44 years old. Australia assigns 31 percent of the passing grade for Ph.D. and 23 percent for master and bachelor, whereas Canada assigns 37 percent for Ph.D. and 34 percent for master and bachelor's degrees. In the US, there is no formal point system, but 16 percent of total immigration goes through an employment-based program for skilled workers. This program selects holders of advanced degrees, people with exceptional skills, or immigrants with

<sup>&</sup>lt;sup>20</sup> The passing grade was 70 out of 100 between 1999 and 2002.

employers who demonstrate that no American could fill the job. There are no points for language skills, age, experience, or motivation. Understanding the impact of these wide differences is critical. Borjas (1993) finds that American immigrants to Canada (who went through the point system) perform worse than Canadian immigrants to the US (who did not go through the point system). When comparing Canada, Australia and the US, Borjas (1991) and Antecol et al. (2003) find that immigrants to Australia perform better than immigrants to Canada or the US. However, it is difficult to attribute these findings to the point system, since geographic, economic and political conditions differ greatly across countries.

Finally, a third strand of the literature has compared immigrants who went through the point system to immigrants coming before the introduction of the points system (Green and Green, 1995), or to immigrants who immigrated through other systems, e.g., family reunification and refugees, in Canada and Australia (Borjas 1993; DeSilva 1997; Barett 1998; Miller 1999; Cobb-Clark 2000, 2003; Wanner 2003; Sweetman and Warman 2012). Overall, a consensus emerged from this literature that immigrants selected through the point system are more educated than others. The effects on labour market outcomes are more ambiguous. One set of papers argues that the returns to education of the selected immigrants are positive (Beach et al 2011), and that the difference in earnings with other immigrants persists over time (Abott and Beach 2011; Sweetman and Warman 2012). Other papers find that the returns to education are very small (Aydemir 2011) and that the earnings of the different classes converge rapidly (De Silva 1997; Wanner 2003). These negative findings may not be evidence that the point system is inappropriate, rather that the points are set on the inappropriate categories. In any case, this important literature does not answer the particular question we ask, i.e., does a change in points affect the immigrants' composition.

To answer our question of interest, the ideal experiment would assign different point systems to randomized group of immigrants, and follow their labour market success. In the absence of such an experiment, we use the following identification strategy.

# **3.3 Identification Strategy**

### 3.3.1 The 2001 QC Reform

In this paper, we exploit the fact that QC is the only province in Canada which can set its own point system<sup>21</sup>. In 2001, QC dramatically changed its point system, while ROC did not. Our identification strategy is to compare the characteristics and labour market performance of immigrants who immigrated to QC before and after 2001, relative to those immigrating to ROC before and after 2001. The reform, which we describe in greater detail below, changed points for three categories: education, language, and "adaptability".

Table 1 shows that the points for education increased from 25 to 32 percent (as a proportion of the passing grade<sup>22</sup>) in QC in 2001. The reform specifically focused on bachelor's degrees. Out of the 6 modifications to the points on education, 5 concerned bachelor's degrees (or other degrees at the bachelor level, e.g., postsecondary 3 years, a second university specialty of 1 or 2 years, which can be a second bachelor). Only one modification awarded one more point (out of 60) to Master students. No changes were made to Ph.D. or Medical degree. Moreover, a new section on spouse's education was added in the auxiliary Grille d'employabilité et de Mobilité Professionnelle that granted extra points only to a bachelor education. No extra points were granted to spouses with a Master, Ph.D., or MD. If the point system is effective, immigrants should be more educated, in particular have more bachelor's degrees after the implementation of this reform.

Points for French increased in 2001 in QC by 5 percentage points, and did not change in ROC, as either a first or second language (ROC does not make a distinction between the two official languages, French and English).

Finally, Table 1 shows that the points for "adaptability" decreased in QC in 2001 by 17 percentage points, but stayed constant in ROC. The "adaptability" category consists in a subjective assessment of the ability of an immigrant to successfully integrate into the QC society following an interview with the applicant. Adaptability points are given based on personal qualities, motivation, knowledge of QC, visits to QC, and a connection with a resident in QC. In 2001, there was a decrease in the first two sub-categories only: personal qualities (23 to 10 percent), and motivation (8 to 3 percent). Appendix 1 details all the criteria used to evaluate the immigrants' personal qualities and motivation.

<sup>&</sup>lt;sup>21</sup> In 1991, the Canada-Quebec agreement granted QC the exclusive right to select its immigrants and design its own point system based essentially on the same major characteristics as the ones used in ROC (Kostov 2008).

<sup>&</sup>lt;sup>22</sup> We report the changes in the point system as a proportion of the passing grade to account for the changes in the passing grade. For single applicants in Quebec, the passing grade changed from 65 (out of 115 total points available) in 1996 to 60 (out of 106) in 2001. For married applicants, it changed from 70 (out of 132) to 68 (out of 123).

In summary, the 2001 QC reform made three changes: more points for education and French knowledge, and fewer for adaptability. There were no significant changes in other categories, as shown in Table 1.

As explained above, there were no significant changes to the point system in ROC in 2001. In 2002, ROC implemented a reform, called the Immigration and Refugee Act (IRPA). For the purpose of this paper, we will ignore this change since only 8 percent of the federal skilled workers immigrants who arrived to Canada in 2003 were evaluated under the IRPA (CIC 2010).

If the point system works, the 2001 QC reform may attract more educated, French-speaking, less "adaptable" immigrants. This may have profound positive consequences for the labour market performance of these new immigrants. The existing literature on immigration usually finds that more educated immigrants perform better on the labour market (Beach et al, 2011). Speaking the local language may be beneficial, and is posited by immigration officials as being beneficial<sup>23</sup>.

Yet, the existing literature from QC has found that foreign bachelor degree holders fared relatively poorly on the QC labour market (Godin and Pinsonneault, 2004; Renaud and Cayn, 2006). Using QC's administrative data on 1,579 immigrants, they find that the time to first employment is higher for bachelor's degree holders (20 weeks) than for any other education category, including high school graduates (9 weeks). Explanations for their findings are that foreign bachelor's degree holders. By increasing points on the French language, the reform may attract cohorts of immigrants speaking French only. If bilingualism, not speaking French only, is required on the labour market, the new immigrants may be less employable. Moreover, a drop in points on "adaptability" may hurt integration if these subjective assessments were effectively screening for unobservables.

Finally, immigrants with a foreign bachelor's degree and speaking French may not be highly substitutable for other workers and hence may be in a very narrow labour market. An increase in their number will inevitably lead to worse labour market outcomes, all else equal. Overall, it is unclear whether this reform will improve labour market outcomes. Our empirical analysis below will look at the impact of the reform on immigrants' education and language skills, as well as their labour market outcomes.

<sup>&</sup>lt;sup>23</sup> Plan Stratégique 2001-2004, Ministère des Relations avec les Citoyens et de l'Immigration.

### 3.3.2 Exogeneity of the Reform

A concern for our analysis is that the reform may have been endogenous to labour market conditions of the time. For example, it could be that QC's labour market was on a worse trend than in ROC, and that this prompted the QC government to enact this reform. If immigrants entering QC after 2001 are found to do worse than earlier immigrants, it might be due to these changing labour market conditions, not to the point system. In this case, one would be tempted to conclude that changing points in this manner was detrimental to immigrants, when in fact, the reform may have been beneficial.

In this particular case, the 2001 QC reform was not designed to counter a worsening economic trend. This reform was designed to achieve the primary objective of the new immigration policy decided in 2001: "increase the French-speaking immigration, while maintaining the socioeconomic requirements (which favor a rapid integration to the job market)"<sup>24</sup>. This new immigration policy was decided by the governing party, the Parti Québécois, whose two first priorities are independence from Canada and protection of the French language (programme du Parti Québécois, 2001). The control of immigration was seen as a key issue to realize these goals. Statistical analysis of the results from the 1995 QC independence referendum (50.6 percent said no to independence) revealed that 60 percent of French speakers voted for independence, while only 5 percent of non Francophones did (Drouilly 1996). In 1995, Jacques Parizeau, the Premier of Quebec, stated that the next referendum would be successful with only a few percentage more of French speakers (Cardinal 2005). The underlying motivations for the 2001 reform were thus more political and cultural than economic. In this sense, we consider this reform as exogenous from labour market conditions of the time.

## **3.4 Data**

We use the confidential microdata files of the 2006 Canadian Census to identify individuals who immigrated to Canada between 1999 and 2003<sup>25</sup>. The reform was implemented in September 2001.

<sup>&</sup>lt;sup>24</sup> In French: "hausser l'immigration francophone, tout en maintenant les exigences socio-économiques (qui favorisent l'intégration rapide au marché du travail)", Plan Stratégique 2001-2004, Ministère des Relations avec les Citoyens et de l'Immigration.

<sup>&</sup>lt;sup>25</sup> We focus on the period 1999-2003 since there were no changes to point system in ROC over this period. In 1997-1998, ROC shifted from using the Canadian Classification Dictionary of Occupations (CCDO) to National Occupation Classification (NOC) in order to describe the different occupations. In 2002, ROC introduced the
Immigrants who landed in QC after 2001 went through the new point system, and form our treatment group. Immigrants who landed in QC in 2000 (or in ROC) did not go through this new point system<sup>26</sup>.

The Census does not provide information on whether the immigrant went through the point system, or immigrated through the family reunion or refugee programs. As noted above, most of the immigrants to Canada are selected through the point system (more than 55 percent each year, CIC 2007). To further ensure that the majority of our sample immigrated through the point system, we restrict the sample in three ways. First, we restrict our sample to those with at least a high school degree, since the QC point system excludes others<sup>27</sup>. Second, we keep only those who can conduct a conversation in at least one of the official languages, since those who cannot are unlikely to go through the point system. Third, we restrict the sample to household heads<sup>28</sup>, since nonhousehold heads are likely to immigrate as a dependent, or through family reunification. In our triple differences (as will be explained in greater detail below), we will relax these three restrictions one by one to create control groups of immigrants who have not gone through the point system.

The 2006 census reports education and language abilities in 2006, not at the time of immigration. It is possible that education in 2006 differs from education at the time of immigration, if immigrants acquired additional education in Canada. Since we are interested in the effect of a change in the point system on immigrants' characteristics at the time of immigration, we further restrict our sample to those who earned their highest degree outside Canada and those aged between 25 and 45 at the time of immigration since individuals are most likely to have completed their formal schooling at that age. We present results with and without this restriction.

The Census does not specify the province of landing. It is possible that the province of residence observed in the 2006 census differs from the province of landing at immigration. However, in our

Immigration and Refugee Act (IRPA) which altered the point system. However, only 8 percent of the federal skilled workers who immigrated to ROC in 2003 were assessed under the IRPA (CIC 2010).

<sup>&</sup>lt;sup>26</sup> It is unclear whether immigrants who landed in 2001 went through the new point system, and we will thus analyze them separately.

<sup>&</sup>lt;sup>27</sup> In ROC, the point system does not automatically exclude applicants with less than a high school degree. It merely gives them zero points on the education category. However, getting zero points on education, and knowing at most one official language (only 2 percent of immigrants to ROC with less than a high school degree speak the two official languages of Canada, 70 percent speak only one language and the rest do not speak any of the official languages) leads to a grade less than the minimum passing grade in all possible scenarios. Thus, the restriction of the sample to individuals with less than a high school degree is valid for ROC.

<sup>&</sup>lt;sup>28</sup> In the data, we use the concept of "main household maintainer", defined in the census as the person that contributes the most towards shelter expenses.

sample, the proportion of immigrants who migrated in and out of QC is small. Only 3.2 percent of immigrants to QC migrated in or out of the province in the last year, and 13.7 percent in the last 5 years. In addition, Okonny-Myers (2010) uses the longitudinal Immigration Database (IMDB) to show that 90 percent of the skilled workers immigrating to QC in 2000 were still in QC in 2006.

After applying these restrictions, the sample consists of 35,327 observations. Table 2 reports descriptive statistics for our sample, in QC and ROC, before the reform (the years 1999 and 2000) and after the reform (the years 2002 and 2003). We exclude the year 2001 since the reform was only partially implemented in 2001. The proportion of immigrants with a high school degree dropped from 18 to 11 percent in QC. In ROC, the drop was from 13 to 11 percent over the same period. Thus, the difference-in-differences estimate is a 5 percentage point decrease. A naive t-test of this estimate, i.e., not accounting for serial correlation within provinces, provided in the last column of Table 2, shows a statistically significant difference. The rest of Table 2 shows the same analysis for other characteristics. After the reform, immigrants to QC are 4 percentage points more likely to hold a bachelor degree, 3 percentage points to speak French only<sup>29</sup>, 4 percentage points less likely to speak English only, and 9 percentage points less likely to be employed.

Figures A2.1 to A2.5 in Appendix 2 show the same data by year in graphs. Figure A2.1 shows the percentage of people with a university degree who immigrated to QC between 1999 and 2003. While there were no significant changes to Master, Ph.D. and medical degrees, the proportion of immigrants with a bachelor's degree increased in 2001 and even more so in 2002. Figure A2.2 shows the share of immigrants to QC and ROC holding a bachelor's degree. In 1999, 27 percent of QC immigrants held a bachelor's degree, compared to 34 percent in ROC. The gap widened in 2000, before closing down after 2001. In other words, QC was strongly catching up to ROC after 2001. Figure A2.3 shows that the proportion of QC immigrants speaking only French increased in 2001 and even more so in 2002. Figure A2.4 shows employment rates of immigrants across time. Despite immigrants being more educated and speaking more French, the official language of QC, there is no improvement in the employment rate of immigrants who landed after 2002. In fact, the employment rate seems to decrease. Figure A2.5 shows no discernible effect on earnings after 2001 in QC, since the gap between QC and ROC remains the same over time.

<sup>&</sup>lt;sup>29</sup> French only is a dichotomous variable equal to 1 if the immigrant is able to conduct a conversation in French but not in English.

In the next section, we present our formal difference-in-differences methodology that allows us to control for potential confounding factors, and address the issue of serial correlation within provinces.

### 3.5 Methodology

Our empirical analysis exploits the fact that the selection process changed in QC in 2001, while that of ROC did not. To determine the effects of the change in points on immigrants' skills and labour market outcomes, we perform the following difference-in-differences analysis:

$$y_i = PROV_i + YEAR_i + \gamma_0 QCx1999_i + \gamma_1 QCx2001_i + \gamma_2 QCx2002_i + \gamma_3 QCx2003_i + \theta X_i$$
$$+ u_i$$

where *i* corresponds to individual *i*.  $y_i$  is the variable of interest (education, language, employment, or earnings), measured in 2006. *PROV<sub>i</sub>* are a set of provincial fixed effects. For example, it includes  $QC_i$ , a dichotomous variable equal to 1 if individual *i* resides in QC, 0 otherwise. *YEAR<sub>i</sub>* are a set of dichotomous variables for the year of immigration of immigrant *i* between 1999 and 2003. For example, *1999<sub>i</sub>* is a dichotomous variable equal to 1 for an individual immigrating in 1999, 0 otherwise.

We further include all interactions of  $QC_i$  and year of immigration dummies, except for  $QC*2000_i$ , the reference period before the reform.  $QC*2001_i$  is an interaction term between  $QC_i$  and  $2001_i$ , and isolates in the data those immigrants who reside in QC and immigrated in 2001. The coefficient  $\gamma_1$  of  $QCx2001_i$  thus looks at the impact of the reform in 2001 when the reform was partially implemented (i.e., after September of that year).

The coefficients of interest are  $\gamma_2$  and  $\gamma_3$ . They measure the changes in characteristics of immigrants to QC after the reform, relative to the same changes in ROC. The strength of this difference-in-differences methodology is that it controls for any provincial and time fixed effects. In other words, the analysis controls for the fact that QC is systematically different from ROC (by comparing immigrants within QC before and after the reform), and for the fact that later cohorts have less time to integrate or face different macroeconomic conditions (by comparing the same cohort of immigrants in different places).

The remaining identification assumption is the common time effects assumption: for our results to be valid, QC and ROC must be on similar trends. In other words, in the absence of the reform, QC would have evolved the same way as ROC. We address this concern in three ways. First, we look at pre-reform trends visible in  $QCx1999_i$ . No changes were made to the point system in QC and ROC in 1999. We thus expect  $\gamma_0$  to be not significantly different form zero. Second, we include as control variables,  $X_i$ , provincial macroeconomic variables (provincial unemployment rate and provincial average earnings) to directly control for the fact that QC may have been on a different time trend than ROC. Third, we provide triple differences estimates, described in greater detail below.

In all the regressions, we cluster the standard errors by province, the level at which the reform took place (Moulton 1990), to deal with the within cluster serial correlation problem that might occur in the difference-in-differences estimation (Bertrand, Duflo and Mullainathan 2004). However, Cameron, Gelabach and Miller (2008) show that using clustering of that sort leads to over-rejection when the number of clusters is small. Using Monte Carlo simulations with 10 clusters (equal to the number of provinces used in our analysis) and different error structures and cluster sizes, they show that the OLS standard errors reject the null at a rate of 10.6 percent to 77 percent. Even after correcting for clustering, the cluster-robust standard errors reject the null 8.2 percent to 18.3 percent. We follow Cameron et al (2008), and use the wild cluster bootstrap. With 10 clusters, they show that this technique rejects the null at a rate of 4.5 percent to 6.4 percent, not significantly different from 5 percent. In our analysis, we use the 6-point weight distribution proposed by Webb (2014).

### **3.6 Results**

### 3.6.1 Immigrants' Characteristics

We start by considering the effect of the reform on education and French language. In Column (1) of Table 3, the dependent variable is a dichotomous variable equal to 1 if the highest degree attained is high school, 0 otherwise. We only report the coefficients of the interaction between  $QC_i$  and the year of immigration dummies.  $\gamma_0$ , the coefficient of  $QCx1999_i$ , shows that QC and ROC were on a similar trend before the reform.  $\gamma_1$ , the coefficient of  $QCx2001_i$ , shows no significant effect after a partial implementation of the reform in 2001.  $\gamma_2$  and  $\gamma_3$ , the coefficients of  $QCx2002_i$ 

and  $QCx2003_i$ , show a negative and significant effect of the reform: compared to immigrants who landed in 2000, the reference period, immigrants who landed in QC in 2002 are 6 percentage points less likely to be high school graduates. In 2003, the coefficient is smaller and still significant. This is expected considering the partial<sup>30</sup> implementation in ROC of the IRPA reform in 2003, which increased points for education from 23 to 33 percent.

Column (2) shows the results for immigrants with a bachelor's degree. Immigrants who landed in QC in 2002 are 6 percentage points more likely to hold a bachelor's degree<sup>31</sup>.

Concerning language skills, Column (3) of Table 3 shows that immigrants who landed in 2002 were 4 percentage points more likely to speak only French, and 5 percentage points less likely to speak only English. There is no change for immigrants who speak both French and English.

Throughout Table 3, all coefficients pertaining to the 1999 period are not significantly different from zero, confirming that QC and ROC were on similar trends before the reform. All coefficients pertaining to the 2001 period are either not significantly different from zero, or smaller than those for 2002. This is expected since the reform was implemented only after September 2001.

Overall, these results show that the point system significantly affected the composition of immigrants: immigrants were more educated and spoke more French following an increase in points for education and French. We next ask whether this translated into better labour market outcomes for immigrants.

#### **3.6.2 Labour Market Outcomes**

After studying the effect of the 2001 change in the point system on immigrants' education and language skills, we now turn to their labour market performance. Column (1) of Table 4 does not show a positive impact of the reform on employment. If anything, the effect is negative in 2002, although the coefficient is not significant for the year 2003. There is no effect on earnings, as can be seen by the insignificant coefficients in Column (2) of Table  $4^{32}$ .

As noted above, our sample only includes immigrants who earned their highest degree outside Canada, since one of our main questions of interest was whether a change in the point system

<sup>&</sup>lt;sup>30</sup> Only 8 percent of immigrants in 2003 came under the IRPA.

<sup>&</sup>lt;sup>31</sup> The full set of variables as well as the results for the other educational categories (diplomas and certificates below university, Master, Ph.D. and medical degrees) are presented in Table A3.1.

<sup>&</sup>lt;sup>32</sup> The full set of variables is reported in Table A3.2.

attracted more educated immigrants. This restriction might conceal effects of the point system change if the reform attracted immigrants who decide to pursue their education in Canada. Getting more education in Canada may be associated with better labour market outcomes, if Canadian education is of greater quality, or if Canadian employers discriminate less against immigrants with Canadian degrees. By selecting these immigrants out of our sample, we might underestimate the true effects of the reform.

To address this issue, we replicate our methodology after including in our sample the immigrants who decided to study in Canada (see in appendix 4, Tables A4.1 and A4.2). As was the case in the restricted sample, this unrestricted sample shows that immigrants after 2002 were more educated, spoke more French, and did not have better labour market outcomes. Therefore, our results are not sensitive to our focus on individuals who studied outside of Canada.

Throughout Table 4, the coefficients pertaining to  $QCx1999_i$  are not significant, indicating that QC and ROC were on similar trends before the reform. Nonetheless, it is possible that QC experienced a negative shock precisely in 2001, which would explain all the results. To address this issue, we provide in the next section evidence from a triple differences analysis.

### **3.7 Triple Differences**

An issue with the difference-in-differences methodology is that QC may have been on a declining trend relative to ROC. This may explain the stagnating economic performance of immigrants, and mask beneficial effects of the change in points. To address this concern, we consider additional control groups of immigrants to QC who did not go through the point system, as well as natives. In a sense, these groups provide a falsification exercise for our theory, because immigrants not assessed in the point system, or natives, should not be affected by the QC 2001 reform. If their labour market performance is affected, it must mean that other factors are at play, and confound the estimates of the reform.

We first consider immigrants with no secondary education (aged between 20 and 64 at time of immigration). Those individuals cannot immigrate through the point system since the point system excludes such individuals. We thus define  $Assessed_i$ , a dichotomous variable equal to 0 for those immigrants without a secondary education, 1 for our sample identified above. We then perform the following triple difference analysis:

$$y_{i} = PROV_{i} + YEAR_{i} + \gamma_{0}QCx1999_{i} + \gamma_{1}QCx2001_{i} + \gamma_{2}QCx2002_{i} + \gamma_{3}QCx2003_{i}$$
$$+ \gamma_{4}Assessed_{i} + \gamma_{5}QCxAssessed_{i}$$
$$+ \gamma_{6}1999xAssessed_{i} + \gamma_{7}QCx1999xAssessed_{i}$$
$$+ \gamma_{8}2001xAssessed_{i} + \gamma_{9}QCx2001xAssessed_{i}$$
$$+ \gamma_{10}2002xAssessed_{i} + \gamma_{11}QCx2002xAssessed_{i}$$
$$+ \gamma_{12}2003xAssessed_{i} + \gamma_{13}QCx2003xAssessed_{i}$$
$$+ \theta X_{i} + u_{i}$$

 $y_i$  are labour market outcomes. We cannot look at education as an outcome since the control group of immigrants has no education, by definition. Moreover, the point of the triple difference analysis is to address the issue of a potentially deteriorating labour market in QC, thus we focus on labour market outcomes.

The main coefficients of interest are  $\gamma_{11}$  and  $\gamma_{13}$ , the triple differences estimates. For the common time effects assumption to hold, and the validity of the falsification exercise, we need  $\gamma_2$  and  $\gamma_3$ , i.e., the impact of the reform on the immigrants not exposed to the reform, to be not significantly different from zero. In Table 5, we report these four coefficients, along with  $\gamma_7$  and  $\gamma_9$  to verify that the trends between the treatment and control groups were not different before the reform.  $X_i$  includes the provincial macroeconomic variables used above (provincial unemployment rate and the provincial average earnings).

Columns (1) and (2) of Table 5 present the findings for employment and earnings:  $\gamma_2$  and  $\gamma_3$ , the coefficients of  $QCx2002_i$  and  $QCx2003_i$ , are not significantly different from zero. This indicates that immigrants not going through the point system are not doing worse in QC after 2001. This falsification exercise confirms that it is reasonable to assume that QC and ROC would have been on similar trends had the reform not been implemented. The triple differences coefficients  $\gamma_{11}$  and  $\gamma_{13}$  of  $QCx2002xAssessed_i$  and  $QCx2003xAssessed_i$  are not significantly positive, as found above, confirming that the reform had no positive impact on labour market outcomes of immigrants.

In columns (3) and (4), we consider another falsification exercise by looking at another group of immigrants that did not go through the point system: immigrants aged between 45 and 64 at the time of immigration, who have at most a high school degree, and are not able to conduct a conversation in either French or English. This group cannot get the passing grade of the point

system. We replicate our triple differences methodology, and find that this group was not affected by the reform, as expected.

In columns (5) and (6), we look at spouses of immigrants who immigrated to Canada at least one year after the household head (20 to 64 years old at immigration). These spouses are most likely sponsored by the household head to immigrate under the family class. As they immigrated one year after the household head, they did not contribute to the points under the skilled worker program. These spouses are unlikely to have gone through the point system, and thus provide another falsification test. Their performance on the labour market should be unaffected by the 2001 QC reform. Columns (5) and (6) show that this is the case.

Lastly, we consider a control group consisting of natives. A major obstacle in using this control group is the unavailability of a year of immigration. Instead, we use their year of entry to the labour market as an equivalent<sup>33</sup>, and compare natives who entered the labour market before and after 2001 in QC and ROC. Since natives did not immigrate, the changes in the point system should not affect their labour market outcomes. This is confirmed in columns (7) and (8) of Table 5.

Additionally, in all four control groups,  $\gamma_7$  and  $\gamma_9$ , the coefficients for  $QCx1999xAssessed_i$  and  $QCx2001xAssessed_i$  are not significant, confirming that the treated and control groups were not on different trends before or during the implementation of the reform.

Overall, these triple differences results show that QC and ROC were on similar time paths for non-assessed immigrants, and the same is probably true for assessed immigrants. Moreover, these results confirm that the reform had no positive impacts on employment and earnings, despite immigrants being more educated and speaking more French.

### **3.8 Discussion**

Our results show that immigrants held more bachelor's degrees and spoke more French after the reform, but that this did not translate into better labour market outcomes. Our triple difference estimates show that these results are not driven by a differential trend in QC after the reform, since other categories of immigrants not assessed through the point system, as well as natives did not fare differently in QC after the reform.

<sup>&</sup>lt;sup>33</sup> proxied by age -- years of education -6.

An explanation for these findings is that the reform specifically focused on bachelor's degrees, and the returns to holding a foreign bachelor's degree are low. In Table 6, we show the results of a simple OLS regression of labour market outcomes on basic characteristics of immigrants to QC. When they land, employment rates for bachelor's degree holders are 12 percentage points less than for holders of diplomas and certificates below university, the omitted category<sup>34</sup>. This is in line with the existing literature which documents the difficulties of foreign bachelor degree holders to find a job (Godin and Pinsonneault, 2004; Renaud and Cayn, 2006).

Moreover, the returns to speaking French only are low. Table 6 shows that those speaking French only are 8 percentage points less employed than bilingual immigrants, all else equal. Immigrants with a foreign bachelor degree and speaking French only thus have difficulties integrating in the QC labour market. Their issues may have been compounded by the increase in their number after the reform: if those particular immigrants are not highly substitutable for other workers and hence are in a very narrow labour market, an increase in their number will lead to worse labour market outcomes, all else equal.

A final explanation may be that the drop in points for "adaptability" had negative labour market consequences. If these qualitative interviews were screening accurately for unobservables, then we would expect a drop in points for that category to be associated with worse labour market outcomes.

In any case, despite the absence of better labour market outcomes, these results are exactly in line with the intended goal of the reform. As explained above, the primary objective of the reform was to "increase the French-speaking immigration, while maintaining the socioeconomic requirements"<sup>35</sup>. We find more French-speaking immigrants and comparable labour market outcomes.

### **3.9** Conclusion

Following a reform in QC in 2001 giving more points for education and French language, we find that immigrants held more bachelor's degrees and spoke more French. Despite low returns to foreign bachelor's degrees and the French language on the QC job market, and an additional

<sup>&</sup>lt;sup>34</sup> The worse performance of bachelor's degree holders is specific to immigrants. Bachelor's degree holders who were born in QC fare better than high school or certificate holders (as can be seen in Table A5 in the appendix). <sup>35</sup> Plan Stratégieure 2001, 2004, Minister deg Paletieure aug las Citerums et de Warmigratieur

<sup>&</sup>lt;sup>35</sup> Plan Stratégique 2001-2004, Ministère des Relations avec les Citoyens et de l'Immigration.

decrease in points for adaptability, we find no deterioration in labour market outcomes of immigrants. Our paper is the first to use an intra-national change in points, which allows us to use a difference-in-differences analysis to study the effects of a change in the point system on immigrants' characteristics and labour market outcomes. Moreover, to address the fact that QC and ROC may be culturally and politically different, and thus on different trends, we use a triple differences analysis. This methodology uses additional control groups within QC, i.e., natives and categories of immigrants who could not pass the point system, but who immigrated through other programs.

From the point of view of the government, whose goal was to increase French-speaking immigrants with no worsening in labour market outcomes, the reform was an unambiguous success. One may wonder what the labour market effects would have been if more points had been assigned to characteristics fetching higher returns in the job market. For example, in the case of QC, advanced university degrees (Master, Ph.D., MD), and bilingual immigrants perform better than other immigrants. However, in 2001, a Master or Ph.D. holder earned only 3 percent more of the passing grade than a bachelor's degree holder. Speaking English only earned 10 percent, while speaking French earned 30 percent. An interesting avenue of future research would be to study the labour market effects of reforms assigning more points to these categories.

# Tables

	(	QC	ROC
	before 2001	after 2001	1999-2002
Education	25	32	23
Language			21
French	29	34	
English	10	11	
Adaptability			14
Personal Qualities	23	10	
Motivation	8	3	
Knowledge of QC	3	3	
Visit to QC	9	10	
Connection with a resident in QC	5	5	
Age	19	21	14
Training	8	8	26
Experience	17	19	11
Employment	23	25	14
Financial autonomy	1	2	
Occupational Demand			14
Demographic			14

# Table 1: Summary of the Point System in QC and ROC from 1999-2002(percentage of the passing grade)

	(1)	(2)	(3)	(4)	(5)
	Ç	QC	RO	DC	DID
	1999-2000	2002-2003	1999-2000	2002-2003	
High school	0.18	0.11	0.13	0.11	-0.05
	(0.39)	(0.31)	(0.31)	(0.32)	[-4.00]***
Diplomas & Certificates	0.28	0.28	0.20	0.20	0.001
	(0.45)	(0.45)	(0.40)	(0.40)	[0.04]
Bachelor	0.28	0.36	0.35	0.39	0.04
	(0.45)	(0.48)	(0.48)	(0.49)	[2.40]**
Master	0.13	0.14	0.19	0.19	0.01
	(0.33)	(0.35)	(0.40)	(0.39)	[1.08]
Ph.D.	0.05	0.03	0.04	0.03	-0.01
	(0.21)	(0.28)	(0.19)	(0.17)	[-1.25]
Medical Degree (MD)	0.01	0.01	0.02	0.02	0.001
	(0.11)	(0.11)	(0.13)	(0.13)	[0.11]
French Only	0.23	0.26	0.001	0.001	0.03
	(0.42)	(0.44)	(0.03)	(0.03)	[2.40]**
English Only	0.25	0.20	0.95	0.94	-0.04
	(0.43)	(0.40)	(0.22)	(0.24)	[-2.60]***
French & English	0.52	0.54	0.05	0.06	0.004
	(0.50)	(0.45)	(0.22)	(0.24)	[0.22]
Employment	0.83	0.72	0.88	0.85	-0.09
	(0.37)	(0.45)	(0.32)	(0.36)	[-6.01]***
Log Earnings	10.07	9.86	10.38	10.16	0.01
	(1.05)	(1.04)	(0.96)	(0.98)	[0.24]
Observations	1755	3033	11361	11034	27183

#### Table 2: Descriptive Statistics from 1999 to 2003

**Notes:** Mean (standard deviation in parentheses). The first two columns report the mean and standard deviation in QC for the periods 1999-2000 and 2002-2003, respectively. Columns 3 and 4 reports the statistics for ROC. Column 5 reports the difference-in-differences estimates; t-values are reported in brackets. \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1. High school, Dipl & Cert, Bachelor, Master, Ph.D. and MD are dichotomous variables equal to 1 if the highest degree earned is a high school, a diploma or certificate below bachelor, a Bachelor's degree, a Master's degree, a Ph.D. degree and a medical degree respectively. French only is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in French but not in English. English only is a dichotomous variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant can conduct a variable equal to 1 if the immigrant is employed. Log Earnings is the logarithmic transformation of the earnings of employed immigrants

Dependent	(1)	(2)	(3)	(4)	(5)
Variable	High School	Bachelor	French only	English only	Fr & Eng
QCx1999	-0.04	0.01	0.004	0.01	-0.02
	(0.29)	(0.36)	(0.29)	(0.24)	(0.26)
QCx2001	-0.05	0.03	0.03	-0.04	0.01
	(0.23)	(0.47)	(0.04)**	(0.50)	(0.59)
QCx2002	-0.06	0.06	0.04	-0.05	0.01
	(0.03)**	(0.02)**	(0.001)***	(0.02)**	(0.16)
QCx2003	-0.07	0.03	0.03	-0.03	-0.003
	(0.05)*	(0.05)*	(0.03)**	(0.04)**	(0.46)
Observations	35327	35327	35327	35327	35327

#### Table 3: Effect of the 2001 QC Reform on Immigrants' Education and Language

**Notes:** Wild cluster bootstrap p-values at the province level are reported in parentheses, \*\*\*p<0.01, \*\*p<0.05, \*p<0.1. All regressions include province fixed effects, year fixed effects, provincial unemployment rate and provincial average earnings. High School and Bachelor are dichotomous variables equal to 1 if the highest degree earned is a high school or a Bachelor's degree respectively. French only is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in French but not in English. English only is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in English but not in French. Fr & Eng is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in English but not in French. Fr & Eng is a dichotomous variable equal to 1 if the immigrant sto QC who landed in 1999. It provides a falsification exercise by looking at pre-reform trends. QC\*2001, QC\*2002 and QC\*2003 provide the impact of the reform in 2001, 2002 and 2003. QC\*2000 is the reference category.

	(1)	(2)
Dependent Variable	Employment	Log Earnings
QCx1999	0.02	-0.06
	(0.58)	(0.59)
QCx2001	-0.04	-0.01
	(0.54)	(0.45)
QCx2002	-0.06	-0.06
	(0.07)*	(0.16)
QCx2003	-0.07	-0.02
	(0.17)	(0.48)
Observations	35327	28386

### Table 4: Effect of the 2001 QC Reform on Immigrants' Labour Market Outcomes

Notes: Wild cluster bootstrap p-values at the province level are reported in parentheses, \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1. All regressions include province fixed effects, year fixed effects, provincial unemployment rate and provincial average earnings. Employment is a dichotomous variable equal to 1 if the immigrant is employed. Log Earnings is the logarithmic transformation of the earnings of employed immigrants. QCx1999 isolates immigrants to Quebec. QCx1999 isolates immigrants to QC who landed in 1999. It provides a falsification exercise by looking at pre-reform trends. QCx2001, QCx2002 and QCx2003 provide the impact of the reform in 2001, 2002 and 2003. QC\*2000 is the reference category

			Additional control	group for the triple d	ifferences estime	ation (Assessed=0		
	Immigrants ag	ged between 20	Immigrants a	aged between 45	Spouses age	ed between 20	Natives w	ho are
	and 64 at time	of immigration	and 64 at tim	e of immigration	and 64 at time	of immigration	household h	eads and
	with no secon	dary education	with at most a	high school degree	who immigra	ted at least one	entered the lab	our market
			no knowled	ge of Fr or Eng	year after the	household head	between 1999	and 2003
	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8) I og
Dependent Variable	Employment	Log Earnings	Employment	Log Earnings	Employment	Log Earnings	Employment	Earnings
QCx2002	0.004	-0.31	-0.004	0.81	0.03	-0.26	-0.01	-0.01
	(0.88)	(0.13)	(0.79)	(0.53)	(0.47)	(0.24)	(0.30)	(0.78)
QCx2003	-0.06	0.09	-0.01	0.15	-0.04	-0.20	-0.01	0.06
	(0.50)	(0.46)	(0.83)	(0.71)	(0.16)	(0.25)	(0.46)	(0.46)
QCx1999xAssessed	0.05	0.01	-0.16	-0.44	0.06	-0.04	0.02	-0.03
	(0.52)	(0.89)	(0.73)	(0.15)	(0.55)	(0.72)	(0.39)	(0.66)
QCx2001xAssessed	-0.09	0.07	-0.09	-0.67	-0.05	0.25	-0.05	-0.03
	(0.63)	(0.50)	(0.47)	(0.67)	(0.49)	(0.27)	(0.52)	(0.28)
QCx2002xAssessed	-0.07	0.26	-0.07	-0.87	-0.09	0.22	-0.07	-0.07
	(0.62)	(0.20)	(0.34)	(0.68)	(09.0)	(0.19)	(0.48)	(0.47)
QCx2003xAssessed	-0.01	-0.11	-0.07	-0.17	-0.03	0.19	-0.07	-0.09
	(0.67)	(0.35)	(0.50)	(0.70)	(0.23)	(0.32)	(0.16)	(0.43)
Observations	48272	34187	41067	29596	45101	33286	132628	114289
<b>Notes:</b> Wild cluster b provincial unemploym is the logarithmic trantifies the years 2002 and 2	ootstrap p-value ent rate and pro sformation of th 2003, respective	s at the province vincial average ea e earnings of emp ly. QCx1999xAs	level are reporte urnings. Employm ployed immigrant ssessed, QCx2001	d in parentheses. All tent is a dichotomous s. QCx2002 and QCx IxAssessed, QC2002	regressions incl variable equal to 2003 provide th xAssessed and	ude province fixe o 1 if the immigrar e impact of the re QCx2003xAssess	ed effects, year fix at is employed. Lo <sub>i</sub> form on the contro sed are the triple	ed effects, g Earnings I group in difference

**Table 5 : Triple Differences Estimation** 

75

estimates. The omitted category is QCx2000xAssessed.

	(1)	(2)
Dependent Variable	Employment	Log Earnings
High School	-0.11	0.01
	(0.06)*	(0.17)
Bachelor	-0.12	-0.14
	(0.04)***	(0.13)
Master	0.01	0.15
	(0.05)	(0.16)
Ph.D.	0.15	0.42
	(0.08)*	(0.33)
Medical Degree	0.01	0.08
	(0.07)	(0.19)
YSM	0.03	0.07
	(0.01)***	(0.02)***
High School x YSM	0.004	-0.06
	(0.02)	(0.04)
Bachelor x YSM	0.03	0.05
	(0.01)**	(0.03)
Master x YSM	0.004	0.03
	(0.01)	(0.04)
Ph.D. x YSM	-0.03	0.05
	(0.02)	(0.07)
MD x YSM	-0.03	0.05
	(0.02)	(0.07)
French Only	-0.01	0.0004
	(0.02)	(0.05)
French & English	0.08	0.20
	(0.02)***	(0.05)***
Observations	6092	4286

Table 6:	Returns to	o Edu	cation	for	Immigran	ts in	QC
							× -

Robust standard errors in parentheses, \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1. All regressions include age, age squared, gender marital status and presence of kids. High school, Bachelor, Master, Ph.D. and MD are dichotomous variables equal to 1 if the highest degree earned is a high school, a Bachelor's degree, a Master's degree, a Ph.D. degree and a medical degree respectively. The omitted category is dipl & cert a dichotomous variable equal to 1 if the highest degree is a diploma or certificate below bachelor. French only is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in French but not in English. French \& English is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in both French and English. The omitted category is English only, a dichotomous variable equal to 1 if the immigrant can conduct a conversation in French

# Figures



Figure 1: Canadian permanent residents by category from 1998 to 2003 Source: CIC facts and figures 2007

# Appendix

# **Appendix 1: Adaptability**

The "Ministere d'immigration et communautes culturelles" (MICC) specifies, in the "Guide des procedures d'immigration" rules upon which points should be granted to each category of the point system. The following is a detailed explanation of the adaptability criterion which is assessed through an interview.

- Personal Qualities: assessment of the candidate's:
  - Ability to prove his/her achievements and accomplishments during an interview
  - Knowledge of the difficulties of immigration project (financial, family or professional)
  - Understanding of the values of QC society
  - Signing the "Déclaration sur les valeurs communes de la société québécoise"
  - Intention to learn French if he/she does not already know it.
  - Motivation: The steps taken by the applicant to facilitate socioeconomic integration:
    - His/Her efforts to get a job in QC (e.g. applying for jobs)
    - His/Her efforts to improve language proficiency in English or French
    - His/Her efforts to obtain a license to practice if he/she intends to exercise a regulated profession in QC
    - Other personal approaches showing efforts for integration (searching for a place to live, a school for children, etc..).
- Knowledge of QC:
  - Knowledge of the labour market
  - Knowledge of the economic sector
  - Knowledge of the living conditions
- Visit to QC:
  - Visit to QC before applying to immigration. Points are awarded depending on the duration and purpose of the visit.
- Connection with a resident in QC:
  - The presence of a close family member holding the Canadian citizenship or permanent residency and residing in QC.



# **Appendix 2: Immigrants' Characteristics and Labour Market Outcomes from 1999 to 2003**

Figure A2.1: Immigrants to QC from 1999 to 2003 by education level



Figure A2.2: Percentage of immigrants with bachelor's degree in QC and ROC from 1999 to 2003



Figure A2.3: Percentage of immigrants with only French knowledge in QC and ROC from 1999 till 2003



Figure A2.4: Employment of immigrants to QC and ROC from 1999 to 2003



Figure A2.5: Log Earnings of immigrants to QC and ROC from 1999 to 2003

		Tal	<u>ble A3.1: Efi</u>	fect of the 200	01 QC Refor	<u>m on Immig</u>	rants' Educatio	on and Langu	age
	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8)	(6)
Dependent	High	Dipl &		к И	х И	k Y	х 2	English	n. T
Variable	School	Cert	Bachelor	Master	Ph.D.	MD	French only	only	Fr & Eng
QC	0.10	0.05	-0.07	-0.08	-0.02	0.02	0.22	-0.67	0.45
	$(0.01)^{**}$	(0.16)	(0.29)	(0.20)	(0.16)	(0.11)	$(0.000)^{***}$	$(0.000)^{***}$	$(0.002)^{***}$
1999	0.03	0.01	-0.02	-0.03	-0.004	0.01	-0.0001	0.01	-0.01
	$(0.06)^{*}$	(0.21)	(0.18)	(0.12)	$(0.02)^{**}$	$(0.03)^{**}$	(0.67)	$(0.08)^{*}$	$(0.08)^{*}$
2001	-0.01	-0.02	0.02	0.01	-0.003	0.01	0.001	0.004	-0.01
	(0.41)	(0.26)	$(0.06)^{*}$	(0.79)	(0.14)	$(0.06)^{*}$	(0.17)	(0.69)	(0.49)
2002	-0.01	-0.01	0.02	0.004	0.0004	-0.0004	-0.002	0.01	-0.01
	(0.14)	(0.62)	$(0.04)^{**}$	(0.95)	(0.92)	(0.93)	(0.13)	(0.11)	(0.15)
2003	-0.01	-0.003	0.03	-0.004	-0.01	-0.0001	-0.001	-0.01	0.01
	$(0.03)^{**}$	(0.93)	$(0.08)^{*}$	(0.75)	(0.20)	(0.78)	(0.44)	(0.17)	(0.17)
QCx1999	-0.04	-0.02	0.01	0.02	0.01	0.02	0.004	0.01	-0.02
	(0.29)	(0.57)	(0.36)	(0.20)	(0.37)	(0.48)	(0.29)	(0.24)	(0.26)
QCx2001	-0.05	0.03	0.03	-0.002	-0.01	-0.003	0.03	-0.04	0.01
	(0.23)	(0.29)	(0.47)	(0.82)	(0.41)	(0.30)	$(0.04)^{**}$	(0.50)	(0.59)
QCx2002	-0.06	-0.003	0.06	0.004	-0.01	0.01	0.04	-0.05	0.01
	$(0.03)^{**}$	(0.84)	$(0.02)^{**}$	(0.80)	(0.21)	(0.25)	$(0.001)^{***}$	$(0.02)^{**}$	(0.16)
QCx2003	-0.07	-0.002	0.03	0.02	-0.01	0.02	0.03	-0.03	-0.003
	$(0.05)^{*}$	(0.85)	$(0.05)^{*}$	(0.60)	(0.52)	(0.54)	$(0.03)^{**}$	$(0.04)^{**}$	(0.46)
Unemployment	0.01	0.01	0.001	-0.01	-0.01	0.01	0.002	-0.01	0.01
Rate	(0.19)	(0.79)	(0.92)	(0.52)	$(0.05)^{*}$	$(0.02)^{**}$	$(0.048)^{**}$	$(0.07)^{*}$	$(0.08)^{*}$
Average	5.72E-06	-5.03E-06	3.2E-07	-3.59E-06	-4.75E-06	6.25E-06	4.57E-07	-6.73E-07	2.16E-07
Earnings	$(0.06)^{*}$	(0.64)	(0.82)	(0.74)	(0.11)	(0.17)	(0.17)	(0.66)	(0.89)
Newfoundland	-0.04	-0.11	-0.29	0.07	0.14	0.24	-0.02	0.01	0.01
	$(0.01)^{**}$	(0.60)	$(0.07)^{*}$	(0.49)	(0.11)	$(0.02)^{**}$	$(0.03)^{**}$	(0.77)	(0.87)

# Appendix 3: Detailed Results of the Effects of the 2001 QC Reform

			1	N TICKI NICH	עוווווועע				
	(1)	(2)	(3)	(4)	(5)	(9)	(2)	(8)	(6)
Prince Edward	0.01	0.30	-0.20	-0.08	-0.02	-0.01	-0.01	0.13	-0.13
	(0.49)	(0.45)	(0.51)	(0.25)	(0.63)	(0.53)	(0.24)	(0.0)*	(0.12)
Nova Scotia	0.03	-0.03	-0.11	-0.01	0.07	0.04	-0.04	-0.004	0.01
	(0.13)	(0.82)	(0.22)	(0.72)	(0.35)	(0.11)	(0.43)	(0.76)	(0.73)
New Brunswick	0.03	-0.03	-0.17	-0.01	0.09	0.08	0.03	-0.06	0.03
	(0.10)	(0.80)	(0.22)	(0.67)	(0.30)	(0.12)	$(0.04)^{**}$	(0.15)	(0.34)
Manitoba	0.13	0.14	-0.12	-0.16	-0.04	0.04	0.01	-0.01	0.001
	$(0.05)^{*}$	(0.64)	(0.45)	(0.42)	$(0.07)^{*}$	(0.11)	(0.16)	(0.68)	(0.93)
Saskatchewan	0.09	0.001	-0.16	-0.08	0.05	0.09	0.01	-0.03	0.02
	$(0.08)^{*}$	(96.0)	(0.32)	(0.59)	(0.39)	(0.11)	(0.14)	(0.27)	(0.49)
Alberta	0.03	0.03	0.02	-0.07	-0.02	0.01	0.004	-0.02	0.01
	(0.21)	(0.83)	(0.40)	(0.39)	$(0.03)^{**}$	$(0.08)^{*}$	(0.20)	(0.19)	(0.27)
<b>3ritish Columbia</b>	0.03	-0.02	0.02	-0.03	-0.02	0.01	-0.001	0.03	-0.03
	(0.16)	(0.84)	(0.84)	(0.34)	(0.23)	(0.12)	(0.55)	(0.11)	(0.12)
Observations	35327	35327	35327	35327	35327	35327	35327	35327	35327

	(1)	(2)
Dependent Variable	Employment	Log Earnings
QC	-0.08	-0.13
	(0.27)	(0.11)
1999	-0.02	0.06
	(0.19)	(0.25)
2001	-0.02	-0.04
	(0.02)**	(0.08)*
2002	-0.04	-0.12
	(0.02)**	(0.07)*
2003	-0.06	-0.22
	(0.01)**	(0.02)*
QCx1999	0.02	-0.06
	(0.58)	(0.59)
QCx2001	-0.04	-0.01
	(0.54)	(0.45)
QCx2002	-0.06	-0.06
	(0.07)*	(0.16)
QCx2003	-0.07	-0.02
	(0.17)	(0.48)
Unemployment Rate	0.01	-0.04
	(0.12)	(0.39)
Average Earnings	1.68E-06	5.53E-06
	(0.70)	(0.82)
Newfoundland	-0.03	0.72
	(0.37)	(0.20)
Prince Edward	-0.12	-0.15
	(0.21)	(0.51)
Nova Scotia	-0.03	0.07
	(0.36)	(0.75)
New Brunswick	-0.06	0.51
	(0.23)	(0.37)
Manitoba	0.06	-0.14
	(0.33)	(0.72)
Saskatchewan	0.07	0.04
	(0.30)	(0.83)
Alberta	0.06	0.06
	(0.05)*	(0.59)
British Columbia	-0.04	-0.04
	(0.47)	(0.71)
Observations	35327	28386

 Table A3.2: Effect of the 2001 QC Reform on Immigrants' Labour Market Outcomes

**Notes:** All regressions are clustered at the province level. Wild cluster bootstrap p-values are reported in parentheses, \*\*\*p<0.01, \*\*p<0.05, \*p<0.1. All regressions include province fixed effects, year fixed effects, provincial unemployment rate and provincial average earnings. Employment is a dichotomous variable equal to 1 if the immigrant is employed. Log Earnings is the logarithmic transformation of the earnings of employed immigrants. QCx1999 isolates immigrants to Quebec who landed in 1999. It provides a falsification exercise by looking at pre-reform trends. QCx2001, QCx2002 and QCx2003 provide the impact of the reform in 2001, 2002 and 2003. QCx2000 is the reference category

Domondont	(1) U:ab		(3)	(4)	(2)	(9)	(7)	(8) Eaclish	(6) (6)
Variable	School	Cert	Bachelor	Master	Ph.D.	MD	only	only	Eng
( ) (									
QCx1999	-0.03	-0.04	0.03	0.03	0.001	0.02	-0.01	-0.01	0.02
	(0.28)	(0.63)	(0.22)	(0.19)	(0.28)	(0.57)	(0.15)	(0.15)	(0.11)
QCx2001	-0.03	-0.0004	0.03	0.01	-0.003	0.004	0.01	-0.03	0.02
	(0.28)	(0.93)	(0.36)	(0.31)	(0.51)	(0.21)	(0.29)	(0.29)	(0.50)
QCx2002	-0.04	-0.02	0.05	-0.01	0.01	0.01	0.03	-0.06	0.03
	$(0.02)^{**}$	(0.33)	$(0.03)^{**}$	(0.64)	(0.16)	(0.31)	$(0.02)^{**}$	$(0.02)^{**}$	$(0.08)^{*}$
QCx2003	-0.04	-0.04	0.03	0.02	0.002	0.02	0.01	-0.03	0.01
	$(0.04)^{**}$	(0.60)	$(0.08)^{*}$	(0.51)	(0.56)	(0.60)	$(0.05)^{*}$	$(0.05)^{*}$	(0.15)
Observations	41637	41637	41637	41637	41637	41637	41637	41637	41637
Notes: All reparentheses, effects, provir Mater, Ph.D. diploma or ce- degree respec conversation i can conduct a the immigrant who landed i QCx2002 and category.	gressions are ***p<0.01, ncial unempl and MD are rtificate belo tively. Fren tively. Fren tively. Iten conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio conversatio	e clustered **p<0.05, oyment rai dichotomo w bachelo ch only is t not in Engli n in Engli t conversa provides a provide the	l at the prov *p<0.1. A te and provii ous variable or, a Bachelc or, a Bachelc s' a dichoton glish. Englis sh but not ii tions in both a falsificatio impact of th	ince level II regression ncial avers s equal to or's degree nous varis h only is a 1 French. J 1 French a on exercis e reform in	. Wild clu ions inclu age earnin 1 if the h 3 Master able equa 1 dichotorr Fr & Eng nd Englis se by loo n 2001, 20	uster boot ide provij igs. High : ighest deg 's degree, l to 1 if 1 nous varia is a dicht h. QCx19 king at p )02 and 20	strap p-val nce fixed school, Dip gree earned a Ph.D. de the immigi the immigi ble equal to otomous va 999 isolates bre-reform 003. QCx20	ues are repo effects, yea l & Cert, Ba is a high so gree and a 1 ant can col o 1 if the imu riable equa timmigrants trends. QC 000 is the re	

Appendix 4: Estimation without Restricting the Sample to Immigrants who Studied Outside Canada

	(1)	(2)
Dependent Variable	Employment	Log Earnings
QCx1999	0.01	-0.03
	(0.60)	(0.53)
QCx2001	-0.05	-0.01
	(0.45)	(0.47)
QCx2002	-0.06	-0.01
	(0.06)**	(0.61)
QCx2003	-0.07	0.001
	(0.21)	(0.94)
Observations	41637	33337

#### Table A4.2: Effect of the 2001 QC Reform on Immigrants' Labour Market Outcomes

**Notes:** All regressions are clustered at the province level. Wild cluster bootstrap p-values are reported in parentheses, \*\*p<0.01, \*p<0.05, \*p<0.1. All regressions include province fixed effects, year fixed effects, provincial unemployment rate and provincial average earnings. Employment is a dichotomous variable equal to 1 if the immigrant is employed. Log Earnings is the logarithmic transformation of the earnings of employed immigrants. QCx1999 isolates immigrants to Quebec who landed in 1999. It provides a falsification exercise by looking at pre-reform trends QCx2001, QCx2002 and QCx2003 provide the impact of the reform in 2001, 2002 and 2003. QCx2000 is the reference category

# **Appendix 5: Returns to Education for Natives in QC**

	(4)	
	(1)	(2)
Dependent Variable	Employment	Log Earnings
Below High School	-0.22	-0.39
	(0.00)***	(0.00)***
High School	-0.05	-0.14
	(0.00)***	(0.00)***
Bachelor	0.03	0.36
	(0.00)***	(0.00)***
Master	0.05	0.45
	(0.00)***	(0.00)***
Ph.D.	0.08	0.65
	(0.00)***	(0.00)***
Medical Degree	0.05	0.67
	(0.00)***	(0.00)***
Allophones	-0.08	0.04
	(0.02)***	(0.07)
French Only	0.11	0.05
	(0.00)***	(0.01)***
Fr & Eng	0.13	0.15
	(0.00)***	(0.01)***
Observations	77498	61454

#### Table A5: Returns to Education for Natives in QC

**Notes:** Robust standard errors in parentheses, \*\*\* p<0.01, \*\* p<0.05, \*p<0.1. All regressions include age, age squared, gender, marital status and presence of kids are included. High school, Bachelor, Master, Ph.D. and MD are dichotomous variables equal to 1 if the highest degree earned is a high school, a Bachelor's degree, a Master's degree, a Ph.D. degree and a medical degree respectively. The omitted category is dipl & cert, a dichotomous variable equal to 1 if the highest degree is a diploma or below bachelor. French only is a dichotomous variable equal to 1 if the immigrant can conduct a conversation in French but not in English. Fr&Eng is a dichotomous variable equal to 1 if the immigrant can conduct conversations in both French and English. The omitted category is English only: a dichotomous variable equal to 1 if the immigrant can conduct a conversation in English to 1 if the immigrant can conduct a conversation in both French and English. The omitted category is English only: a dichotomous variable equal to 1 if the immigrant can conduct a conversation in English to 1 if the immigrant can conduct a conversation in English. The omitted category is English only: a dichotomous variable equal to 1 if the immigrant can conduct a conversation in English but not French

## **Connecting Text**

After considering the determinants and returns of human capital in the first two essays, the last essay explores a specific component of human capital. Non-cognitive skills are a relatively recent and understudied factor of human capital in the economics literature. This thesis has already considered non-cognitive skills in the first essay where we studied the effects of maternal care on parent-reported measures of non-cognitive outcomes for children aged 2-3 years. In the third essay, we consider other measures of non-cognitive skills and present experimental evidence on personality traits and risk and ambiguity preferences.

The link between personality traits and labour market outcomes is well established in the literature. The same is true for the link between individual preferences and labour market outcomes. However, the relationship between personality traits and individual preferences under uncertainty is still understudied. The last essay in this thesis fills this gap.

# **Chapter 4**

# **Experimental Evidence on Personality Traits and Preferences**

# **4.1 Introduction**

Characteristics intrinsic to individual personality play a fundamental role in economic decisionmaking in ways not captured by markets (e.g. prices and constraints). Heterogeneity in preferences such as risk and time preferences, considered by most economists to be exogenous, influence economic behaviours from human capital investment to labour market participation to portfolio investment decisions (Guiso and Paiella 2005; Eckel, Johnson and Montmarquette 2005). Similarly, labour economists have shown the importance of non-cognitive traits in determining educational attainment, wages and employment outcomes (Heckman, Strixrud and Urzua 2006; Muller and Plug 2006; Lindqvist and Vestman 2011). Among such non-cognitive traits, economists have been increasingly considering personality attributes, such as the Big Five<sup>36</sup> developed by psychologists (openness, conscientiousness, extraversion, agreeableness and neuroticism) to explain a host of different economic behaviours. Most recently, Prevoo and ter Weel (2015) show that conscientiousness in particular highly correlates with wages and other socio-economic outcomes. Nonetheless, the relationship between personality traits and economic preferences is still not well established (Borghans et al 2008; Prevoo and ter Weel (2015)). Since both preferences and non-cognitive traits influence the same economic outcomes, an understanding of the relationship between preferences and personality traits is critical yet understudied.

In this paper, we fill this gap by studying the relationship between personality traits and preferences towards uncertainty<sup>37</sup>. Observing both individuals' uncertainty attitudes and personality types in a unique setting and dataset is at best rare. In order to help fill this gap, we run an experiment to check whether the Big-Five personality traits are related to risk and ambiguity

<sup>&</sup>lt;sup>36</sup> The Big Five personality traits are detailed in Appendix 1.

<sup>&</sup>lt;sup>37</sup> While decomposing how much preferences versus personality traits affect economic outcomes is interesting, it is outside the scope of our paper.

preferences. In the experiment, subjects make choices that disclose their risk and ambiguity preferences. Afterwards, they participate in an unstructured discussion. The discussion (chat) is internet based and done in small groups of three individuals. The subjects are given the chance to reconsider their earlier risk and ambiguity choices after the chat. Through content analysis, the chats allow us to measure subjects' personality traits. The particular setting of the experiment gives us the opportunity to study the relationship between non-cognitive traits and preferences before and after the chat. The second stage of the experiment is particularly interesting, since it allows us to determine which individual personality traits are more likely associated with a change in behaviour and whether the attributes of the other group members affect the individual's behavior after the chat. To our knowledge, this is the first paper that looks at the effect of others' personality traits on the individual's decision-making under uncertainty.

We find that personality traits are correlated with risk and ambiguity preferences before the chat. Specifically, conscientiousness is negatively related to risk and ambiguity aversion and agreeableness is negatively related to ambiguity aversion. We also show that the probability of changing decisions after the chat is affected by individuals' non-cognitive traits but not by the traits of the other group members. The latter do affect the direction of the change, albeit asymmetrically. Other chat group members' traits only affect the probability of becoming more risk or ambiguity averse but do not have an effect on the probability of becoming less risk or ambiguity averse. Specifically, openness to experience of the other group members plays a role in individuals becoming more risk averse after the chat as well as becoming more ambiguity averse, and agreeableness of the other group members affect the probability of becoming more ambiguity averse.

Our results complement other studies that find a relation between the Big-Five factor model and risk preferences (Becker et al. 2012; Nicholson et al. 2006). However, the evidence is mixed on which traits affect risk preferences and in which direction. Regarding ambiguity preferences, only two papers study the effect of the Big-Five personality traits. Borghans et al (2009) find no effect of personality traits whereas Rustichini et al (2012) show that extraversion is negatively related to ambiguity aversion.

Our results have important implications for both the economics and human resources disciplines. In economics, the correlation between personality traits and risk and ambiguity preferences is crucial for screening and contract designs. Specifically, employers and insurance

companies are interested in preferences; however, these are harder to observe during interviews than personality attributes. The fact that the personality traits of other group members affect the preferences of individuals might be of interest in the human resources domain, especially with the growing interest in team work.

The paper proceeds as follows: in Section 4.2, we present a selective literature review. Sections 4.3 and 4.4 describe the experimental design and procedure, respectively. In Section 4.5, we provide descriptive statistics of our sample and measures of preferences and personality traits. Section 4.6 presents the results and Section 4.7 concludes.

### **4.2 Literature Review**

Our paper contributes to the small but growing literature linking personality traits (e.g. Big Five) to individual preference parameters linked to decision-making under uncertainty. The majority of this literature considers only risk preferences, and the results tend to be mixed. Lauriola and Levin (2001) find that different Big-Five traits affect risk preferences differently depending on the domain. Specifically, they find that openness is positively but neuroticism negatively associated with risk taking in a gains domain among a sample of 76 Italian individuals. However, neuroticism is associated with more risk taking among the sample in the loss domain.

Becker et al. (2012) use both experimental and survey data from Germany to examine this relationship. General conclusions from their study show that risk taking behaviour is positively associated with openness and extraversion and negatively associated with conscientiousness, agreeableness and neuroticism. Nevertheless, this association is weak as the correlations are low in the three data sets they consider<sup>38</sup>.

Ertac and Gurdal (2012) extend the analysis to a group setting by examining the effect of personality traits on the willingness to take risky decisions on behalf of a group. They find that traits can have differential effects on preferences depending on the gender composition of the group. For instance, they find that higher levels of conscientiousness make men less cautious but women more cautious.

Very few papers extend this analysis further into the realm of decision-making under uncertainty so as to consider the relationship between personality traits and ambiguity preferences and the

<sup>&</sup>lt;sup>38</sup> These results are similar to Nicholson et al. (2006) and Dohman (2010).

existing results are mixed<sup>39</sup>. Borghans et al. (2009) find results on risk that are consistent with Becker et al. (2012) but no statistically significant relationship between the Big Five and ambiguity aversion. Rustichini et al. (2012) meanwhile do find a negative relationship between extraversion and ambiguity aversion.

In summary, the results from the existing literature point to more risk taking behaviours among individuals who score higher on openness and extraversion and less so for the remaining three traits. How these relate to ambiguity preferences is largely understudied. Our paper contributes to this literature by considering both risk and ambiguity preferences.

### **4.3 Experimental Design**

### 4.3.1 Overview

The experimental design is similar to the one used in Engle, Engle-Warnick and Laszlo (2011). It consists of three stages. In the first stage, we use risk and ambiguity instruments to reveal the subjects' attitudes towards uncertainty. The distinction between risk and ambiguity is driven by knowledge of the probability distribution of the outcomes. Risk pertains to situations in which the realization of the outcome is unknown and the probability distribution of the outcome and their probability distribution function are unknown. The design of our instruments is described in Engle, Engle-Warnick and Laszlo (2011) and is an experimental application of the ambiguity model in Klibanoff, Marinacci and Mukerji (2005). More explicitly, we can consider the ambiguity model to be approximated by a model of compound risk (Halevy 2007; Abdellaoui, Klibanoff and Placido 2015)<sup>40</sup>.

In the second stage, subjects participate in an unstructured online chat. This online chat is intended to provide a motive for social exchange. Engle, Engle-Warnick and Laszlo (2011) analyse the role that social learning plays in this exchange, and measure the effect of this social learning on measured attitudes towards risk and ambiguity. In this paper, we aim to identify the role that

<sup>&</sup>lt;sup>39</sup> A simple distinction that we use here between risk and ambiguity is that in both cases while the realization of future outcomes is uncertain, the probability distribution function of outcomes is considered known under standard risk situations but unknown in ambiguous situations. For more discussion on these differences, please see Engle-Warnick, Escobal and Laszlo (2011) and Engle-Warnick and Laszlo (2013).

<sup>&</sup>lt;sup>40</sup> We do not reproduce the theoretical underpinnings of our instruments, as this is done in Engle, Engle-Warnick and Laszlo (2011) for the same experiment

individual personality traits play in measured risk and ambiguity preferences. We are able to identify personality traits by analysing the content of what the subject says in the chat, traits which can then be correlated against risk and ambiguity attitudes. The group discussion provides a unique opportunity for us to also correlate the subjects' risk and ambiguity preference with the personality traits of other subjects within the group. Specifically, since the last stage provides subjects the possibility to revise their decisions in the risk and ambiguity instruments, we can evaluate whether individual versus others' personality traits trigger a change in measured risk and ambiguity preferences.

#### 4.3.2 Stage 1

In stage 1, subjects responded to risk and ambiguity instruments consisting of 12 and 13 binary choices respectively. Figure 1 presents the risk instruments. The individual has to make 12 choices, one in each row. Each row consists of two gambles, and each gamble has two equally probable outcomes. For example, in row 1, subjects have to choose between the left and the right lotteries. The left gives them \$26 with certainty whereas the right lottery gives them the chance to win \$24 or \$29 with a probability of 0.5 each.

As we move down the rows, the expected utility as well as the variance of the right gamble increase. So a utility maximizer individual will switch at some point from the left row to the right one. This switch-over point enables us to calculate an interval estimate of his risk preferences. Roughly speaking, the later the switch-over, the more risk averse the individual is. We use the number of safe choices (left row choices) as a measure of the individual's risk aversion (Holt and Laury 2002)

For ambiguity preferences, we use the instruments presented in Figure 2 that consist of 13 binary choices. Each individual has to make 13 choices. Similar to the risk instrument, each row consists of two gambles. The gamble on the left is taken from the risk instrument and the one on the right presents the same outcomes but with unknown probability to the player. The left gamble costs the subject \$1. For example, in row 2 in Figure 2 the participant can, on the one hand, pay \$1 and play the left lottery that gives equal probability of winning \$24 or \$29. On the other hand, he/she can choose the right gamble at no cost, and get the chance to win \$24 or \$29 but at unknown probabilities. Therefore, for each one of the 13 instruments, the individual has to make a choice between two possible gambles. The first is a risky gamble where the realization of the outcomes

is uncertain but the outcomes' probabilities distribution is known (in our case 50-50). The individual has to pay \$1 if he/she chooses this gamble. The other gamble is similar to the first one in terms of the possible outcomes, however, both the outcomes' realization and their probability distribution are uncertain. We use the number of times the individual pays to avoid ambiguity as a measure of the individual's ambiguity aversion (Engle-Warnick, Escobal and Laszlo 2011).

#### 4.3.3 Stage 2

In Stage 2, after answering all the risk and ambiguity instruments, subjects participated in an online chat in groups of three. The participants were instructed that they will be participating in an unstructured internet based chat that lasts 10 to 15 minutes. During the chat, the participants have to type their messages into the computer and can view the chat history at all time. An example of a chat is presented in Appendix 2.

### 4.3.4 Stage 3

After completing the chat, respondents were instructed that they can revise the decisions they made in stage 1. They were handed in a new copy of the risk and ambiguity instruments along with their original copy and a different pen color to prevent cheating. The participants were notified that only one of stage 3 decisions will be paid. The probability of choosing a specific decision was constant.

### **4.4 Experimental Procedure**

For stage 1, an experimenter read the instructions out loud to the participants. Afterwards, participants were given enough time to individually make decisions regarding 12 risk and 13 ambiguity instruments. The decisions were made in a paper and pencil format. Subjects were instructed that they will only be paid for one of those 25 decisions.

For the payment, each individual blindly chose a chip from a bag containing 25 chips<sup>41</sup> to determine the gamble to be paid for. Then, the subject declared loudly his/her winning color: yellow or blue. For the risk instrument, the participant drew a chip from a bag containing 5 blue chips and 5 yellow chips. For the ambiguity instrument, the bag contained 10 chips, however; the

<sup>&</sup>lt;sup>41</sup> 12 chips representing the 12 risky gambles and 13 chips representing the ambiguous gambles.

individual did not know the exact number of blue or yellow chips. The number of yellow chips in this bag was determined beforehand using a uniform distribution. Each participant had the right to check the color of the chips in the ambiguity bag after the experiment, if he/she asked for.

After completing stage 1, the respondents were put randomly in groups of 3 to participate in an online chat. We used an open-source internet based chat software. The software contained two windows, a lower one where the chatter can enter his/her messages and an upper one where he/she can check the chat history. Individuals were instructed that the chat was unrestricted, so they can talk about anything they like except their identity and profanity. Each chat lasted 10 to 15 minutes. Once the chats were completed, an experimenter handed out new copies of the risk and ambiguity instruments used in stage 1. Subjects were asked to fill them the way they desire. They were given a pen of different color than the one used in stage 1 and were instructed that they will be paid for exactly one of the 25 gambles.

After finishing stage 3, subjects filled a socio-demographic survey and were paid in cash for their participation. 122 individuals participated in the experiment and earned on average \$40 along with the \$10 show up fee.

### 4.5 Data Description

#### 4.5.1 Descriptive Statistics

Our sample, described in Table 1, consists of 122 subjects. The average subject is 26.3 years old, equally likely to be male or female. Almost half our subjects are working at the time of the experiment, 31% attained undergraduate education and 65% attained graduate education (the omitted category consists of those who did not attain any university education). We asked subjects about previous experiment experience: 81% of our subjects had already participated in experiments but only 37% participated in lottery experiments before. We use the dwelling value in the area defined by the first three characters of the postal code as a proxy of wealth (the Forward Sortation Area - FSA). This data is collected from the 2006 census. The average dwelling value in our sample is \$188,842.

#### 4.5.2 Choices in the Risk and Ambiguity Instruments
Figure 3 presents a histogram of the number of safe choices made in the risk instruments. The number of safe choices an individual can make is on the horizontal axis. The range is from 0 to 12. The vertical axis shows the frequency or the number of subjects making each choice. The blue bars correspond to the choices made before the chat and the red ones to the choices made after the chat. The figure shows a bi-modal distribution of the number of safe choices with the first mode occurring on 0 and the second one around 5 or 6. Figure 4 reports the histogram for the number of times paid to avoid ambiguity. The distribution is roughly bi-modal in this case too with the first mode at 0 and the second at 8.

Since we are also interested in the change in the individual's behavior after the chat, we report in Figure 5 the difference in the number of safe choices the subject made before and after the chat. This allows us to study the change in the revealed risk preferences after the chat. The horizontal axis ranges from -12 to 12. A negative value corresponds to an individual becoming less risk averse. The figure shows that the majority of participants did not change the number of safe choices they made pre and post chat. However, there is a slightly larger number of subjects becoming less risk averse than those becoming more risk averse.

The difference in the number of times the individual paid to avoid ambiguity is reported in Figure 6. Here again, the number of individuals becoming less ambiguity averse is slightly larger than those becoming more ambiguity averse. Nevertheless, most subjects did not change their revealed ambiguity preferences.

#### 4.5.3 The Big-Five Personality Traits

The Big-Five personality factors consist of five dimensions: Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism (OCEAN). Costa and McCrae (1988, 1992a) argue that all personality attributes can be mapped to one or more of those facets. We extracted our subjects' personality traits by analyzing the chat content. Four expert coders were trained to analyze the chat content and identify statements that reveal each trait. These coders were all graduate students in economics, and they followed standard protocols for content analysis coding. They were trained by the researchers to understand the meaning and sub-characteristics of each personality trait<sup>42</sup>.

Specifically, Openness was identified by the tendency to be informed, creative, insightful, and curious; Conscientiousness by the tendency to be thorough, organized, controlled, dependable, and decisive; Extraversion by the tendency to be sociable, assertive and to have positive energy; Agreeableness by the tendency to be accepting, conforming, trusting and nurturing and Neuroticism by the tendency to be depressed, anxious, insecure, vulnerable and hostile. Each trait was assigned the value of 1 if all descriptors are available and none is violated, 0 if at least one of the trait's descriptors is violated and 0.5 if there is no violation of the trait descriptors but not all of them are present.

Figures 7 to 11 show the distribution of the five traits among our subjects. The distributions of all the traits, except for neuroticism, are spread over the whole range of possible outcomes. The distribution of neuroticism is mainly concentrated at the lower end of the range implying that most individuals were not coded as neurotic.

The chat length ranged from a minimum of 24 lines to a maximum of 139 lines, with an average of 68 lines. Although the chat was unstructured, most chats were about the risk and ambiguity instruments used in the experiment. The inter-coder pairwise correlations are reported in table 2 for each trait. Even though the correlations are not very high, the majority is significant except for a few correlations for the neuroticism trait. In the rest of the paper, we use the average rater.

### 4.6 Results

#### 4.6.1 Pre-Chat Analysis

To analyze the pre-chat behavioral choices, we run the following regression:

$$Y_i = X'_i \alpha_1 + \sum_{j=1}^5 \beta_j TRAIT_{ji} + \epsilon_i \tag{1}$$

<sup>&</sup>lt;sup>42</sup> Personality traits questionnaires are the commonly used way to measure personality traits. We used chat content analysis instead since we are interested in knowing how individuals would perceive the personality traits of a subject after a chat. In the future, we would like to study the correlation between our measure and personality traits measured using questionnaires.

where *i* corresponds to the individual,  $X'_i$  is a vector of the individual's demographic characteristics. We consider the Big-Five personality traits where *j* represents the trait itself:  $j=\{1,2,3,4,5\}=\{$ neuroticism, extraversion, agreeableness, openness, conscientiousness $\}$ . The dependent variable  $Y_i$  is the number of safe choices the individual makes in the case of risk and the number of times he/she pays to avoid ambiguity in the case of ambiguity. Since both variables are count variables, we use a Poisson model to estimate our regressions. We run these regressions for the average rater. All our regressions are estimated using session clustered standard errors.

Table 3 reports the results of these regressions. Most demographic variables are insignificant. However, age is negatively correlated with the pre-chat risk and individuals whose highest level of schooling is graduate are more ambiguity averse. Turning into the traits variables, conscientiousness is negatively related to the number of safe choices before chat. This could be explained by the proposition that more conscientious individuals put high emphasis on their achievements and thus are willing to take more risk in order to get to their goals. Lauriola and Levin (2001) suggest that individuals with low impulsivity (conscientiousness is negatively associated with impulsivity) study each and every option they have very deeply which would create a conflict of interest and an unpleasant feeling so they would resort to the riskless option or In they could behave according to a defense mechanism, such as rationalization, which can lead to risk-taking" (p.217). This result complements the findings of Preevo and ter Weel (2015) who show that conscientiousness explains many socio-economic outcomes such as wages, health and saving behavior. They state that "empirical knowledge is too limited to judge how conscientiousness relates to the concepts and parameters economists typically model to predict outcomes" (p.920). In the case of ambiguity, agreeableness is negative and significant. Agreeable individuals are honest, trustful and optimistic about the human nature which might provide them with more ability to cope with ambiguity.

#### 4.6.2 Post-Chat Analysis

Next, we study the post-chat decisions. We are interested in both the change in decision and its direction. For this reason, we consider three dependent variables. The first is a binary variable equal to 1 if the number of safe choices is different after the chat than the one before the chat. The second dependent variable is again a dummy equal to 1 if the difference in the number of safe choices before and after the chat is positive, i.e. if the individual becomes more risk averse. This

dummy is equal to 0 if there was no change in the number of safe choices. Note that in this case, we only consider individuals who either became more risk averse or did not change their risk preferences after the chat. The third dependent variable is a dummy equal to 1 if the individual becomes less risk averse after the chat, 0 if he does not change his behavior. The same applies to the ambiguity measures. Since all dependent variables are binary variables, we estimate our regressions using a probit model.

For each specification, we run two regressions. The first is similar to the one used to analyse the pre-chat decisions, equation (1). Equation (2) adds the traits of the other individuals in the group.

$$Y_{ig} = X'_{ig}\alpha_1 + \sum_{j=1}^5 \beta_j TRAIT_{ji} + \sum_{j=1}^5 \theta_j OTRAIT_{jig} + \epsilon_{ig}$$
(2)

where *g* corresponds to the chat group of individual *i*.  $OTRAIT_{jig}$  is the prevalence of each personality trait among the individuals other than individual *i* in group *g*. As above, *j* represents the trait itself:  $j=\{1,2,3,4,5\}=\{$ neuroticism, extraversion, agreeableness, openness, conscientiousness $\}$ . The prevalence is calculated as the sum of each traits among the individuals in the group other than individual *i* according to the average rater.

Table 4 reports the results for the change in risk and ambiguity decisions. Men are less likely than women to change their risk decisions after the chat. However, this result is not significant when we add the attributes of the other group members. Agreeable individuals are more likely to change the number of safe choices they choose. This result might be explained by the trusting and compliant nature of agreeable individuals. Interestingly, the attributes of the other group members do not affect the probability of changing the risk preferences after the chat.

Regarding ambiguity preferences, singles and individuals who participated in experiments before are more likely to change their ambiguity decisions. Neurotic individuals and those who are conscientious are less likely to change their decisions after the chat. Those results are significant at the 5% and 1% level respectively, even after controlling for the traits of the other members in the group. In the case of ambiguity, the chat can be considered a way to exchange valuable information which might decrease the level of ambiguity. However, neurotic individuals are anxious, tense and worrying, so it is unlikely that they will integrate any of the information they get in the chat into their decision which explains their lower probability of changing the ambiguity decisions after the chat. Conscientious individuals are very thorough and ordered. It is

unlikely that they will receive additional information to help them resolve the ambiguity and will thus be less likely to change their ambiguity decisions after the chat.

In summary, the individual's personality traits but not the attributes of the group members affect the probability of the individual's changing his/her behaviour regarding risk and ambiguity after the chat. However, this analysis cannot distinguish whether group members' personality traits cause changes in opposing directions that cancel each other out. We turn our attention then to the direction of this change i.e. whether the individual becomes more or less risk/ambiguity averse and whether the attributes of other group members affect the probability of changing preferences in a specific direction rather than the absolute change. Table 5 illustrates the results for risk preferences. The first two columns correspond to a positive difference in the number of safe choices, i.e. the individual becoming more risk averse. The other columns correspond to the negative difference, so that the post-chat number of safe choices is lower than the pre-chat one, meaning that the individual becomes less risk averse. Individuals who participated in experiments before are more likely to become more risk averse after the chat. Men are less likely to become less risk averse after the chat and thus more likely not to change their decision after the chat confirming our results in Table 4. Turning to personality traits, column 1 shows that agreeable individuals are more likely to become more risk averse after the chat, but, this result does not hold after adding the attributes of the other group members. However, agreeableness is positive and significant in columns (3) and (4) meaning that agreeable individuals are more likely to become less risk averse after the chat. So, the earlier significance that we got for agreeableness on the probability of changing risk decisions seems to emerge from the agreeable individuals becoming less risk averse. In addition, the higher the number other open individuals in the group, the more likely the individual will become more risk averse after the chat.

We report the results for ambiguity preferences in table 6. Single individuals are more likely to become more ambiguity averse after the chat, whereas working individuals, graduates, undergraduates and those who participated in an experiment before are more likely to become less ambiguity averse. As we have shown earlier, neurotic and conscientious individuals are less likely to change their ambiguity decisions. Here, we see specifically, that conscientious individuals are less likely to become more ambiguity averse, whereas neurotic individuals are less likely to become less ambiguity averse. This is in line with our earlier explanation that conscientious individuals are very thorough and comprehensive, so it is unlikely that the chat will provide them

with new information to change their ambiguity decisions; and neurotic individuals are pessimistic, anxious and insecure so it would be very unlikely that they will become less ambiguity averse even after the chat. Moving to the attributes of group members, the prevalence of neuroticism and agreeableness among the other group members are negatively related to the individual becoming more ambiguity averse, whereas the prevalence of openness to experience is positively related to the probability of becoming more ambiguity averse. None of the group members' attributes is significant in the case of becoming less ambiguity averse.

## **4.7 Conclusion**

Using a laboratory experiment, we study the effect of the Big-Five personality traits on risk and ambiguity preferences. We show that conscientiousness is negatively related to risk and ambiguity aversion and agreeableness is negatively related to ambiguity aversion. In addition, we study the effect of the personality traits of other group members on the individual's decision, through studying the change in the individual's risk and ambiguity decisions after participating in an unstructured group chat. We show that the attributes of other group members do not affect the probability of a change in the decisions per se, however, it affects the direction of the change. Particularly, the other member's attributes play a role in the probability of becoming more risk or ambiguity averse, but do not affect the probability of becoming less risk or ambiguity averse.

The relationship between personality traits and preferences provides additional support for integrating the non-cognitive traits into the economic discipline. Our results have several implications. On the academic side, this relationship might prove helpful in contract designs and screening strategies. In addition, the result that the attributes of other group members affect the individual's decision might be used in peer effects studies. On the policy side, early childhood interventions that attempt to influence non-cognitive skills might have an additional effect on future economic outcomes through their effect on risk and ambiguity preferences.

Our aim in this study was not to explain the effect of each specific trait on preferences, we were rather more interested in showing that a relationship exists between the two. Future research can attempt to explain and theorize the effect of the individual's as well as others' personality traits on preferences. In addition, in our paper, we measured personality traits through analyzing the chat content. Future studies can use personality traits questionnaire and check whether the relationships we find hold in other contexts.

# Tables

Table 1: Summary Statistic	CS
----------------------------	----

Variable	Mean	Std.Dev.
Age	26.31	6.53
Gender (Male = 1)	0.50	0.50
Employed	0.48	0.50
Highest level of schooling attained: Graduate	0.31	0.47
Highest level of schooling attained: Undergraduate	0.65	0.48
Participated in an experiment before	0.81	0.39
Participated in a lottery experiment before	0.37	0.49
Average value of dwelling in Forward Sortation Area (CDN\\$)	188, 841.6	75, 460.49
N. 100		

N = 122

	Rater A	Rater B	Rater C	Rater D
Neuroticism				
RaterA	1			
RaterB	0.1942	1		
RaterC	0.4474*	0.2018	1	
RaterD	0.2951*	0.1808	0.4957*	1
Extraversion				
RaterA	1			
RaterB	0.4261*	1		
RaterC	0.4430*	0.3717*	1	
RaterD	0.5125*	0.3026*	0.2995*	1
Extraversion				
RaterA	1			
RaterB	0.4164*	1		
RaterC	0.3730*	0.3588*	1	
RaterD	0.4221*	0.3267*	0.4026*	1
<b>Openess</b> to Experience	2			
RaterA	1			
RaterB	0.5112*	1		
RaterC	0.4529*	0.3881*	1	
RaterD	0.3211*	0.2127	0.4640*	1
Conscientiousness				
RaterA	1			
RaterB	0.4286*	1		
RaterC	0.4911*	0.4607*	1	
RaterD	0.4816*	0.5390*	0.6070*	1

## **Table 2: Intercoder Correlations**

Notes: \* implies significance at the 1% level. N= 122

	# of safe choices	# of times paid
Individual is neurotic- average rater	0.15	0.26
	(0.55)	(0.41)
Individual is extraverted- average rater	0.14	0.24
	(0.26)	(0.34)
Individual is agreeable- average rater	-0.05	0.48
	(0.24)	(0.17)***
Individual is open- average rater	0.21	-0.18
	(0.39)	(0.38)
Individual is conscientious- average rater	-0.60	-0.41
	(0.35)*	(0.23)*
Age	-0.04	-0.01
	(0.02)**	(0.01)
Gender (Male=1)	-0.23	-0.08
	(0.14)*	(0.07)
Marital status (Single = 1)	-0.08	-0.18
	(0.18)	(0.12)
Employed	0.17	0.08
	(0.13)	(0.13)
Highest education: Graduate	0.39	1.27
	(0.30)	(0.64)**
Highest education: undergrad	0.07	1.02
	(0.34)	(0.70)
Ever participated in experiment	0.18	0.17
	(0.12)	(0.21)
Ever participated in lottery experiment	-0.03	-0.13
	(0.19)	(0.11)
Log average FSA dwelling value	0.09	0.02
	(0.18)	(0.15)
F-test for individual big five	8.99*	53.36***

Table 3: Pre-chat Ch	naracteristics (Poiss	on Estimates)
----------------------	-----------------------	---------------

Notes: Standard errors clustered by session in parentheses. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1. N = 113

	change in # of safe choices		change in # of	times paid
	(1)	(2)	(3)	(4)
Individual is neurotic- avg. rater	-0.38	-0.65	-0.62	-0.75
-	(0.56)	(0.55)	(0.29)**	(0.30)**
Individual is extraverted- avg. rater	-0.01	-0.01	-0.12	-0.10
-	(0.32)	(0.25)	(0.36)	(0.36)
Individual is agreeable- avg. rater	0.51	0.52	0.40	0.38
	(0.21)**	(0.25)**	(0.30)	(0.37)
Individual is open- avg. rater	-0.51	-0.61	0.04	0.08
	(0.42)	(0.41)	(0.33)	(0.30)
Individual is conscientious- avg. rater	-0.05	0.12	-0.57	-0.62
	(0.26)	(0.28)	(0.21)***	(0.21)***
Other neurotic group members- avg. rater		0.36		-0.02
		(0.24)		(0.24)
Other extraverted group members- avg.		0.11		-0.08
Rater		(0.13)		(0.19)
Other agreeable group members- avg.		-0.13		-0.07
Rater		(0.20)		(0.18)
Other open group members- avg. rater		0.31		0.38
		(0.19)		(0.25)
Other conscientious group members- avg.		0.10		-0.18
Rater		(0.18)		(0.22)
Number of safe choices- Pre-chat	0.04	0.05		( )
	(0.01)***	(0.01)***		
Number of times paid- Pre-chat			0.03	0.02
*			(0.02)	(0.02)
Age	-0.01	-0.01	-0.01	-0.01
-	(0.01)	(0.01)	(0.01)	(0.01)
Gender (Male=1)	-0.12	-0.11	-0.07	-0.05
	(0.06)**	(0.07)	(0.12)	(0.12)
Marital status (Single = $1$ )	0.19	0.21	0.20	0.22
	(0.13)	(0.12)*	(0.11)*	(0.12)*
Employed	0.10	0.15	0.09	0.09
	(0.13)	(0.15)	(0.06)	(0.07)
Highest education: Graduate	-0.09	-0.34	0.32	0.27
	(0.30)	(0.25)	(0.34)	(0.37)
Highest education: undergraduate	-0.04	-0.32	0.28	0.20
	(0.21)	(0.25)	(0.35)	(0.39)
Ever participated in experiments	0.11	0.19	0.28	0.30
	(0.09)	(0.12)	(0.15)*	(0.15)**
Ever participated in lottery experiments	0.02	0.03	-0.17	-0.17
	(0.11)	(0.10)	(0.12)	(0.11)
Log average FSA dwelling value	-0.09	-0.16	0.11	0.09
	(0.19)	(0.22)	(0.15)	(0.15)
Pseudo R-squared	0.20	0.25	0.21	0.23
F-test individual Big Five	14.9***	10.6**	158.15***	382.41***
F-test group Big Five		6.59		6.85

Notes: Standard errors clustered by session in parentheses. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1. N=113

	Positive change		Negative change	
	$(1) \qquad (2)$		(3)	(4)
Individual is neurotic- avg rater	-0.44	-0.48	-0.16	-0.51
	(0.42)	(0.40)	(0.57)	(0.46)
Individual is extraverted- avg_rater	-0.11	-0.0001	0.06	-0.02
	(0.33)	(0.31)	(0.27)	(0.21)
Individual is agreeable- avg_rater	0.41	0.24	0.43	0.49
	(0 14)***	(0.15)	$(0\ 20)^{**}$	(0.24)**
Individual is open- avg_rater	-0.33	-0.27	-0.47	-0.51
	(0.25)	(0.24)	(0.33)	(0.36)
Individual is conscientious- avg rater	-0.31	-0.42	0.12	0.31
	(0.25)	$(0\ 21)^{**}$	(0.18)	(0.19)
Other neurotic group members- avg rater	(0.20)	0.10	(0.10)	0.28
		(0.19)		(0.29)
Other extraverted group members- avg rater		-0.03		0.18
		(0.18)		(0.15)
Other agreeable group members- avg rater		0.06		-0.17
		(0.11)		(0.18)
Other open group members- avg. rater		0.29		0.18
		(0.17)*		(0.20)
Other conscientious group members- avg. rater		0.01		0.10
		(0.09)		(0.20)
Number of safe choices- Pre-chat	0.01	0.01	0.05	0.05
	(0.02)	(0.01)	(0.01)***	(0.01)***
Age	-0.01	-0.01	-0.001	0.001
5	(0.01)	(0.01)	(0.01)	(0.001)
Gender (Male=1)	-0.09	-0.05	-0.09	-0.09
	(0.05)*	(0.06)	(0.04)**	(0.05)*
Marital status (Single $= 1$ )	0.13	0.05	0.14	0.16
	(0.11)	(0.09)	(0.10)	(0.10)
Employed	0.15	0.19	0.03	0.08
	(0.12)	(0.12)*	(0.11)	(0.13)
Highest education: Graduate	-0.12	-0.19	-0.02	-0.27
	(0.15)	(0.11)*	(0.40)	(0.30)
Highest education: undergraduate	-0.07	-0.18	0.04	-0.27
	(0.13)	(0.17)	(0.35)	(0.41)
Ever participated in experiment	0.18	0.15	0.03	0.13
	(0.06)***	(0.07)	(0.11)	(0.12)
Ever participated in lottery experiment	0.03	0.10	-0.004	0.004
	(0.08)	(0.09)	(0.13)	(0.14)
Log average FSA dwelling value	-0.15	-0.13	-0.004	-0.07
	(0.09)	(0.07)*	(0.19)	(0.23)
Observations	83	83	92	92
Pseudo R-squared	0.30	0.36	0.20	0.24
F-test individual Big Five	43.99***	149.69***	6.26	9.70*
F-test group Big Five		22.73***		3.08

Notes: Standard errors clustered by session in parentheses. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

	Positive change		Negative change	
	(1) $(2)$		(3)	(4)
Individual is neurotic- avg. rater	-0.38	-0.45	-0.64	-0.88
-	(0.23)	(0.37)	(0.36)*	(0.45)*
Individual is extraverted- avg. rater	-0.06	-0.12	-0.05	0.01
-	(0.29)	(0.34)	(0.31)	-0.27
Individual is agreeable- avg. rater	0.37	0.42	0.35	0.13
	(0.24)	(0.26)	(0.35)	(0.41)
Individual is open- avg. rater	0.08	0.19	-0.08	-0.02
1 0	(0.22)	(0.18)	(0.33)	(0.29)
Individual is conscientious- avg. rater	-0.64	-0.71	-0.31	-0.34
C C	(0.19)***	(0.14)***	(0.21)	(0.29)
Other neurotic group members- avg. rater		-0.61		0.19
		(0.23)***		(0.29)
Other extraverted group members- avg. rater		-0.51		0.20
		(0.16)***		(0.20)
Other agreeable group members- avg. rater		-0.18		0.04
		(0.16)		(0.18)
Other open group members- avg. rater		0.38		0.34
		(0.21)*		(0.23)
Other conscientious group members- avg. rater		-0.20		-0.18
		(0.18)		(0.18)
Number of times paid- Pre-chat	0.003	-0.004	0.04	0.03
A A A A A A A A A A A A A A A A A A A	(0.02)	(0.01)	(0.02)**	(0.02)
Age	-0.003	-0.01	-0.004	-0.001
6	(0.01)	(0.01)	(0.01)	(0.01)
Gender (Male=1)	-0.06	-0.08	-0.03	-0.01
	0.13)	(0.14)	(0.09)	(0.06)
Marital status (Single $= 1$ )	0.23	0.28	0.02	0.05
	(0.09)***	(0.08)***	(0.10)	(0.08)
Employed	0.001	-0.09	0.09	0.14
	(0.08)	(0.07)	(0.05)**	(0.07)**
Highest education: Graduate	0.16	0.18	0.99	0.99
C	(0.28)	(0.23)	(0.03)***	(0.002)***
Highest education: undergraduate	0.072	0.01	0.93	0.90
0	(0.22)	(0.24)	(0.18)***	(0.04)***
Ever participated in experiment	0.11	0.08	0.28	0.30
	(0.09)	(0.09)	(0.10)***	(0.08)***
Ever participated in lottery experiment	-0.05	-0.05	-0.18	-0.15
	(0.09)	(0.08)	(0.11)	(0.09)*
Log average FSA dwelling value	0.12	0.07	0.10	0.10
	(0.11)	(0.08)	(0.15)	(0.13)
Observations	80	80	90	90
Pseudo R-squared	0.25	0.34	0.27	0.33
F-test Individual Big Five	19.77***	22.4***	20.48***	19.99***
F-test group Big Five		21.8***		21.08***

Table 6: Direction of the Change in the Number of Times Paid (Probit Marginal Effects)

**Notes:** Standard errors clustered by session in parentheses. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

# Figures

Choose the Lottery Tou Trefer in Each Now of the Table						
	Left L	ottery			Right Lottery	
Decision	5 chances	5 chances In 10	Your Choice		5 chances	5 chances In 10
1	\$26	\$26	Left	Right	\$24	\$29
2	\$24	\$29	Left	Right	\$22	\$32
3	\$22	\$32	Left	Right	\$20	\$35
4	\$20	\$35	Left	Right	\$18	\$38
5	\$18	\$38	Left	Right	\$16	\$41
6	\$16	\$41	Left	Right	\$14	\$44
7	\$14	\$44	Left	Right	\$12	\$47
8	\$12	\$47	Left	Right	\$10	\$50
9	\$10	\$50	Left	Right	\$8	\$53
10	\$8	\$53	Left	Right	\$6	\$56
11	\$6	\$56	Left	Right	\$4	\$59
12	\$4	\$59	Left	Right	\$2	\$62

Decision Task #1 Choose the Lottery You Prefer In Each Row of the Table

	Left Lottery					Right Lottery		
Decision	5 chances In 10	5 chances In 10	Lottery Cost	Your Choice		? chances In 10	? chances In 10	Lottery Cost
1	\$26	\$26	\$1.00	Left	Right	\$26	\$26	\$0.00
2	\$24	\$29	\$1.00	Left	Right	\$24	\$29	\$0.00
3	\$22	\$32	\$1.00	Left	Right	\$22	\$32	\$0.00
4	\$20	\$35	\$1.00	Left	Right	\$20	\$35	\$0.00
5	\$18	\$38	\$1.00	Left	Right	\$18	\$38	\$0.00
6	\$16	\$41	\$1.00	Left	Right	\$16	\$41	\$0.00
7	\$14	\$44	\$1.00	Left	Right	\$14	\$44	\$0.00
8	\$12	\$47	\$1.00	Left	Right	\$12	\$47	\$0.00
9	\$10	\$50	\$1.00	Left	Right	\$10	\$50	\$0.00
10	\$8	\$53	\$1.00	Left	Right	\$8	\$53	\$0.00
11	\$6	\$56	\$1.00	Left	Right	\$6	\$56	\$0.00
12	\$4	\$59	\$1.00	Left	Right	\$4	\$59	\$0.00
13	\$2	\$62	\$1.00	Left	Right	\$2	\$62	\$0.00

Decision Task #2 Choose the Lottery You Prefer In Each Row of the Table

Figure 2: Ambiguity Preference Elicitation Instrument



Figure 3: Histogram of the number of safe choices



Figure 4: Histogram of the number of times paid to avoid ambiguity



Figure 5: Change in the number of safe choices before and after the chat



Figure 6: Change in the number of times paid to avoid ambiguity before and after the chat



Figure 7: Neuroticism Distribution



**Figure 8: Extraversion Distribution** 



Figure 9: Agreeableness Distribution



### Figure10: Openness Distribution



Figure 11: Conscientiousness Distribution

# **Appendix 1: The Big Five Personality Traits**

The Big-Five personality traits is a model listing five main dimensions that are sufficient to describe the human personality. Below, we list the main facets of each dimension as described in McCrae and John (1992)

- Neuroticism: its main facets are anxiety, hostility, depression, self-consciousness, impulsiveness, and vulnerability. Neurotic individuals are anxious, self-pitying, tense, touchy, unstable and worrying
- Extraversion: its main facets are warmth, gregariousness, assertiveness, activity, excitement seeking and positive emotions. Extravert individuals are active, assertive, energetic, enthusiastic, outgoing and talkative
- Agreeableness: its main facets are trust, straightforwardness, altruism, compliance, modesty and tender-mindedness. Agreeable individuals are appreciative, forgiving, generous, kind, sympathetic and trusting.
- Openness to experience: the main facets of this dimension are fantasy, aesthetics, feelings, actions, ideas, values. Open individuals are artistic, curious, imaginative and insightful.
- Conscientiousness: its main facets are competence, order, dutifulness, achievement striving, self-discipline and deliberation. Conscientious individuals are efficient, organized, reliable, thorough and responsible.

A	ppen	dix	2:	Exam	ple	Chat

Chat group	Line	Time	Lab ID	comment
2	1	5:40	labo06	hello!
2	2	5:45	labo04	hey
2	3	5:48	labo05	hey
2	4	6:16	labo04	so what did you choose for decision 1
2	5	6:45	labo06	I chose left choice
2	6	7:02	labo04	ok me too cuz safer
2	7	7:05	labo05	I choose right cause at the end we can earn more
2	8	7:06	labo06	ya
2	9	7:07	labo05	ahahah
2	10	7:29	labo05	sorry
2	11	7:52	labo04	well I chose left for the first one then right for all the others
2	12	7:52	labo04	cuz true that we can earn more
2	13	8:03	labo04	with 50% chance
2	14	8:03	labo06	like me
2	15	8:21	labo06	I did same choices
2	16	8:48	labo05	what teh difference between the first and the rest
2	17	8:53	labo05	?
2	18	9:17	labo04	for the first, it's sure that you get 26 if left
2	19	9:52	labo05	you are right sorry, I was checking the wrong page, ahahah
2	20	10:01	labo04	haha
2	21	10:33	labo04	OK, we go on to decision 2 now
2	22	10:37	labo05	OK
2	23	11:01	labo04	for that one, right for the first one for sure cuz at no cost
2	24	11:15	labo05	exact
			labo06	about decision task 2 for the first decisions I choose right
2	25	11:16	11 04	choices
2	26	11:55	labo04	OK for the rest, I chose randomly
2	27	11:55		I think we made the same choices
2	28	12:00	labo06	so smart!
2	29	12:10	labo05	I keep right up to answer 4
2	30	12:14	labo04	cause of the cost
2	31	12:20	labo05	me5
2	32	12:55	labo04	me and 6 have he same mentality
2	33	13:04	labo05	what a team
2	34	13:06	labo05	!
2	35	13:28	labo06	LoL

2	36	13:57	labo05	For the rest I would say fifty fifty
2	37	14:06	labo05	four in the left, four in the right
2	38	14:17	labo06	it's the best way
2	39	15:19	labo05	but the last one is definitely right because of the cost again
2	40	15:24	labo06	maybe your right!
2	41	15:35	labo04	not really cuz have to think about probability right
2	42	15:43	labo06	I understand!
2	43	15:56	labo05	yep, but you have no control over that
			labo04	there might be 1 chance out of 10 for right and 5 out of 10
2	44	15:57		for left
2	45	16:25	labo04	I wanted to keep my 50% chance so I chose left
2	46	16:29	labo06	:)
2	47	16:57	labo04	haha, I didn't know we could use smiley face
2	48	17:08	labo04	just the button send and scroll bar
2	49	17:09	labo06	so cool!
2	50	17:15	labo05	yes and there might be 9 chance out of 10
				for right and 5 out of 10 for left
2	51	17:21	labo05	you are soulmates
2	52	17:26	labo05	ahahah
2	53	17:31	labo06	ahahah
2	54	17:31	labo04	hahaha
2	55	17:55	labo05	:)
2	56	18:13	labo04	ya, so for decision 2, we should only choose randomly cuz
				we don't have control on anything
2	57	18:24	labo05	that's what I think
2	58	18:28	labo06	kind of
2	59	19:31	labo05	but choose 50% left and 50% right answers
2	60	19:58	labo04	ya, now that's what I did for the final decision
2	61	20:12	labo06	bye!
2	62	20:17	labo05	good chat
2	63	20:17	labo05	bye
2	64	20:20	labo04	1-5 right, then 6, 8, 10, 12 left
2	65	20:20	labo06	bye
2	66	21:00	labo05	yep

## Conclusion

The three essays of this thesis use natural and laboratory experiments to study human capital determinants, returns and components. We start by studying the determinants of children's human capital. Specifically, we study the effects of maternal care on child development. Then we consider adult's human capital formation and returns. For this aim, we study the effects of immigration policies on immigrants' stock of human capital and its returns on the labour market. Lastly, we contribute to the literature on non-cognitive skills by providing experimental evidence on personality traits and preferences. Non-cognitive skills are a recently incorporated component of human capital in the economics literature.

The first essay, **"The Impact of Maternal Care on Child Development: Evidence from Sibling Spillover Effects of a Parental Leave Expansion"** studies the effects of maternal care on the cognitive, non-cognitive and health outcomes of children aged 2-3 years old using a parental leave reform implemented in Canada on December 31, 2000. Empirically, the reform increased the time mothers spend with their newborns by 3 months without affecting their income net of taxes, benefits and child care costs (Baker and Milligan 2010). Using the National Longitudinal Survey of Children and Youth, we perform a difference-in-differences methodology to compare children aged 2-3 years old with siblings born before and after the reform relative to children of the same birth cohorts who did not have a sibling in the period surrounding the reform. We find an increase in the time treated children spend with their mothers by 4.5 hours per week. The increase in maternal care is crowding informal care rather than centre-based care. We find no effect on child development in the short-run or the medium run.

Future research should study long term effects of maternal care for this age group and consider older children as well. Most of the outcomes used in this essay were parent reported, future studies should consider replicating the analysis using professional assessment of child development. Lastly, these studies would greatly benefit from measures about the quality of maternal and non-maternal care.

In the second essay, "The Effects of a Change in the Point System on Immigration: Evidence from the 2001 Quebec Reform", we evaluate the point system using a reform implemented in Quebec in 2001. The reform increased the points allocated for the French language and education (specifically bachelor's degrees) and decreased those allocated for a subjective criterion "adaptability". The aim of the policy was to increase the number of French-speaking immigrants without deteriorating their labour market outcomes. Using the 2006 Canadian Census, we employ difference-in-differences and triple differences analysis and show that, in line with the policy objectives, immigrants to Quebec after the reform held more Bachelor's degrees and spoke more French. There was no deterioration in their labour market outcomes measured by employment and earnings.

This essay shows clearly that the point system does attract the intended pool of immigrants and can be used to shape immigrant workforce according to policy goals. Many avenues for future research arise. First, one can evaluate the long term labour market outcomes of the treated immigrants as it may take more than five years for immigrants to assimilate properly. Second, future studies should try to employ direct assessment of the language skills and assess the quality of immigrants' education. A third avenue is to study the effects of the change in the point system on natives' labour market outcomes. Lastly, future research should study the labour market effects of reforms assigning more points to categories fetching high returns such as holding a Master's degree or being bilingual.

The last essay, **"Experimental Evidence on Personality Traits and Preferences"**, contributes to the growing literature on non-cognitive skills. We conduct an experiment to fill the gap in the literature on the relationship between personality traits and preferences. In the experiment, we use lottery instruments to measure the subjects' risk and ambiguity preferences. Afterwards, the subjects participate in an unstructured anonymous chat in groups of three, the content of which is used to measure personality traits. Later, the individuals are given the chance to revise their lottery choices.

We find that personality traits are correlated with decisions before the chat. Specifically, conscientiousness is negatively related to risk and ambiguity aversion and agreeableness is negatively related to ambiguity aversion. In addition, individuals' personality traits affect the probability of changing one's decision after the chat, whereas the attributes of other group members only affect the direction of the change.

These results can be used in contract designs and screening strategies. In addition, this essay contributes to the literature on peer effects. Lastly, this essay has important policy implications as early childhood interventions that attempt to influence non-cognitive skills might have an additional effect on future economic outcomes through their effect on risk and ambiguity

preferences. Future research can attempt to formalize the relationship between personality traits and individual preferences and replicate the analysis using personality traits questionnaires.

In summary, in this thesis, we show that maternal care is not a significant determinant of the human capital of children aged 2-3 years. In addition, we find that the point system can be used to determine the human capital stock of immigrants. In our specific case, the returns to the additional education and language knowledge did not have positive returns for immigrants to Quebec. Lastly, we show through experimental evidence, the existence of a relationship between personality traits and individual preferences under uncertainty.

Future research in these areas of human capital should consider other age groups and study other components of parental involvement in child development. Specifically, understanding the quality of parental time is crucial. For immigrants, studying the effects of changes in other categories of the point system is interesting, especially the ones fetching high returns in the labour market. Lastly, the literature on non-cognitive skills is still recent. More research is needed to devise the right measures and formalize the relationship between personality traits and preferences.

# References

- Abbott, M. G., & Beach, C. M. (2011). *Do admission criteria and economic recessions affect immigrant earnings?*. Montreal: Institute for Research on Public Policy.
- Abdellaoui, M., Klibanoff, P., & Placido, L. (2015). Experiments on compound risk in relation to simple risk and ambiguity. *Management Science*, 61(6), 1306-1322
- Almond, D., & Currie, J.(2011). Human capital development before age five. *Handbook of labor economics*, 4, 1315-1486.
- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, 103(484).
- Antecol, H., Cobb-Clark, D.A., & Trejo, S.J. (2003). Immigration policy and the skills of immigrants to Australia, Canada, and the United States. *Journal of Human Resources* 38(1): 192-218.
- Australia Bureau of Statistics. (2013). *Migration, Australia, 2013-14*. Catalogue number 3412.0.
- Aydemir, A. (2011). Immigrant selection and short-term labour market outcomes by visa category. *Journal of Population Economics* 24(2): 451-475.
- Baker, M., Gruber, J., & Milligan, K. (2008). Universal child-care, Maternal Labor Supply, and Family Well-Being. *Journal of Political Economy*, 116(4), 709-745
- Baker, M., Gruber, J., & Milligan, K. (2015). *Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program* (No. w21571). National Bureau of Economic Research.
- Baker, M., & Benjamin, D. (1994). The performance of immigrants in the Canadian labor market. *Journal of labor economics*, 369-405.
- Baker, M., & Milligan, K. (2008). Maternal employment, breastfeeding, and health: Evidence from maternity leave mandates. *Journal of Health Economics*, 27(4), 871-887.
- Baker, M., & Milligan, K. (2008b). How Does Job-Protected Maternity Leave Affect Mothers' Employment?. *Journal of Labor Economics*, 26(4), 655-691.
- Baker, M., & Milligan, K. (2010). Evidence from parental leave expansions of the impact of maternal care on early child development. *Journal of Human Resources*, 45(1), 1-32.

- Baker, M., & Milligan, K. (2015). Parental leave and children's cognitive and behavioral development. *Journal of Population Economics*, 28(2), 373-391.
- Barrett, A. (1998). The effect of immigrant admission criteria on immigrant labour-market characteristics. *Population Research and Policy Review* 17(5): 439-456.
- Baum, C. L. (2003). Does early maternal employment harm child development? An analysis of the potential benefits of leave taking. *Journal of labor Economics*, 21(2), 409-448.
- Baum, C. L. (2004). The long-term effects of early and recent maternal employment on a child's academic achievement. *Journal of Family Issues*, 25(1), 29-60.
- Beach, C. M., Worswick, C., & Green, A.L. (2011). *Toward Improving Canada's Skilled Immigration Policy: An Evaluation Approach*. CD Howe Institute
- Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annual Reviews of Economics*, 4(1), 453-478
- Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics* and *Statistics*, 86(2), 586-601.
- Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program cost--benefit analysis using data from the age-40 follow up. *Journal of Human Resources*, 41(1), 162-190.
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 1165-1188.
- Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3), 491-507.
- Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119 (1), 249--75.
- Bettinger, E., Haegeland, T., & Rege, M. (2014). Home with mom: the effect of stay-athome parents on children's long-run educational outcomes. *Journal of Labor Economics*, 32(3), 443-467.
- Bratsberg, B., & Ragan Jr, J. F. (2002). The impact of host-country schooling on earnings: a study of male immigrants in the United States. *Journal of Human resources*, 63-105
- Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: a production function approach. *World development*, 32(1), 1-13.

- Borghans, L., Duckworth, A. L., Heckman, J. J., & ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 972-1059
- Borghans, L., Heckman, J. J., Golsteyn, B. H., & Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3), 649-658.
- Borjas, G. J. (1991). *Immigration and self-selection*. In Immigration, trade, and the labour market, 29-76. University of Chicago Press.
- Borjas, G. J. (1993). *Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and the United States.* In Small Differences That Matter: Labour Markets and Income Maintenance in Canada and the United States, 21-44. University of Chicago Press.
- Caille, J. P. (2001). Scolarisation à 2 ans et réussite de la carrière scolaire au début de l'école élémentaire. *Éducation et formations*, 7-18.
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414-427.
- Cardinal, M. (2005). *Breaking Point: Quebec, Canada, The 1995 Referendum*. Montreal: Bayard Canada Books.
- Charach, A., Lin, E., & To, T. (2010). Evaluating the hyperactivity/inattention subscale of the National Longitudinal Survey of Children and Youth. Health reports, 21(2), 45.
- Chiteji, N. (2010). Time-preference, non-cognitive skills and well-being across the life course: Do non-cognitive skills encourage healthy behavior?. *The American economic review*, *100*(2), 200.
- Citizenship and Immigration Canada (2007) Facts and Figures: Immigration Overview Permanent and Temporary Residents. <u>http://publications.gc.ca/collections/collection\_2009/cic/Ci1-8-2007E.pdf</u>. Accessed 18 March 2016.
- Citizenship and Immigration Canada (2010) Evaluation of the Federal Skilled Worker Program. <u>http://www.cic.gc.ca/english/pdf/research-stats/FSW2010.pdf</u>. Accessed 18 March 2016.
- Cobb-Clark, D. A. (2000). Do selection criteria make a difference?: visa category and the labour market status of immigrants to Australia. *Economic Record* 76 (232): 15-31.
- Cobb-Clark, D. A. (2003). Public policy and the labour market adjustment of new immigrants to Australia. *Journal of Population Economics* 16(4): 655-681.

- Costa Jr, P. T., & McCrae, R. R. (1988). From catalog to classification: Murray's needs and the five-factor model. *Journal of personality and social psychology*, 55(2), 258.
- Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual differences*, 13(6), 653-665.
- Dahl, G. B., & Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *The American Economic Review*, 102(5), 1927-1956.
- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological bulletin*, 131(4), 483.
- De Silva, A. (1997). Earnings of immigrant classes in the early 1980s in Canada: A reexamination. *Canadian Public Policy/Analyse de Politiques* 179-202.
- Dohmen, T., Falk, A., Huff man, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3), 1238–60.
- Drouilly, P. (1996). Le référendum du 30 octobre 1995 : une analyse des résultats. L'année politique au Québec. Presses de l'Université de Montréal.
- Eckel, C., Johnson, C., & Montmarquette, C. (2005). Saving decisions of the working poor: Short-and long-term horizons. *Emerald Group Publishing Limited*, 10, 219-260.
- Engle, J., Engle-Warnick, J., & Laszlo, S. (2011). *Social Exchange and Risk and Ambiguity Preferences*. CIRANO working paper. No. 2011s-46
- Engle-Warnick, J., J. Escobal & S. Laszlo. (2011). Ambiguity Aversion and Portfolio Choice in Small-Scale Peruvian Farming. *B.E. Journal of Economic Analysis and Policy*, 11(1):
- Engle-Warnick, J., & S. Laszlo. (2013). *Learning-by-doing in an Ambiguous Environment*. CIRANO working paper.
- Ertac, S., & Gurdal, M.Y. (2012). *Personality, Group Decision-Making and Leadership*. Koç University-TUSIAD Economic Research Forum Working Paper.
- Felfe, C., Nollenberger, N., & Rodríguez-Planas, N. (2015). Can't buy mommy's love? universal childcare and children's long-term cognitive development. *Journal of Population Economics*, 28(2), 393-422.
- Friedberg, R. M. (2000). You can't take it with you? Immigrant assimilation and the portability of human capital. *Journal of Labor Economics*, *18*(2), 221-251.

- Godin, J.F., & Pinsonneault G. (2004). *L'insertion en emploi des travailleurs admis au Québec en vertu de la grille de sélection de 1996*. Direction de la population et de la recherche, Relations avec les citoyens et immigration Québec.
- Grant, M. L. (1999). Evidence of new immigrant assimilation in Canada. *Canadian Journal of Economics*, 930-955.
- Green, A. G., & Green D. A. (1995). Canadian immigration policy: The effectiveness of the point system and other instruments. *Canadian Journal of Economics* 1006-1041.
- Gregg, P., Washbrook, E., Propper, C., & Burgess, S. (2005). The Effects of a Mother's Return to Work Decision on Child Development in the UK. *The Economic Journal*, 115(501), F48-F80.
- Grove, W. A., Hussey, A., & Jetter, M. (2011). The gender pay gap beyond human capital heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, *46*(4), 827-874.
- Guiso, L., & Paiella, M. (2004). *The role of risk aversion in predicting individual behaviours*. Centre for Economic Policy Research.
- Gupta, N. D., & Simonsen, M. (2010). Non-cognitive child outcomes and universal high quality child-care. *Journal of Public Economics*, 94(1), 30-43.
- Haeck, C., Lefebvre, P., & Merrigan, P. (2015). Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics*, *36*, 137-157.
- Hanratty, M., & Trzcinski, E. (2009). Who benefits from paid family leave? Impact of expansions in Canadian paid family leave on maternal employment and transfer income. *Journal of Population Economics*, 22(3), 693-711.
- Halevy, Y. (2007) Ellsberg revisited: An experimental study. *Econometrica*, 75(2): 503-536.
- Havnes, T., & Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 97-129.
- Heckman, J.J., Humphries, J.E. & Mader, N. (2011a). The GED. In: Hanushek, E.A., Machin, S., Wößmann, L. (Eds.), *Handbook of the Economics of Education* (3). North Holland, Elsevier, Amsterdam, pp. 423–484 (Chapter 9).
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, *19*(4), 451-464.

- Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. & Yavitz, A.Q. (2010a). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94 (1–2), 114–128.
- Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A. & Yavitz, A.Q. (2010b). Analyzing social experiments as implemented: A reexamination of the evidence from the High-Scope Perry Preschool Program. *Quantitative Economics* 1 (1), 1–46.
- Heckman, J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics* 24(3):411--482.
- Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4), 800-802.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644-1655.
- Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education review*, *21*(6), 589-598.
- Karoly, L. A., Greenwood, P. W., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., Sanders, M. & Chiesa, J. (1998). *Investing in Our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions*. Santa Monica, Calif.: Rand Corporation
- Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under uncertainty. *Econometrica*, 73, 1849-1892.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.
- Kostov, C. (2008). Canada-Quebec immigration agreements (1971--1991) and their impact on federalism. *American Review of Canadian Studies* 38(1) 91-103.
- Kottelenberg, M. (2014). *Universal child-care, Obesity, and the BMI Distribution*. Queen's Economics Department Working Paper.
- Kottelenberg, M. J., & Lehrer, S. F. (2012). *The Gender Effects of Universal Child Care in Canada: Much Ado about Boys?*. Department of Economics, Queen's University, Kingston (mimeo).
- Kottelenberg, M. J., & Lehrer, S. F. (2013). New evidence on the impacts of access to and attending universal child-care in Canada. *Canadian Public Policy*, 39(2), 263-286.

- Kottelenberg, M. J., & Lehrer, S. F. (2014). Do the Perils of Universal Childcare Depend on the Child's Age?. *CESifo Economic Studies*, ifu006.
- Kramer, M. S., Matush, L., Vanilovich, I., Platt, R. W., Bogdanovich, N., Sevkovskaya, Z., ... & Shapiro, S. (2007). Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *The American Journal of Clinical Nutrition*, 86(6), 1717-1721.
- Kramer, M. S., Aboud, F., Mironova, E., Vanilovich, I., Platt, R. W., Matush, L., ... & Shapiro, S. (2008). Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Archives of general psychiatry*, 65(5), 578-584
- Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences*, 31(2), 215-226.
- Lefebvre, P., & 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., & Verstraete, M. (2009). Dynamic labor supply effects of childcare subsidies: Evidence from a Canadian natural experiment on low-fee universal childcare. *Labor Economics*, 16(5), 490-502.
- Lindqvist, E., and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 101-128
- Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education review*, 26(1), 52-66.
- Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance?. Economics of Education review, 26(1), 33-51.
- Marshall K. (2008). Father's use of paid parental leave. Perspect labor Income, 20, 5-14
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2), 175-215.
- McWhinney, M. (1998). A selection criteria chronology, 1967-1997: Critical changes in *definition, the point system and priority processing*. Strategic Research and Review Branch, Citizenship and Immigration Canada.
- Miller, P. W. (1999). Immigration policy and immigrant quality: The Australian point system. *American Economic Review* 192-197.

- Milligan, K., & 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), 305-326.
- Milligan, K., & Stabile, M. (2011). Do Child Tax Benefits Affect the Well-being of Children? Evidence from Canadian Child Benefit Expansions. *American Economic Journal: Economic Policy*, 175-205.
- Ministere de l'Immigration et des Communautes Culturelles. (2000). *Plan Strategique 2001-2004*. QC. <u>http://www.midi.gouv.qc.ca/publications/fr/planification/PlanStrategique20012004-Integral.pdf</u>. Accessed March 18, 2016.
- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, 334-338
- Mueller, G., and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review*, 3-22.
- Nicholson, N., Soane, E., Fenton-O'Creevy, M., & Willman, P. (2005). Personality and domain-specific risk taking. *Journal of Risk Research*, 8(2), 157-176.
- NICHR Early Child Care Research Network. (2003). Does amount of time spent in child care predict socioemotional adjustment during the transition to kindergarten?. *Child Development* 74(4), 976-1005.
- NICHD Early Child Care Research Network (2004). Type of child care and children's development at 54 months. *Early Childhood Research Quarterly*, 19(2), 203–230.
- Okonny-Myers, I. (2010). *The interprovincial mobility of immigrants in Canada*. Citizenship and Immigration Canada, Citoyenneté et Immigration Canada, Research and Evaluation.
- O'Shea, E. (2009). *Missing the Point (s): The declining fortunes of Canada's economic immigration program.* Transatlantic Academy Paper Series: 19.
- Parti Quebecois. (2001). Programme du Parti Quebecois. <u>http://www.democratie-nouvelle.qc.ca/wp-content/uploads/2013/01/parti\_quebecois-programme\_2001.pdf</u>. Accessed March 18, 2016.
- Phipps, S. A. (2000). Maternity and parental benefits in Canada: Are there behavioral implications?. *Canadian Public Policy/Analyse de Politiques*, 415-436.

- Prevoo, T., and ter Weel, B. (2015). The importance of early conscientiousness for socioeconomic outcomes: evidence from the British Cohort Study. *Oxford Economic Papers*, 67(4), 918-948.
- Quinn, P. J., O'Callaghan, M., Williams, G. M., Najman, J. M., Andersen, M. J., & Bor, W. (2001). The effect of breastfeeding on child development at 5 years: a cohort study. *Journal of pediatrics and child health*, 37(5), 465-469.
- Ray, R. (2008). *A detailed look at parental leave policies in 21 OECD countries*. Washington, DC: Center for Economic and Policy Research.
- Renaud, J., & Cayn, T. (2006). Un emploi correspondant à ses compétences?: les travailleurs sélectionnés et l'accès à un emploi qualifié au Québec. Ministère de l'immigration et des communautés culturelles,
- Ruhm, C. (2004). Parental employment and child cognitive development. *Journal of Human Resources*, 39(1), 155--192.
- Ruhm, C. (2008). Maternal employment and adolescent development. *Labor Economics*, 15(5), 958--983.
- Rustichini, A., DeYoung, C. G., Anderson, J., & Burks, S. V. (2012). *Toward the integration of personality theory and decision theory in the explanation of economic and health behaviour*. Discussion Paper series, Forschungsinstitut zur Zukunft der Arbeit. No. 6750
- Schönberg, U., & Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics*, 32(3), 469-505.
- Statistics Canada (2003). *Microdata User Guide: National Longitudinal Survey of Children and Youth*. Technical report, September 2002, Statistics Canada, Ottawa. Last accessed: 18 March 2016.
- Sweetman, A., & Warman, C. (2012). *The structure of Canada's immigration system and Canadian labour market outcomes.* No. 1292. Queen's Economics Department Working Paper.
- Ter Weel, B. (2008). The noncognitive determinants of labor market and behavioral outcomes: Introduction to the symposium. *Journal of Human Resources*, *43*(4), 729-737.
- Thomas, D., & Strauss, J. (1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77(1), 159-185.
- Waldfogel, J. (1999). The impact of the family and medical leave act. *Journal of Policy Analysis and Management*, 18(2), 281-302.

- Waldfogel, J., Han, W. J., & Brooks-Gunn, J. (2002). The effects of early maternal employment on child cognitive development. *Demography*, 39(2), 369-392.
- Wanner, R. A. (2003). Entry class and the earnings attainment of immigrants to Canada, 1980-1995. *Canadian Public Policy/Analyse de Politiques*, 53-71.
- Webb, M. D. (2014). *Reworking wild bootstrap based inference for clustered errors*. No. 1315). Queen's Economics Department Working Paper.