UNIVERSITY OF CALGARY

Econometric Analysis in Health and Healthcare Dynamics

by

Mahmood Zarrabi Ghahvehchi

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

CALGARY, ALBERTA

SEPTEMBER, 2009

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled “Econometric Analysis in Health and Healthcare Dynamics” submitted by Mahmood Zarrabi Ghahvehchi in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Supervisor, Christopher Auld, Associate Professor
Department of Economics

Mingshan Lu, Associate Professor
Department of Economics

Daniel V Gordon

Daniel Gordon, Professor
Department of Economics

Christopher Bruce, Professor
Department of Economics

Jenny Godley

Jenny Godley, Assistant Professor
Department of Sociology

Jane C Ruseski

External Examiner, Jane Ruseski, Assistant Professor
University of Alberta, Department of Economics

Aug 14 2009

Date

Abstract

This thesis consists of three studies that apply econometric techniques to issues of health and health care. The first study examines the long-term effect of cigarette prices on the probability and intensity of smoking. Using linear probability model, OLS and multinomial logit regressions, I test the effect of cigarette prices one faced as a youth on smoking habits as an adult. I find cigarette prices can deter and defer a youth's propensity to initiate smoking or to transition to daily smoking. However, the long-term effect of cigarette prices is weak but detectable. The second study employs a treatment effect model to examine the causal effect of parental leave on child outcomes. Using an endogenous switching regression model and a natural experiment on parental leave in Canada, I distinguish the average treatment effect from the effect of treatment on treated. I find there is substantial heterogeneity in the effect of parental leave, in which the unobservable heterogeneity effect of parents' skills for caring for their child and returning to work accounts for the main effect of parental leave. Moreover, I find longer parental leave attributes to a higher child development and temperament scores. The effect is more pronounced among the parents who took a long parental leave, controlling for unobservable factors affecting parents' propensity to take a long parental leave. The last chapter examines Wagner's Law for explaining the increases in the share of health expenditure as a proportion of national income in Canada. I use recent advances in time series econometric techniques to test panel stationarity and cointegration between health expenditure and income for the period 1975 to 2006 in a panel of the ten Canadian provinces. I use two specification approaches: First, using a dynamic panel, I find health expenditure has income elasticity in the range 0.47 to 0.61 and is not a luxury good. Second, using an error correction model, I find the growth rate of health expenditure does not change in periods of economic strength and weakness. Both findings show Wagner's Law is not credible for explaining the increases in health expenditure over time.

Acknowledgment

I am very grateful to my supervisor, Dr. Christopher Auld, for his invaluable comments and suggestions on my PhD dissertation. My study would have never ended up without his contributions and his good judgments on econometric methods and health research. He encouraged me to learn more and dig deeply into the literature. I also am grateful for comments and supports from my co-supervisors, Dr. Mingshan Lu and Dr. Daniel Gordon. I thank Dr. Herbert Emery for his supports and ideas for improving the last chapter of the dissertation.

I am truly indebted to my family that have always supported and encouraged me to move forward, in particular my dad, Ali-Akbar, who did not live to see the blossom of his patience and ambitions. I am heartily grateful for ever lasting love and support from my wife, Mahnoush Rostami, who has sacrificed a lot for giving me an opportunity to continue my education.

Mahmood Zarrabi
August 2009

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

This thesis consists of three studies that apply micro- and macro-econometric analyses in issues of health and health care dynamics. The first study examines the long-term effect of cigarette prices on the probability and intensity of smoking. The second study employs a treatment effect model to examine the causal effect of parental leave on child outcomes. The last chapter examines Wagner's Law to explain the increases in health expenditure over time in Canada.

Smoking control policies are legitimate if they effectively reduce the smoking rate of the population. It is argued that if a smoking control policy prevents youths from smoking, it will reduce the smoking rate of the population in the long-term. In the first study, I test this argument using a sample of 95,000 individuals from the Canadian Community Health Survey. I examine the effect of cigarette prices one faces as a youth on the probability and intensity of smoking as an adult. I find cigarette prices in youth have detectable but weak effects on smoking habits in later life. Using a simulation method, I predict a counter factual 50% increase in cigarette prices in youth will reduce the proportion of the smokers aged 20 to 29 years old by roughly 2%.

Increases in the female labour force participation rate over the last several decades have substantially affected the family environment and consequently might have affected children's health and development outcomes. Literatures in psychology and sociology find a mixed effect of the mother working on children's outcomes, from an adverse effect to no effect, and even a positive effect during childhood. The controversial results probably have arisen because of complexity in the correlation between the mother working and children's outcomes. For the second study, I try to disentangle the effect of unobservable heterogeneity in parents' skills for balancing home tasks and work responsibilities from the causal effect of parental employment on children's outcomes. I estimate the effect of parental leave using a flexible functional form of an endogenous switching regression model with error terms that follow a trivariate t-student distribution. Exogenous variation in parental work status arises from a natural experiment driven by reform in the parental leave mandates in Canada on December 31, 2000. I examine the short and medium term effects of a parental leave between 7 to 12 months on a variety of children's outcomes, including health, temperament, behaviour, milestone achievements, cognitive development, literacy, breastfeeding, parenting and family functioning. Using data of the National Longitudinal Survey of Children and Youth for the period 1996 to 2005 in Canada, I distinguish the average treatment effect from the effect of treatment on treated. I find that a long parental leave does not have any contemporaneous effect, but some positive and a few negative effects appear in later lives of children when aged 2 to 5 years old. In particular, I find cognitive development, breastfeeding and temperament of the children improve with a longer parental leave, while aggressive behaviours, family functioning and hostile parenting scores are reduced with a longer parental leave. Moreover, I find the negative effects disappear in later lives of the children.

The increases in health expenditure have been at the core of an enormous number of studies. Wagner's Law predicts that health is a luxury good with an income elasticity greater than 1 and so it predicts the share of health expenditure as a percentage of national income will grow over time as the economy expands. This is a major policy concern, because a higher proportion of health expenditure is associated with higher cost contamination in a publicly financed healthcare system. In the last study, first I employ recent advances in time series econometrics to test for stationarity and cointegration between health expenditure and income. Second, I test Wagner's Law using a panel data from the ten Canadian provinces. Two identification strategies are adopted to test Wagner's Law. First, using a dynamic panel, I test for the long-term income elasticity of health expenditures by subcategories, including source of finance and use of funds. Second, using an error correction model, I test for asymmetry relationship between growth rates of income and health expenditures in periods of economic strength and weakness. I find total health expenditure has a long-term income elasticity that falls in the range 0.47 to 0.61, which varies with model specification. I find health is not a luxury good, but hospital and physician expenditures, which account for almost 50% of total health expenditure, have long-term income elasticities above or close to unity. Furthermore, the relationship between the growth rate of income and that of health expenditure does not change with periods of economic strength and weakness, and so I conclude Wagner's Law is not credible for explaining the increases in health expenditure in Canada.

Overall, the findings in this thesis stress that individual unobservable factors contribute significantly to health behaviours and outcomes, such as smoking habits and children's health and development, and the effect is heterogeneous across individuals. Moreover, I find that health is not a luxury good and demand for healthcare services more likely will increase over time corresponding to the long-term growth path of the economy rather than by short-term changes in economic growth.

Chapter 2

The long-term effect of cigarette price on smoking: Evidence from the CCHS

2.1 Introduction

The prevalence of smoking among youths and adults has been at the core of an enormous number of studies. Less effort, however, has been taken to examine the relationship between youth and adult smoking. Becker and Murphy (1988) argue that smoking addiction requires current smoking to be correlated to past smoking, and that the degree of addictiveness could vary from one person to another. On the other hand, the relationship between youth and adult smoking might be explained by the effect of an unobservable factor that persistently affects smoking behaviours over time. Assuming the effect of unobservable factor accounts for a substantial correlation between youth and adult smoking, one concludes that discouraging youths from smoking would result in youth smoking initiation to be postponed without a substantial effect on the smoking rate of the adult population. Such a conclusion has very important policy implications, as it stresses the effectiveness of youth smoking control policies. This study proposes to test the effect of cigarette prices one faces as a youth on smoking behaviours as an adult in order to examine the long-term effect of cigarette prices.

I am particularly interested in the effect of cigarette prices more than the effects of other smoking control policies, because among different policies that prevail in most developed countries, such as tobacco taxes, clean indoor air laws, restrictions on cigarette use for the youth, and health warnings, there is a universal agreement that cigarette tax (price) is the most effective policy to control smoking (Warner, Chaloupka *et al.* ,1995). Moreover, I turn my attention to the long-term effect of cigarette prices because the literature finds most smokers start smoking in their youths, high cigarette prices reduce smoking initiation among youths, and youth smoking is more price sensitive than adult smoking[1]. Given that, there is a common belief that preventing youths from smoking will reduce the smoking rate of the adult population. This paper contributes to the literature by testing this argument. To do this, I exploit the high variation in cigarette prices in Canada, to get robust estimates of effects of cigarette prices, and take into account all cigarette prices in the entire youth that contribute to a youth smoking decision.

Cigarette prices in Canada increased sharply during the period 1991 to 1994 and then dropped dramatically among the eastern provinces, which in result cigarette prices have become very disproportionate across the provinces and very volatile over time. On the other hand, Statistics Canada conducted the Canadian Community Health Survey (CCHS) ten years after the sharp changes in' cigarette prices for covering health information of Canadians. The high variation in cigarette prices and the CCHS survey create a sample of adolescents and adults who faced disproportionate cigarette prices as youths, which suits my estimations for verifying the long-term effect of cigarette prices on smoking.

Empirical studies concerning the contemporaneous effect of cigarette prices conclude that cigarette demand is inelastic, with estimated elasticities in the range -0.7 to -0.5 among youths and in the range -0.25 to 0.0 among adults (Evans *et al.*, 1999;

[1] Lewit, Coate, and Grossman (1981); Evans and Huang (1998); Harris and Chan (1999); Tauras and Chaloupka (1999); Cawley, Markowitz, and Tauras (2004)

Chaloupka and Warner, 2000). In this study, I examine the long-term cigarette price elasticity of adult smoking. I use data of a pooled sample of individuals from the CCHS cycles 2001, 2003, and 2005; and take into account the effect of the average cigarette prices at age 14, 14 to 16, or 12 to 18 years (henceforth called price 14, price 14-16 and price 12-18, respectively) that respondents faced during period 1979 to 2004 in their youths. Linear probability, OLS, and multinomial-logit regression models are employed to examine the long-term effect of cigarette prices on the probability of smoking, the smoking intensity, and the smoking type of respondents aged 19 to 40 years old.

I find cigarette prices in early life can deter and defer a youth's propensity to transition into daily smoking. However, the long-term effect of cigarette prices on the probability of smoking is weak, but detectable. That is, a 10% increase in the cigarette price 12-18 will reduce the probability of smoking in adulthood by roughly 1%. Moreover, I find cigarette prices in youth have an adverse effect on the smoking intensity of the smoker. Furthermore, I find the cigarette price 12-18 has more influence on smoking habits in later life than the price 14-16 and in turn than the point price 14.

2.2 Previous Studies

The long-term effects of policies encountered early in life are often of policy and academic interest; for instance, the effect of early initiation on adolescent smoking (Auld, 2005), the long-term effect of youth tobacco control (Glied, 2002), the long-term effect of minimum wage on onset labour market outcomes (Neumark and Nizalova, 2004), the long-term effect of legalizing abortion on the crime rate (Donohue and Levitt, 2001), and the long-term effect of youth unemployment (Moraz and Savage, 2001).

The major body of the literature on the effect of cigarette price on smoking is related to the contemporaneous effect of cigarette price on smoking initiation and consumption of cigarettes. Using retrospective data, Foster and Jones (1999) find higher cigarette taxes are associated with later initiation. Chaloupka and Wechsler (1997) predict that a 75% price increase in a pack of cigarettes would reduce the number of smokers aged 18 to 24 years by over 1.2 millions. Warner *et al.* (1995) conclude that an increase in cigarette price is the most effective policy to influence the smoking of groups of people for whom education has been less effective. They conclude that"… the value of increased taxation in discouraging children from becoming addicted to nicotine was potentially the most powerful argument supporting increased taxes" (p.386). Gruber and Koszegi (2000) show that if the decision to smoke is time-inconsistent, then prohibiting youth from smoking may make them better off in later life. In contrast, Douglas and Hariharan (1994) and Douglas (1998) find that current cigarette prices are uncorrelated with smoking initiation.

Some studies have tested the causal effect of past prices on current smoking. Auld (2005) uses a dynamic structural model to decompose the youth smoking pattern over time into the correlation of smoking over time (addiction) and an unobservable heterogeneous effect of smoking intensity. He considers smoking at age 14 as an endogenous treatment on subsequent smoking, and uses cigarette price at age 14 as an instrumental variable that affects smoking at that age, while it does not have a direct effect on smoking in early adolescence. He uses the Youth Smoking Survey of Canada and employs an endogenous switching binary response regression, and finds that smoking is highly addictive for all respondents, but for those who were observed initiated early it is less addictive. He concludes that a smoking control policy would reduce the smoking rate of the population if it could deter the smoking initiation of early initiators, but not by a large magnitude. Glied (2002) uses cross-cohort correlation to examine the effect of smoking at age 21 on smoking at age 30 and 40. She finds that for every 100 smokers at age 21 there will be 75 smokers at age 30 and 55 smokers at age 40. She also uses a longitudinal analysis to test the effect of cigarette tax at age 14 on overall smoking behaviours, quitting and initiation habits as an adult. She takes advantage of the characteristics of the National Longitudinal Survey of Youth (NLSY79) in the United States to observe the actual cigarette taxes individuals faced at age 14 and changes in their smoking behaviours over time. She finds that a cigarette tax at age 14 has substantial effect on contemporaneous smoking, but that the effect is attenuated by adulthood. She concludes that the difference in cigarette taxes the respondents faced at age 14 has no effect on their smoking behaviour by age 40. Laux (2000) argues that those who faced high taxes in their youths may be more reluctant to initiate smoking as adults, and adult smokers who faced high taxes as youths are likely to have begun smoking in later adolescent. In contrast, some researchers find youth tax policy has no effect after adolescent (Orphnides and Zerovs, 1995; Survanovic *et al.*, 1999; Gruber and Koszegi, 2000). Decicca, Knedel, and Mathios (2002) use a model of onset smoking and discrete-time hazard with state fixed effects, and find that tax has no effect on the onset smoking between eighth and twelfth grades.

2.3 Data

This paper uses a pooled sample of individuals from the Canadian Community Health Survey (CCHS) and employs price of cigarettes as a proxy for cigarette taxes in

Canada. Cigarettes prices in Canada were subjected to sharp changes in early 1990s as a result of large changes in the federal and provincial cigarette taxes at the time. Statistics Canada conducted the CCHS survey 10 years later than the increases in the taxes to collect information on health status of Canadian 12 years and older. The tax changes provided lots of variation in the cigarette prices faced by cohort of Canadian youths who were adolescent/adult in the CCHS cycles. I extract this advantage of the CCHS and the tax changes in Canada to test the long-term effect of cigarette taxes (prices) on smoking.

2.3.1 The Canadian Community Health Survey

The CCHS is a cross-sectional survey that has been conducted biannually by Statistics Canada since 2001. The CCHS collects information on health status, health care utilization, and health determinants for the Canadian population. It operates on a two-year collection cycle. The first year of the survey cycle employs a large sample designed to provide reliable estimates at health region levels[2]. The second year of the survey cycle is a smaller sample and is designed to provide provincial level results on focused health topics. The CCHS targets persons aged 12 years and older who are living in private dwelling. Three cycles of the data are available at this time, cycles 2001, 2003, and 2005, which are used in this study.

The key advantages of using CCHS data are that first, the CCHS sample sizes are very large, thereby enabling researchers to derive reliable estimates even from their sub-samples. For example, this study restricts the sample to individuals aged 19 to 40 years old and stratifies its estimated models by gender. Second, tax changes in 1991 to 1994

[2] There are 122 health regions across the ten provinces and 1 health region per territory in Canada in 2003. This paper excludes information on the territories.

provided lots of variation in the cigarette prices faced by a cohort of Canadian youths who were adolescent/adult in the CCHS cycles.

Smoking habits tested in this paper include experimenting with a whole cigarette for the first time, youth daily smoking, youth deterred smoking, adult daily smoking, adult occasional smoking, and the smoking intensity. The smoking status of a person is ascertained by a series of questions in the CCHS. A person is considered as a daily smoker if he smoked each day in the last month prior to the interview date. The number of cigarettes that a daily smoker smokes is used to measure his smoking intensity. The smoking type of a respondent, including never smoked, former smoker, or current smoker, is obtained as a discrete variable that takes on the value 0, 1, or 2, respectively. The CCHS categorizes a former smoker as a person who has smoked at least 100 cigarettes in his life or used to be an occasional/daily smoker, but reported that he is not smoking at the time of interview. A current smoker in the CCHS is a person who is identified as an occasional or daily smoker at the time of interview. A never-smoked person is one who fits into neither the current smoker nor the former smoker category.

Table 2.1 displays the descriptive statistics of the variables used in this paper. The ethnicity of a respondent is ascertained from a question addressing his country of birth. Immigration status of a person is derived based on his current age and year of landing in Canada. A dummy variable is assigned to those immigrants who landed in Canada in their youth, so that I can observe the cigarette prices they faced as youths. That is, if a model is set to estimate the effect of cigarette prices at ages 14 to 16 on the probability of smoking at age 40 in year 2005, the dummy variable is set to take on value 1 if the adult is an immigrant who landed in or prior to 1979. The variable takes on value 0 if the adult is not an immigrant or he is an immigrant who landed in 1982 or later. Otherwise, I set the value of the dummy variable as missing; the missing value indicates that the adult is an immigrant who landed in Canada in a year within the period 1980 to 1981 and so a part of the cigarette prices he faced at ages 14 to 16 is not observable to econometricians. A similar procedure has been done for all immigrants aged 19 to 40 years old in all three cycles.

2.3.2 Cigarette Prices and Taxes in Canada

Over the last two decades, tobacco in Canada has been subject to volatile taxes, both over time and across the provinces. Particularly, the federal government of Canada raised tobacco excise tax and duty in 1991, which as a result the price of a pack of 200 cigarettes rose from roughly \$35 in 1990 to almost \$45 to \$50 in 1991. Studlar (2002) notes that Canadian tobacco taxes from 1984 to 1991 quadrupled while American taxes increased by less than 50%; as a result taxes in Canada averaged about seven times the US level, and consequently, raised tobacco smuggling sailed from the United States into Canada. Cigarette prices in Canada remained high until 1994 when the federal government reduced tobacco excise tax and duty. This was followed by reductions in retail tobacco taxes among the eastern and central provinces, except for Newfoundland, to control cigarette smuggling. These led to almost \$14 to \$21 reduction in the price of a pack of 200 cigarettes in Ontario, Quebec, New Brunswick, Prince Edward Island, and Novo Scotia, while the price of cigarettes remain relatively high in the western provinces and in Newfoundland (Hamilton *et al.*, 1997). Figure 2.1 displays the changes in Canadian cigarette prices over time and across the provinces from 1979 to 2004. The increases in tobacco taxes from 1991 to 1994, which were higher than the average, and then the sharp rollback in the prices in 1994 among the eastern and central provinces created a bulge in tobacco price trend. The substantial variations in cigarette prices in the 1990s and the CCHS survey which collected information on health status of Canadian 10 years later tailor an experiment that suits my estimates to identify the long-term effect of cigarette prices on smoking. I use the cigarette prices at three age groups in youth to ensure that they recover the effects of all prices attributed to a youth's smoking decision:

1. the effect of cigarette price at age 14,
2. the effect of average cigarette prices at ages 14 to 6,
3. and the effect of average cigarette prices at ages 12 to 18.

This paper uses cigarette price as a proxy of cigarette taxes. Cigarette prices in the Canadian provinces are highly correlated with the taxes, as changes in cigarette taxes are mainly reflected in the price change. The cigarette taxes in Canada include federal excise duty, federal excise tax, provincial tobacco tax, wholesale/retail margin, provincial sales tax, and the federal Goods and Services Tax (GST). From which, for example, a calculation for the share of cigarette taxes in the final price of cigarettes compiled by the Non-Smoker's Rights Association (2003) shows that cigarette taxes account for 80, 76, 68, and 69% of the cigarette prices in British Columbia, Alberta, Ontario, and Quebec in 1997, respectively.

I use the consumer price index for tobacco, which is obtained from Statistics Canada (CANSIM), to measure changes in cigarette prices over time. Statistics Canada began to collect cigarette prices at the provincial level since 1979. I am able to observe cigarette prices in the three age groups, 14, 14 to16 and 12 to18 years, for all respondents if I know which provinces they resided in as youths. Thus, I use a sample of individuals aged 19 to 40 years old in the period 2001 to 2005. For instance, a 40 year old respondent in 2005 was 14 years old in 1979, so I can find the cigarette price he faced at age 14 if I know which province he resided in at that age.

To obtain cigarette prices in youth, I assume individuals in my sample did not move across the provinces since youth. At first glance, this assumption seems to be very strong, but two considerations persuade me to rely on it. First, although cigarette prices in Canada have substantially varied over time, there had not been much disparity in the prices across the provinces during the period 1979-1994 (Figure 2.1). Second, for the rest of the time period (1995 to 2004), a calculation from the National Population Health Survey in Canada shows roughly 10% of the respondents aged 12 to 40 years old in 1994 had moved across the provinces from 1994 to 2003. In addition, since cigarette prices in each province relative to the prices in nearby provinces are not very disproportionate, moving to a near by province, which is more likely to take place, is less problematic to my estimations than moving from the west to the east and vice versa.

Figure 2.1 consists of four panels displaying the real cigarette price index by province from 1979 to 2005, and the cigarette prices respondents faced at age 14, 14 to16 or 12 to 18 over the same period.

2.4 Econometric Methods

The conventional statistical techniques to test the effect of cigarette price on the prevalence of smoking are probit, logit and linear probability models. Some studies, however, have tried to distinguish between participation and consumption of smoking, and so they used a Heckman two-step model or a double-hurdle model (Jones, 1989a and 1989b; Blaylock and Blisard, 1992, and Garcia and Labeaga, 1996). In this study, I am testing the effect of cigarette prices in youth on youth smoking, including the probability of experimenting with a whole cigarette, initiating daily smoking, or deterred daily smoking, and also on adult smoking, including the probability of daily smoking, the smoking intensity, and the smoking type. I employ a linear probability, OLS or multinomial-logit model where the outcome is a binary, continuous or discrete variable, respectively.

2.4.1 Estimate Models

The causal effect of cigarette prices in youth on smoking behaviour in adulthood can be projected by a structural form model given as

$$Smoking_{youth,i} = \alpha + \beta_1 P_{youth,i} + Z_i\gamma_1 + X_{youth,i}\gamma_2 + u_i \tag{2.1}$$

$$Smoking_{adult,i} = \delta + \beta_2 P_{adult,i} + \varphi Smoking_{youth,i} + Z_i\eta_1 + X_{adult,i}\eta_2 + \varepsilon_i \tag{2.2}$$

where $Smoking_{youth,i}$ and $Smoking_{adult,i}$ denote the smoking status of person i as a youth and as an adult, respectively. P_{youth} denotes cigarette price respondent i faced as a youth, P_{adult} denotes current cigarette prices respondent i faces as an adult. Z is a vector of time-invariant explanatory variables, including ethnicity and immigration status that may affect the smoking habits of a respondent, X_{youth} and X_{adult} are time-varying explanatory variables affecting youth and adult smoking, respectively, such as marital status, age, family income, education, province of residence, household size, depression, and pregnancy (for females only). α and δ are intercept terms of the youth and adult smoking models, respectively. u_i and ε_i are error terms indicating idiosyncratic effects on respectively youth and adult smoking. Parameter φ estimates the causal effect of youth smoking on adult smoking. Following Becker and Murphy (1988), φ measures the degree of addictiveness of smoking. The other notations, $\beta_1, \beta_2, \gamma_1, \gamma_2, \eta_1$, and η_2 are structural parameters of the model.

The CCHS only includes retrospective information on the youth smoking of its respondents. To avoid bias estimates that could arise from measurement errors in results of using retrospective information, I use the reduced form model of equations (2.1) and (2.2). The reduced form model is given by substituting equation model (2.1) into (2.2)

$$Smoking_{adult,i} = \mu + \beta_2 P_{adult,i} + \beta_3 P_{youth,i} + Z_i \lambda_1 + X_{adult,i} \lambda_2 + v_i \tag{2.3}$$

where $\mu = \delta + \phi\alpha$, $\beta_3 = \phi\beta_1$, $\lambda_1 = \phi\gamma_1 + \eta_1$, $\lambda_2 = \eta_2$ and $v_i = \phi u_i + \varepsilon_i + \phi\gamma_2 X_{youth,i}$. v_i is the error term of the reduced form model; it consists of X_{youth} plus a linear combination of the idiosyncratic effects. I assume the time-varying explanatory variables of youth smoking model, X_{youth}, are not correlated with cigarettes prices in adulthood. By which, these omitted variables in the reduced form model do not cause endogeneity problem, and so the estimated effect of cigarette prices are consistent and unbiased if the classical linear regression assumptions hold. The parameter of interest in the reduced form model is β_3, which estimates the causal effect of cigarette prices one faced as a youth on his smoking behaviour as an adult. It measures the combined effects of the degree of addictiveness and the contemporaneous prices elasticity of youth smoking. β_3 will equal to 0 if either smoking is not addictive and/or youth smoking is perfectly price inelastic.

Using the equation model (2.1), I estimate the effect of cigarette prices on deterred and deferred youth smoking. The dependent variable in this model is the age of respondent at which he smoked a whole cigarette for the first time, adopted the daily smoking habit, or it is a binary variable indicating the respondent adopted the daily smoke habit at a given age, such that he had not been a daily smoker up to that age. For the long-term effect of cigarette prices, I use the reduced form model (2.3), and use three dependent variables, each indicating different aspects of an adult smoking behaviours. First, I estimate the probability of adult daily smoking using a linear probability model. Second, I employ an OLS regression to estimate the smoking intensity. I use the number of cigarettes one smokes every day as a measure of the smoking intensity. Third, I employ a multinomial-logit model to test the long-term effect of cigarette prices on the smoking type of an individual. I categorize the smoking type of respondents into three groups: current smoker including daily and occasional smokers, former smoker, and never smoked. The dependent variable in the multinomial-logit model is a discrete variable that takes on values 0, 1, and 2 for never-smoked individuals, former smokers, and current smokers, respectively.

Glied (2002) examines the effect of cigarette price a youth faced at age 14 on his smoking behaviour as an adult. However, power of tests that she uses would be low if the ratio of cigarette price to personal income at age 14 is high relative to the ratio in succeeding years, so as a result a 14 years old youth may decide to postpone the decision of smoking initiation if he expects the ratio will reduce in succeeding years. In this study, cigarette prices are highly correlated over time, but they substantially vary from year to year (Figure 2.1). Meanwhile, the hazard of youth smoking initiation stays high for the entire youth, for which I take into account all the cigarette prices in youth by averaging prices in three age groups: 14, 14 to 16 and 12 to 18 years old.

To control the effects of all observable influential factors on smoking, I use a set of regressors in the reduced form model (2.3), including age (22 dummies), family income (11 dummies), education (4 dummies), ethnicity (6 dummies), a discrete variable indicating depressed respondents, household size, and a dummy variable indicating pregnant women. I add two more dummies to control on the time effect of the CCHS's cycles. Error terms in the reduced form model are set to be clustered at province and cycle of the survey. Furthermore, the estimated models are stratified by sex to control for the gender effect, and the estimations are robust to correct the covariance matrix for the heteroscedasticity problem.

To reexamine Glied's (2002) finding regarding the attenuation of the effect of cigarette prices over time, I modify the reduced form model to

$$Smoking_{adult,i} = \mu + \beta_2 P_{adult,i} + \beta_3 P_{youth,i} + \beta_4 P_{youth,i}.age_i + \beta_5 P_{youth,i}.age_i^2 + Z_i \lambda_1 + X_{adult,i} \lambda_2 + \nu_i \tag{2.4}$$

where $P_{youth,i}.age_i$ and $P_{youth,i}.age_i^2$ are interactions between the cigarette prices in youth with age and age-squared variables. Other notations are similar to equation model (2.3).

Having run regression model (2.4), I test changes in the effect of cigarette price with age by taking a derivative from equation (2.4) with respect to youth price

$$\frac{\partial Smoking_{adult,i}}{\partial P_{youth,i}} = \beta_3 + \beta_4.age_i + \beta_5.age_i^2 = 0 \tag{2.5}$$

The null hypothesis of the joint signification test is $H_0 : \beta_3 + \beta_4.age + \beta_5.age^2 = 0$ where age is given, which can be tested using an F-test.

2.5 Results

This section discusses the estimates of the models defined by equations (2.1), (2.3) and (2.4). I categorize the results into three groups. The first group consists of estimates of the effect of cigarette prices on youth and adult smoking. The second group of estimates displays the variation of the effect of cigarette prices with age. The last group illustrates simulated long-term effect of cigarette prices on the prevalence of smoking using counter factual changes in cigarette prices. The large sample size of the pooled CCHS cycles allows me to stratify all estimate models by gender.

2.5.1 Cigarette Price, Youth and Adult Smoking

The contemporaneous effect of cigarette price on youth smoking and its long-term effect on adult smoking are summarized in Tables 2.2 to 2.4. Table 2.2 displays how cigarette price deters or defers youth smoking. I use retrospective information on youth smoking for estimates reported in this table. The first and the second two set columns of Table 2.2 demonstrate the effect of cigarette prices a respondent faced as a youth on his propensity to experiment with a whole cigarette and the probability of become a daily smoker, respectively. I find cigarette prices in youth defer a youth's propensity to experiment with a whole cigarette, in which 10% increase in the average cigarette prices at ages 12 to 18 postpones the age of experimenting with a whole cigarette by 9 to 13 months. The effect on the delayed transition into daily smoking is weaker by 6 to 7 months. The results demonstrate that the cigarette price 12-18 is a more influential factor on youth smoking than the point price 14. The last two columns of Table 2.2 display the effect of cigarette prices one faces in youth on the probability of adopting a daily smoking habit at a given age or older, such that he will not be a daily smoker up to that age. In contrast to previous estimates, the point price 14 is more robust than the price 14-16 or the price 12-18. That is, a 10% increase in the point price 14 prolongs the state of transition into daily smoking by 4 to 7 months while an increase in the price 12-18 in not an effective factor.

Table 2.3 displays the long-term effect of cigarette prices on the probability of smoking and smoking intensity as an adult. The first two columns display the effect of cigarette prices 14, 14-16, and 12-18 on the probability of smoking. Using linear probability model, I find a 10% increase in the price 12-18 reduces the probability of smoking by 0.9% by ages 19 to 40. The effect of the price 14-16 is the same for females, but it is very weak and insignificant for males. Moreover, the point price 14 is not a determinant factor of the probability of smoking in adulthood for both genders.

The long-term effect of cigarette prices on the smoking intensity is displayed in the last two columns of Table 2.3. Using an OLS regression, I find cigarette prices a youth faces will reduce his smoking intensity as an adult, and the effect of the price 12-18 is more robust than the point price 14.

Current cigarette price, however, is not found to be a determinant factor of the prevalence of smoking in adulthood, while it adversely affects an adult's smoking intensity. The estimated effect appears to be more robust on males than females.

Table 2.4 displays the long-term effect of cigarette prices on a person's smoking type, using a multinomial-logit estimate model. The first two columns of the table display the probability that an adult who has never smoked would remain a non-smoker if he had faced 1% higher cigarette prices as a youth. It reports that a 10% increase in the cigarette price 12-18 can increase the probability of remaining a non-smoker by 0.7%. The second set of two columns of Table 2.4 display the effect of cigarette prices in youth on the likelihood of quitting smoking in later life. In this study, a quitter is defined as a former smoker who used to be a daily or an occasional smoker, or a person who has smoked at least 100 cigarettes in his lifetime, but reported that he is not smoking at the time of interview. I find cigarette prices in youth do not affect the propensity to quit smoking in later life. Furthermore, the last two columns of Table 2.4 display the long-term effect of cigarette prices on the likelihood of smoking, both daily and occasionally. Multinomial-logit estimates of adult smoking signify that a 10% increase in the cigarette price 12-18 reduces the probability of smoking in adulthood by almost 1.1% for females and 0.9% for males.

2.5.2 Change in the Effect of Cigarette Prices with Age

This section demonstrates the attenuation of the effect of cigarette prices on the probability of smoking over time as introduced in model (2.5). For this, I test the joint effects of cigarette prices in youth and their interactions with age and age-square where age is given. The results are illustrated in Figure 2.2 by gender. I only post the effects for the price 12-18 for space reserving, the results for the other prices are similar. I do not find any pattern for changes in the effect of cigarette prices in youth with the age of respondents for both males and females.

2.5.3 A Simulation on the Long-Term Effect of Cigarette Prices

As I discussed in the first section of the results, cigarette price in youth is a very weak determinant of smoking behaviours in adulthood, with point elasticity around 0.1. To visualize the magnitude of the long-term effect of the prices, I simulate the proportion of daily smokers using actual and counter factual cigarette prices. The most interesting scenario for a simulation, perhaps, is to determine what the proportion of smokers would be if the federal government of Canada had not raised cigarette taxes in 1991 to 1994. I simulate the proportion of the smokers using the linear probability model, assuming cigarette prices in 1991 to 1994 had remained constant at the 1990-level. The result is illustrated in top panel of Figure 2.3, which displays the difference in the proportion of the smokers for both sexes using the actual and the counter factual cigarette prices for the period 1991-1994. The figure illustrates the difference in the proportion of daily smokers at ages 21 to 29 for respondents who faced the actual and the counterfactual cigarette prices in youth. I find the difference is almost negligible.

In addition, I simulate the proportion of the smokers using a counterfactual 50% increase in cigarette prices in youth. The results are illustrated in bottom panel of Figure 2.3. The figure shows that the proportion of the smokers among adults aged 21 to 29 will drop by roughly 2% if the adults had faced 50% raise in cigarette prices when they were youths.

2.6 Conclusion

In this study, I examine the effect of cigarette prices one faces as a youth on his smoking behaviours as an adult. I exploit the large and disproportionate increases in the cigarette prices across the Canadian provinces and over time for the period 1979 to 2004, particularly the sharp increases in cigarette taxes from 1991 to 1994. Using a pooled sample of individuals in the CCHS cycles 2001, 2003, and 2005, I observe a cohort of individuals who are adolescents and adults in the survey, and faced disproportionate increases in tobacco prices as youths. Besides a simulation method, I try to quantify the long-term effect of cigarette prices in youth using linear probability, OLS, and multinomial-logit regressions. I test the contemporaneous effect of cigarette prices in youth on deferred and deterred youth smoking, and their long-term effect on the probability of smoking, the smoking intensity, and the smoking type in adulthood.

I find that although cigarette price increases can deter and defer youth smoking, their long-term effects are very weak but detectable. I find the probability of daily smoking in adulthood is reduced by 0.9% if average cigarette prices in ages 12 to 18 increased by 10%. I also find cigarette prices in youth can affect the smoking intensity of a person in adulthood. The overall results identify that the long-term effect of the average cigarette prices at ages 12 to 18 or 14 to16 are stronger than point price 14, and are more robust on females than males. This study advances the literature as it finds cigarette prices in the entire youth, average of the prices in ages 12 to 18, is more influential factor on adult smoking than the point price 14, which was used in the literature.

Using a simulation method, I show the proportion of the smokers would not be appreciably different from what currently prevails in Canada if the federal government of Canada had not raised cigarette taxes during the period 1991 to 1994. I also simulate the effect of a 50% increase in cigarette prices, and find that such a sharp rise in cigarette prices will slightly affect the proportion of the smokers in later life.

This paper stresses that a policy of increasing tobacco prices, conducted to reduce the smoking rate of the population by controlling youth smoking, is not very effective. This is, perhaps, because a high correlation between the youth smoking rate and the adult smoking rate does not imply causation between these two. A higher cigarette prices in youth appear to postpone the age at which a youth may experiment smoking or may become a daily smoker, but they do not have a substantial effect on the smoking behaviour of the individual in the long-term. Such a finding suggests that the smoking behaviour of an individual is more likely determined by an unobservable heterogeneous factor that persistently affects the individual smoking behaviour over time.

Table 2.1: Summary statistics of the CCHS cycles 2001, 2003, and 2005 for individuals aged 19 to 40 years old*

	Population		Female		Male	
	Mean	SD	Mean	SD	Mean	SD
Age Smoked First Whole Cigarette	15.11	3.23	14.93	3.06	15.31	3.39
Age Started Smoking Daily	16.67	3.26	16.41	3.19	16.98	3.32
The Proportion of Daily Smokers	0.25	0.43	0.24	0.42	0.27	0.45
The Proportion of Occasional Smokers	0.07	0.26	0.07	0.26	0.08	0.27
Smoking Intensity of a Daily Smoker	14.84	7.7	13.35	7.02	16.33	8.11
The Proportion of Current Smokers	0.33	0.47	0.31	0.46	0.35	0.48
The Proportion of Former Smokers	0.33	0.47	0.34	0.47	0.32	0.47
The Proportion of Never-Smoked Respondents	0.34	0.47	0.35	0.48	0.33	0.47
Current Cigarette Price Index	1.05	0.21	1.05	0.21	1.06	0.21
Cigarette Price at age 14	0.59	0.21	0.6	0.22	0.59	0.21
Cigarette Price at age 14-16	0.62	0.2	0.62	0.2	0.61	0.2
Cigarette Price at age 12-18	0.66	0.17	0.66	0.16	0.65	0.17
Highest level of education completed is less than high school	0.11	0.31	0.09	0.29	0.12	0.33
Highest level of education completed is high school	0.19	0.39	0.18	0.38	0.21	0.41
Achieved some post secondary education	0.11	0.32	0.11	0.32	0.11	0.31
Graduated from an university	0.59	0.49	0.62	0.49	0.56	0.5
Household has no income	0.004	0.065	0.004	0.063	0.005	0.07
Household's income < 5000	0.013	0.11	0.014	0.12	0.011	0.1
Household's income is between 5,000– 9,999	0.032	0.17	0.038	0.19	0.024	0.15
Household's income is between 10,000–14,999	0.05	0.21	0.06	0.24	0.033	0.18
Household's income is between 15,000–19,999	0.062	0.24	0.07	0.26	0.051	0.22
Household's income is between 20,000–29,999	0.11	0.31	0.11	0.31	0.098	0.3
Household's income is between 30,000–39,999	0.12	0.32	0.12	0.33	0.12	0.32
Household's income is between 40,000–49,999	0.11	0.32	0.11	0.31	0.12	0.32
Household's income is between 50,000–59,999	0.14	0.34	0.13	0.34	0.14	0.35
Household's income is between 60,000–79,999	0.16	0.36	0.15	0.34	0.24	0.43
Household's income is >= 80,000	0.21	0.41	0.19	0.39	0.24	0.43
Household size	2.89	1.36	2.97	1.34	2.79	1.38
Immigrant	0.1	0.3	0.1	0.3	0.1	0.3
Married	0.54	0.5	0.57	0.5	0.51	0.50
Pregnant	-	-	0.055	0.23	-	-
Number of Observations	95,408		50,684		44,313	

* Other variables included in the estimate models but not reported in the table are: a set of dummies indicating ethnicity, province of residence, cycle of the data, and age of respondents.

Table 2.2: Effects of cigarette prices in youth on the propensity of experimenting with cigarettes and the probability of transition into daily smoking

	Age Smoked a Whole Cigarette		Age Started Smoking Daily		Deterred Youth Smoking[1]	
	Male	Female	Male	Female	Male	Female
Price of cigarette at age 14	0.03***	0.03***	0.01**	0.02***	0.04***	0.023*
	(.009)	(.006)	(.007)	(.005)	(.01)	(.014)
Number of observations	29,501	32,639	20,319	22,640	44,321	50,687
Average price of cigarettes in ages 14-16	0.04***	0.04***	0.012	0.02***	0.08	0.06***
	(.01)	(.007)	(.009)	(.006)	(.01)	(.02)
Number of observations	29,390	32,552	20,247	22,594	44,100	50,439
Average price of cigarettes in ages 12-18	0.07***	0.05***	0.02**	0.03***	0.044	0.034
	(.001)	(.01)	(.01)	(.009)	(.036)	(.03)
Number of observations	25,132	28,216	17,254	19,519	37,991	43,932

Symbols *,** and *** denote significant at 0.1, 0.05, 0.01significance level, respectively.

1. Deterred smoking is measured by the hazard of falling into daily smoking at time t or later upon on not be a daily smoker up to time t, where t covers a range of ages within which the effect of cigarette prices are taken into account, i.e. 14, 14 to 16, or 12 to 18 years old.

Standard errors are robust and clustered at provinces and cycles of the data in all estimates. The estimated parameters are price elasticities. Figures in parentheses are standard errors. All regression models include immigration status at youth, marital status, country of birth, household size, education (dummies), age (dummies), family income (dummies), province of residence (dummies), depression, pregnancy for female respondents, and two dummy variables indicating cycles of the CCHS data.

Table 2.3: The estimated effects of cigarette prices in youth on the smoking propensity and smoking intensity in adulthood.

	Smoking Propensity[1]		Smoking Intensity[2]	
	Male	Female	Male	Female
Current price of cigarette	0.082	0.101	-0.31***	-.08
	(.12)	(.24)	(.095)	(0.1)
Price of cigarette at age 14	.024	-.01	-.019	-.023*
	(.037)	(.049)	(.014)	(.012)
Number of observations	44,313	50,684	12,044	11,968
Current price of cigarette	0.059	0.11	-.31**	-.072
	(.11)	(.24)	(.098)	(.1)
Average price of cigarettes at age 14-16	-.01	-.092*	-.006	-.034**
	(.03)	(.054)	(.017)	(.017)
Number of observations	44,092	50,436	12,001	11,951
Current price of cigarette	0.15	-.023	-0.3***	-0.023
	(.16)	(.21)	(.09)	(.11)
Average price of cigarettes at age 12-18	-.092	-.093*	-.047**	-.031
	(.06)	(.052)	(.024)	(.027)
Number of observations	37,987	43,929	10,332	10,472

Symbols *,** and *** denote significant at 0.1, 0.05, 0.01significance level, respectively.

1. A linear probability model is used. The dependent variable is a binary response variable indicating a person's smoking status as an adult.

2. An OLS regression is used. The dependent variable is the number of cigarettes a daily smoker smokes every day.

Standard errors are robust and clustered at provinces and cycles of the data. Estimated parameters are price elasticities derived by a linear probability model and an OLS regression. Figure in parentheses is standard error. All regression models include immigration status, marital status, country of birth, household size, education (dummies), age (dummies), family income (dummies), province of residence (dummies), depression, pregnancy for female respondents, and two dummy variables indicating cycles of the CCHS data.

Table 2.4: The estimated effects of cigarette prices in youth on the smoking type in dulthood[1]

	Never Smoked		Former Smoker		Current Smoker	
	Male	Female	Male	Female	Male	Female
Current price of cigarette	0.10	0.076	0.252	0.037	-.335***	-.139
	(.20)	(.15)	(.21)	(.18)	(.13)	(.16)
Price of cigarette at age 14	-.007	0.012	0.009	0.013	-.002	-.03
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
Number of observations	44,263	50,667	44,263	50,667	44,263	50,667
Current price of cigarette	0.101	0.055	0.254	0.037	-.339***	-.112
	(.20)	(.15)	(.21)	(.17)	(.13)	(.16)
Average price of cigarettes at age 14-16	0.041	0.054	-.015	0.024	-.025	-.095**
	(.04)	(.04)	(.04)	(.03)	(.05)	(.04)
Number of observations	44,042	50,420	44,042	50,420	44,042	50,420
Current price of cigarette	0.23	-.03	0.081	0.212	-.296**	-.209
	(.21)	(.18)	(.23)	(.18)	(.15)	(.14)
Average price of cigarettes at age 12-18	0.073	0.076*	0.025	.012	-.094	-.105*
	(.07)	(.05)	(.06)	(.06)	(.06)	(.06)
Number of observations	37,943	43,914	37,943	43,914	37,943	43,914

Symbols *,** and *** denote significant at 0.1, 0.05, 0.01significance level, respectively.

1. Multinomial-Logit estimates of adult's smoking type. Standard errors are robust and clustered at provinces and cycles of the data.

Estimated parameters are price elasticities. Figure in parentheses is standard error. All regression models include immigration status, marital status, country of birth, household size, education (dummies), age (dummies), family income (dummies), province of residence (dummies), depression, pregnancy for female respondents, and two dummy variables indicating cycles of the data.

Figure 2.1: Price of cigarettes by the provinces and the prices that Canadian adults aged 19-40 years in 2001 to 2005 faced in their youths.,

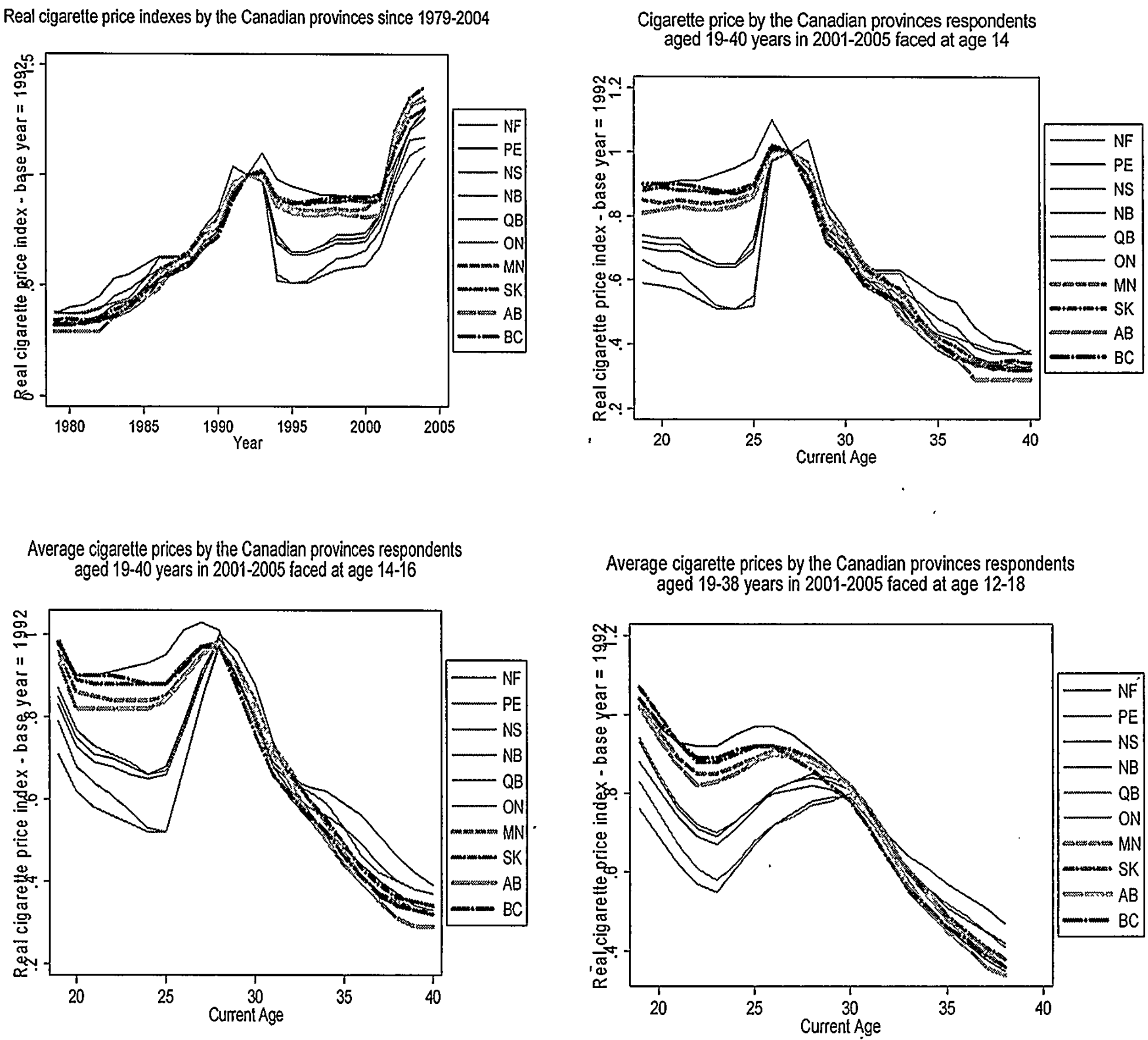
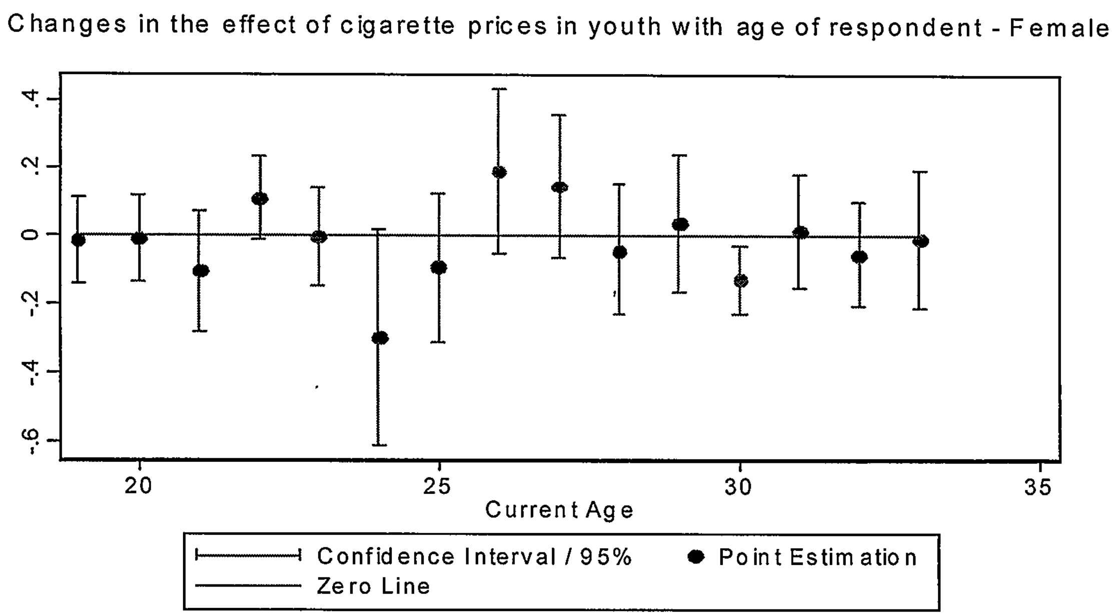


Figure 2.2: Change in the effect of cigarette prices in youth with age



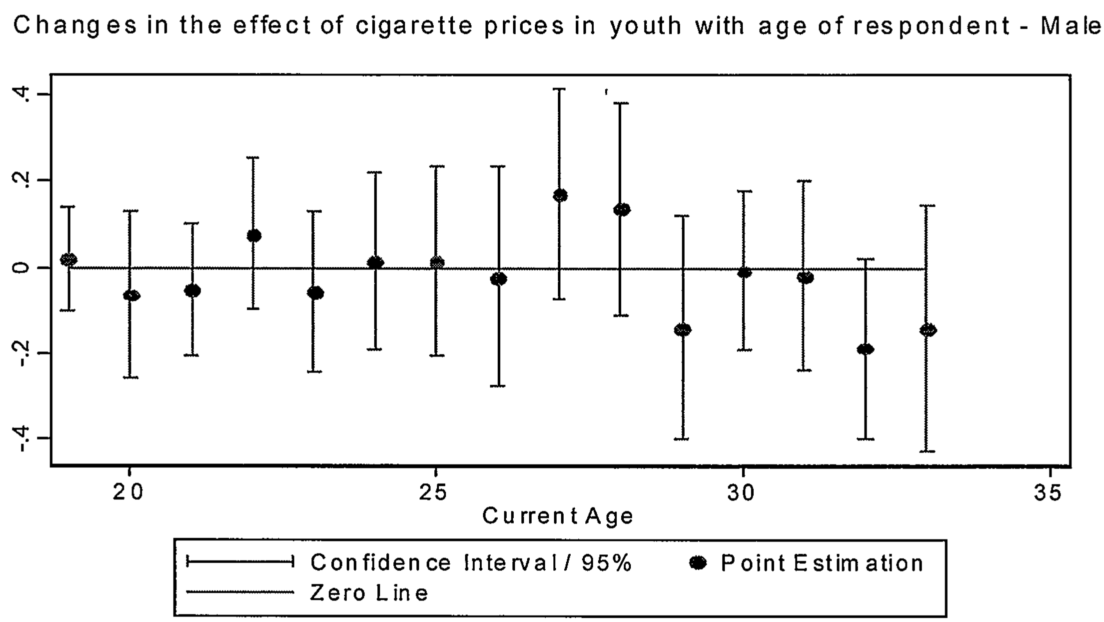
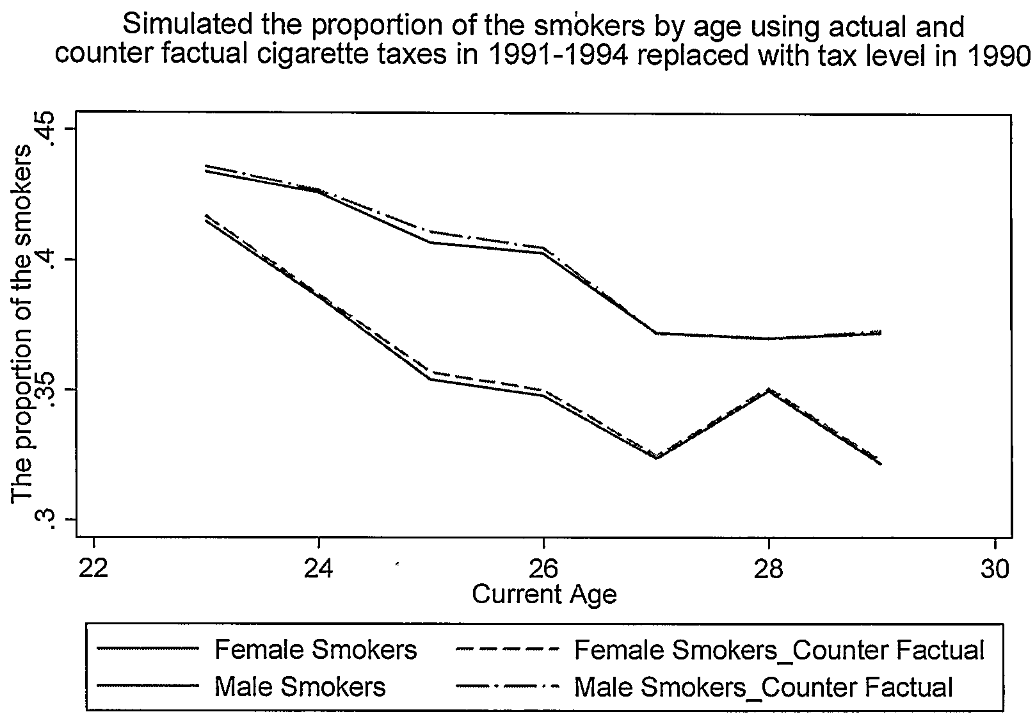
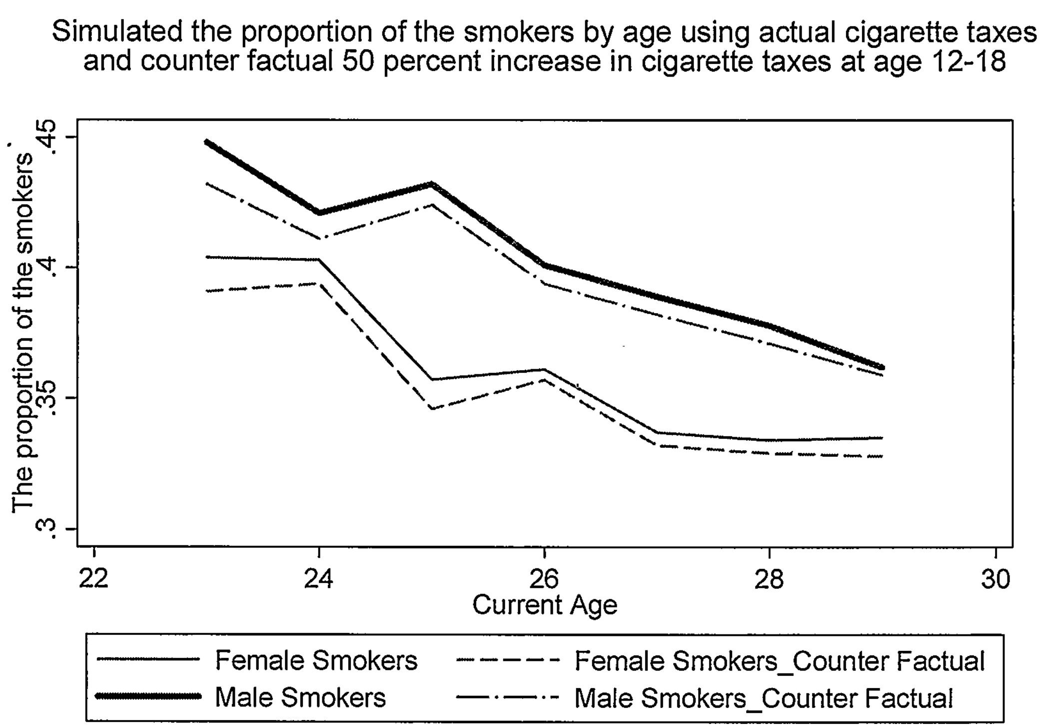


Figure 2.3: Simulated the proportion of the smokers in adulthood using the actual and counterfactual cigarette prices in youth.





Chapter 3

Short and medium term effects of parental leave on child outcomes: Evidence from the NLSCY

3.1 Introduction

Parenting styles and the family environment have substantially changed over the past century. The changes have increased in pace more recently due to rapid transformations in the family structure, economic conditions, and social reforms. Of these, the increases in the female labour force participation rate have had a large effect on the family environment, in which the role of mothers has changed from the main caregiver to an income earner. This has raised a credible policy concern, regarding how maternal employment affects child-mother interactions, and as a results a child's health and development. Within the literature, the effect on a newborn child has been at the core of many studies, because a child in his early life is extremely vulnerable to changes in the home environment. Parents' decisions for choosing between work and care giving simultaneously affect the family income and their child development, but the effect could vary across families, such that parents' abilities and the families' resources for balancing work and home responsibilities are unobservably heterogeneous. This study contributes to the literature by examining the causal effect of parental leave on health and development outcomes for children aged 7 to 72 months old. Moreover, this study advances the literature by controlling for the direction of the causality from the parental employment status to child's outcomes, and taking into account the effect of unobservable heterogeneity in the parents' skills using an endogenous switching regression model with error terms that follow a trivariate t-student distribution.

The majority of mothers are now employed and women's roles have changed both within and outside of the family. The female labor force participation rate has increased from 47% to 65% between 1980 and 1998 in the United States, and from under 50% to 70% between 1976 and 1994 in Canada. The changes are even greater if the employment status of females with children under age six years is only accounted for, from 38% to 59% in the United States and from 42% to 60% in Canada over the same periods. This has been followed by the continuous change in women's role in the family from the main caregiver to an income earner. Many parents now use non-parental childcare arrangements, while the substitute care providers and the caring environment do not necessary hold the minimum quality standards regarding the knowledge, the skills, and the equipments for an effective caregiving.

In psychological research, age zero to three years is a period when a child begins to learn to communicate, to understand informational content, and to develop social interaction (Hoffman, 1980; Belsky, 1988; Coleman, 1988). In particular, it is argued that the first year of a child's life is the most crucial, because the child is extremely attached to his mother and his brain is in its early stages of development toward independently perceiving and interacting with the environment. Research has identified several channels through which child-mother interactions and child's environment affect the child development. In particular, employment status of the mother affects direct and indirect influential factors of child development. Children of employed mothers more likely experience a degree of attachment insecurity if their caregivers do not treat them promptly, consistently, and appropriately; also, a close tie between a caregiver and a child promotes the child's feeling of security and capacity of form trusting relationship which is followed by higher scores in cognitive, emotional and social outcomes in later life (Belsky and Fearon, 2002). Thomas (2005, p.369) notes that "Research has shown that the child's dependency system does not become organized in a lasting form until sometimes near the end of the first year of life when the child begins to walk and talk... Children respond to what they regard as threatening situations in one of three types – [secure, uncertain, and avoidant]. Each type is the result of infants' experiences with the way their most intimate caregivers have reacted in the past."

In this study, I investigate the effect of parental leave on children outcomes. Parental leave is a key factor influencing child-parent interactions during the first year of children's lives. For example, Ruhm (2000a) notes that longer maternal leave is associated with the extended duration of breastfeeding and the greater investment of maternal times in caring for the infants, and so is associated with a lower infant mortality rate. Studies have shown that lower quality of child cares that substitute parental care might adversely affect children's outcomes (Brook-Gunn *et al.*, 2002; Walfogel *et al.,* 2002). Moreover, changes in a family's income due to maternal employment might affect children of high-income families in a different way from those of low-income families. For instance, an increase in family income may have a positive effect on the cognitive and development outcomes of children in low-income families if the parents are low skilled in both home and market activities, and vice versa (Vandell and Ramanan, 1992; Gagne, 2003). Furthermore, the family environment in childhood can have a long-lasting effect on the formation of learning skills and self-esteem (Heckman, 2006, 2007).

In this study, I exploit an expansion in the duration of job-protected and paid parental leave in Canada in 2001, and draw on a sample of children born in a ten-year window that surrounds the reform. The reform on December 31, 2000 resulted in an expansion in the duration of parental leave in Canada from roughly 6 to 12 months. Moreover, Statistics Canada conducted the National Longitudinal Survey of Children and Youth (NLSCY) in 1994 which has been repeated biannually since, in which it covered the periods before and after the reform. The reform created an exogenous change in mothers' propensities to return to work within the first year of their children's lives; and the NLSCY collected information on children born before and after the reform. Thus, the NLSCY and the reform together suit my estimates on the effects of parental leave. Children of the NLSCY and their parents form an appropriate sample in this study if the reform has implied an exogenous change in the duration of parental leave (Figure 3.1)[3].

The rational choice model of family output predicts that parents' decision in respect to timing of return to work after childbirth depends on the parents' abilities, the family's background and resources, and the child's specific needs for care and treatment, if any. Early research on maternal employment tended to used a univariate or direct-effect approach, that is, the difference between children of employed and unemployed mothers were examined without considering the complexity of other factors that might mediate the effects of maternal employment (Gottfried and Gottfried, 1988). The complex pattern of parents' decision for timing of return to work has challenged recent studies for finding the casual effect of parental leave (maternal employment). Some studies have tried to

[3] Baker and Milligan (2007, 2008) find the duration of staying home after giving birth among the mothers who were likely eligible for the leave increased by three to three and half months after the reform.

identify the causal effect of maternal employment on children's outcomes, and have accounted for the effect of unobservable heterogeneity in mothers' skills[4].

Using an endogenous switching regression model and the natural experiment on expansion in parental leave in Canada, I am able to control for unobserved heterogeneity in parents' skills and the direction of the causality from the parental employment to children's outcomes. I test the effect of parental leave on 74 different child outcomes, including health, temperament, development, behaviour, milestone achievements, literacy, parenting and family functioning in order to identify the effect of parental leave on different aspects of a child's development. In particular, I find longer parental leave causes an extended period that the mother stays' home after giving birth, and a longer duration of breastfeeding, in which breastfeeding increases by 4 weeks on average in the whole population after the reform. This is compared with a 25-week increase in the treated population, which is defined by the parents who stayed home between 7 to 12 months after birth. I find parental leave has almost no contemporaneous effect on children, excluding the effect on breastfeeding. In addition, children whose parents took a long parental leave are better off with respect to temperament and cognitive development, but are worse off with respect to aggressive behaviour, family functioning, and hostile parenting when aged two to three years. However, the negative effects appear to be temporary and they disappear later in the children's lives, when aged four to five years.

3.2 Previous Studies

Until recently, the economics literature on maternal employment had mainly focused on the economic consequences of parental leave policies, such as the effect on

---

[4] Lefebvre and Merrigan (2000); Neidell (2000); Gagne (2003); Ruhm (2004); Berger *et al.* (2005); Baker and Milligan (2007, 2008); and Chia (2008).

labour market outcomes, the gender wage differential, and the propensity to have a child[5]. However, availability of longitudinal survey data over the last decade has turned researchers' attentions to the effect of maternity leave on a mother's physical and mental health, and on a child's health and cognitive development.

The majority of the research examining the effect of maternity leave on children's development uses the National Longitudinal Survey of Youth (NLSY) in the United States. It employs the Peabody Picture Vocabulary Test (PPVT) as a measure of cognitive development for young children, and the Behavioral Problem Index (BPI) and school reading and math scores as measures of cognitive outcomes for older children. Studies using the NLSY mainly find an adverse effect of early maternal employment on the children's cognitive development[6]. In particular, Desai *et al.* (1989) find a negative effect on boys in high-income families. An adverse effect is found among children whose mothers work more than 30 hours per week in the early years of the children's lives (Belsky and Eggebeen, 1991; Blau and Grossberg, 1992). Blau and Grossberg (1992) note the negative effect of the mother working in the first year of the children's lives is compensated if the mothers work more weeks in the second and third years than the first year. Han *et al.* (2001) and Waldfogel *et al.* (2002) find a persistent negative effect of early maternal employment on children's cognitive outcomes; but they also find maternal employment in the second and the third year of children's lives has a positive effect on non-Hispanic children, but not on African or Hispanic children. Zick *et al.* (2001) use data from the National Survey of Families and Households to show that parents in employed-mother households are more likely to engage in reading/homework activities

[5] Goldin (1990); Lyness and Judiesch (2001); Drolet (2001); and Hanratty and Trzcinski (2006)
[6] Desai *et al.* (1989); Baydar and Brooks-Gunn (1991); Belsky and Eggebeen (1991); Blau and Grossberg (1992); Vandell and Ramanan (1992); Parcel and Menaghan (1994); and Greenstein (1995)

with their children than parents in households in which the mother stays home. They also note that the mothers' employment status is correlated with fewer behavioural problems and higher grades for children aged 11 and younger. Using data from 900 European American children for the National Institute of Child Health and Human Development Study of Early Child Care, Brooks-Gunn *et al.* (2002) examine the effect of return to work within the first nine months of children's lives on the cognitive scores of the children up to three years old. They find an adverse effect when the mother works 30 hours or more per week, even after controlling for child-care quality, quality of the home environment, and maternal sensitivity. Using the NLSY, Berger *et al.* (2005) stress that returning to work within 12 weeks after giving birth has a negative effect on the duration of breastfeeding, is associated with reduced immunizations, and increases in externalizing behavioural problems. They find the results are more pronounced for women who returned to full-time jobs. Vandell and Ramanan (1992) stress that children of low-income families have benefited from early maternal employment in respect to their scores in the Peabody Individual Achievement Test for Math in the second grade of schooling, but find a negative effect on their PPVT scores. Harvey (1999) examines the effect of early maternal employment on children's outcomes up to age 12 years and finds no significant effects on cognitive and most behavioural outcomes when White, African American, and Hispanic children are analyzed.

The research in Canada has emerged mostly after the National Longitudinal Survey of Children and Youth (NLSCY) became available, particularly using the Early Childhood Development (ECD) cohorts of the NLSCY. Using cycles 1 and 2 of the data, Lipps and Yiptong-Avila (1999) find children who were in a non-parental care arrangement two years earlier are more likely to have top scores in mathematics than those were not. Lefebvre and Merrigan (2000) use cycle 1 of the NLSCY and a family fixed effects model to conclude that children four to five years old who use child-care arrangement appear to have higher PPVT scores, and that child care arrangement is correlated to a better Motor Social Development (PVD) score for children 0-47 months old. Gagne (2003) uses the first three cycles of the NLSCY, controls for family fixed effects and finds an adverse effect of early maternal work on the PPVT scores of children three and half to five years old. However, she notes that the adverse effect is weaker than the effect estimated in the literature. She notes "…children of mothers with above (bellow) average parenting skills and educations have slightly worse (better) PPVT scores when their mothers work full-time." Ram *et al.* (2004) use children of the NLSCY who were three years or less in 1994-1995 and became seven to nine years old in 2000-2001, and find early maternal employment for these children reduced their PPVT scores, but that there was no effect on their math scores. They also stress that children of high-income families are more vulnerable in respect to their verbal skills, and children of low-income families are more vulnerable in respect to their mathematical skills. Chia (2008) attempts to estimate the causal effect of early maternal employment on the risk of having an overweight/obese child and finds that an increase in the mother's work intensity when she returned to work after birth of her child and before the child started school is associated with an increase in the risk of the child becoming overweight or obese later in childhood. Moreover, the literature emphasizes that the specific needs of a child can affect a mother's decision to return to work in different ways. Gould (2004) finds mothers of children whose needs require more time-intensive care are less likely to go back to work in early months of their children's lives. In contrast, mothers of children whose needs require more financial-intensive care are more likely to go back to work in early months of their children's lives.

Some studies have tried to take into account the effect of heterogeneity in the parents' abilities for balancing the care giving and work responsibilities[7]. These studies either use household fixed effects, difference in children outcomes between siblings whose mothers' decisions for returning to work differ between the siblings, or a propensity score matching model to address the observable/unobservable heterogeneity effects of maternal employment. Neidell (2000) uses family fixed effects and concludes uninterrupted time investment of up to one year would offer lasting benefits, particularly for children's non-cognitive outcomes. Ruhm (2004) uses the pre-birth employment of the mothers as an explanatory variable to control on unobserved heterogeneity in the early maternal employment. In respect to the effect of maternal employment in the first year of children's lives, his inferences are consistent with the literature. He finds the mother working in the first year of a child's life is associated with decreases in the verbal ability of three and four year olds. However, in contrast to previous studies, he finds maternal employment in the first three years of a child's life is associated with lower reading and mathematics achievement for the children when aged five and six years old. Baker and Milligan (2005, 2007) use the Labor Force Survey (LFS) and the NLSCY, respectively and exploit the parental leave expansion in Canada to examine the effect of the reform on the duration of breastfeeding and the children's health. They find the parental leave expansion followed by three to three and half months increase in the

[7] Reader is referred to Neidell (2000); Lefebvre and Merrigan (2000); Gagne (2003); Ruhm (2004); Berger *et al.* (2005); Baker and Milligan (2005, 2007, 2008) and Chia (2008)

duration that the mothers, who were more likely eligible for the leave, stay in home and that this increase has prolonged the duration of breastfeeding by one month among these mothers. They argue that since the expansion in parental leave was country-wide and as there was not any control group within Canada, they use three approaches to deal with this challenge. First, they examine inter-cohort changes in the outcomes before and after the expansion. Second, they estimate the effect of parental leave expansion for the full sample of the mothers, who were and were not eligible for the leave, and a sub-sample of the mothers, who were likely eligible for the leave. Then, they compare the results to find out which group drives the estimates. For the final approach, they construct a control group of children 25-36 months old from pre-reform cohorts for treated children of 13-24 months old from after-reform cohorts. They were only able to use the control group method for a few contemporaneous outcomes that they observed for the both groups in their sample, because some questions in the survey are only asked of specific age groups. They do not find any robust evidence that the increase in breastfeeding time has a beneficial effect on the children's health.

In another study, Baker and Milligan (2008) use the NLSCY to test the effect of the parental leave expansion on the usage of non-parental child-care services and a variety of children's development outcomes, including temperament, social/motor development, milestone achievements, family functioning and children's social and family environment. They stress that an increase in maternal leave crowds out care for children in other homes by non-licensed persons, but the effect on children's outcomes is negligible. They observe some improvements in temperament, but it appears as a common trend across all mothers who were or were not eligible for the leave.

3.3 Policy Environment

The optimal duration of parental leave is a decision of parents concerning well-being of their child and fulfilling work responsibilities simultaneously, which both directly or indirectly depend on parents' abilities, the family environment, and labour force legislations. The decision, however, is an endogenous treatment, as a result of the interaction between many observable and unobservable factors. I exploit the reform in parental leave mandates in Canada in 2001 to draw on a sample of children whose parents' decisions in respect to the duration that they stayed in home after birth varied exogenously by the reform.

3.3.1 Parental Leave Policy in Canada

Parental leave is a worldwide policy that varies from long, paid and job-protected in Canada and Europe to short, unpaid, and targeted in the United States. The World Health Organization (2000) recommends that at least 16 weeks of job-leave are necessary for mothers after delivery to retrieve their strength, both physically and mentally, for being able to return to work and for protecting their children's health.

The parental leave policy in Canada is a part of the employment insurance (EI) program in which eligibility for the leave depends on the labour force participation of the parents before the birth of a child. Benefits and income replacement of the leave are financed through the EI program which is a federal program, but the right of return to the previous or a similar job is based on provincial labour standard legislation. The parental leave mandates in Canada have changed since December 31, 2000, in which eligibility for the leave prior to the reform, which was based on minimum 700 hours of work during the last 12 months preceding the leave, was reduced to 600 hours. Benefits remain the same, which are based on earnings over the six months prior to the leave, with an income replacement rate of 55% up to a cap of $39,000. The most significant change by the reform was an increase in the maximum duration of parental leave. The duration of parental leave in Canada is split between the mother and the father of a newborn or adopted child. The duration of leave before the reform was substantially disproportionate across the Canadian provinces, from a low of 18 weeks in Alberta to a high of 35 weeks in Ontario (excluding Quebec that expanded the leave up to 70 weeks in March 1997). The leave prior to the reform was typically an initial 15 weeks of paid benefits for eligible mothers plus additional 10 to 20 weeks which could be split between the mother and the father. The duration of the leave has become more proportionate across the provinces since the effective date of the reform, in which all provinces, excluding Quebec, expanded the leave to 52 weeks, of which the duration that is split between parents rose to 35 paid-weeks plus 2 unpaid-weeks for children born on December 31, 2000 or later[8].

3.3.2 Other Policy Changes

Baker and Milligan (2007, 2008) argue that there have been two major policy changes parallel to the reform in parental leave that might affect children's well being, and so can affect inferences in any study that exploits the 2001 reform for testing the effect of parental leave on children's outcomes. The first is a "$5 a day" universal subsidy program for childcare in Quebec, introduced in 1997 and extended to children aged up to one in September 2000. Baker, Gruber, and Milligan (2005) argue that the program has largely affected non-parental care and is associated with family well-being

[8] Alberta and Saskatchewan delayed adopting the reform, the effective date was February 07, 2001 in Alberta, and June 14, 2001 in Saskatchewan.

in Quebec. The second is an increase in federal National Child Benefits from $605 in 1988 to $1,293 by 2002 for each child. Milligan and Stabile (2007) note that the increase in child benefits had a substantial effect on the employment of single mothers, because low-income families and single mothers were the main beneficiaries of the program. In this study, I exclude children born in Quebec or living in single parent families to avoid the parallel policies affecting my inferences.

3.4 Data

This study uses the NLSCY from 1996 to 2005. The NLSCY is nationally representative survey of children in Canada. The NLSCY was conducted jointly by Statistics Canada and Social Development Canada in 1994 and has been repeated biannually since then. The survey is designed to collect information regarding factors influencing child's social, emotional and behavioural development, and to monitor the impact of these factors over time up to early adulthood of the children. Beginning cycle two of the survey, cycle 1996/97, a cohort of children zero to one year old, the Early Childhood Development (ECD) cohort, has been added to each cycle of the survey, and children of these cohorts are to be followed up to age five. Children of the ECD cohorts mostly appear in three subsequent cycles of the survey. That is, for instance, children of ECD cohort in cycle 2 appear in cycle 3 and 4, and children of the ECD cohort in cycle 3 appear in cycle 4 and 5, and so on.

I draw on a sample of children born in a ten-year window surrounding the five years before and after the reform in parental leave mandates in Canada. I use the last six available waves of the NLSCY: cycle2 (1996/97) cycle3 (1998/99), cycle4 (2000/01), cycle5 (2002/03), cycle6 (2004/05) and cycle7 (2006/07). I observe the children up to age five. Since certain outcomes are only available for specific age groups, I estimate the children's outcomes in four age groups; 7 to 12, 13 to 24, 25 to 48 and 49 to 72 months old.

The ECD cohorts include two files: a primary file that consists of all children, and an education file that includes children four to five years old. Information on child outcomes in the primary files is reported by the Person Most Knowledgeable (PMK) in the household of a child who could be the biological mother, father or anyone else. However, I restrict my sample to children for whom PMKs are the biological parents of the children (98% of the sample) to assure the accuracy of the data; This exclusion also reduces possible bias in my inferences if children of families where biological parents are not present in the household may be exposed to an environment that affects their outcomes in different ways than it affects other children. Moreover, I exclude children living in single parent families (12% of the sample) because the change in federal National Child Benefits program in 2002 affected the employment of single parents (Milligan and Stabile, 2007). I also exclude children born in Quebec (17.7% of the sample) because parental leave in Quebec expanded to 70 months in 1997 which does not coordinate with my sample. In addition, the "$5 a day" program in Quebec in 2000 has affected the well-being of these children (Baker, *et al.*, 2005). I use 74 different child outcomes that are age specific, including health, behaviour, milestone achievements, breastfeeding, literacy, parenting, cognitive development and family functioning. Table 3.6 in Appendix 3.A describes all outcomes used in this study.

The NLSCY is a rich survey that collects information on children development, but does not include complete information on the work status of the mothers and the fathers before birth. Thus, I cannot fully observe parents who were/were not eligible for parental leave. Moreover, the survey does not include complete information on the pattern of parental leave split between the parents. That is, I do not fully observe which parents, the mother or the father, took the leave. To deal with these problems, I restrict my estimations samples to sub-samples of the mothers who returned to work within the first 12 months after giving birth. By which, I anticipate the mothers who were working before birth are more likely to return to work after birth. Although, not all the mothers who work are eligible for leave, Baker and Milligan (2007, 2008) show restricting the sample to the sub-sample of the mothers who returned to work is a proper technique to approximate the effect of parental leave on child outcomes[9]. Moreover, as I do not observe the duration taken by the fathers, my estimates of the effect of parental leave are unbiased only if the pattern of split in parental leave between the mothers and the fathers has not change by the reform.

Figure 3.1 displays changes in the duration that mothers stayed home after birth for the periods before and after the reform (henceforth, the duration refers to parental leave in this paper). Summary descriptive of the NLSCY data by the reform and the age groups of the children are reported in Table 3.1.

3.5 Empirical Method and Identification

Income and the well being of a newborn child are the major concerns of the parents for returning to work after birth. The decision, however, depends on parents'

[9] Baker and Milligan (2007, 2008) use Statistics Canada's (2006) report on the proportions of the mothers with children under age one year old who worked in the 12 months preceding giving birth, and those who were eligible for parental leave between years 2001 to 2005, to adjust their estimates. They derived a lower and upper bounds for their estimates using scaling factors 1.33 and 1.57, respectively which are calculated based on the proportion of the mothers who were working before the birth of their children and those who were eligible for the leave between years 2001 to 2005 in Canada. They find restricting the sample to the mothers who returned to work within the first 12 months of their children's lives gives an estimated effect that falls within the upper and the lower bounds calculated by the scaling factors.

abilities and the family's resources for balancing work and family responsibilities, which are unobservably heterogeneous across families. An appropriate estimate model should be capable of identifying the direction of the causality from parental leave to the child's outcomes, upon controlling for unobservable heterogeneity in the parents' skills. The literature has employed OLS regression, instrumental variable technique, differences in differences method, fixed effects estimates, and propensity score matching regression. However, a more advanced structural model is required to control for both the causality and the heterogeneity in the effect of parental leave on child's outcomes. For this, I employ an endogenous switching regression model to distinguish the effect of treatment on treated (TT) from the average treatment effect (ATE). TT measures the effect of parental leave on children whose parents are randomly selected from the treated population. In this study, the treated population consists of the parents who took a parental leave between 7 to 12 months. ATE measures the effect of parental leave in the whole population. If there is no heterogeneity in the effect of unobservable parents' skills, then one should expect the ATE to be equal to the TT, upon controlling for observable characteristics of parents.

3.5.1 Endogenous Switching Regression under Normality Assumption

Since the reform in December 2000 increased the duration of parental leave on average from 6 to 12 months, I estimate the effect of a parental leave between 7 to 12 months, compared with the effect of a parental leave equal or shorter than 6 months, on children's health and development outcomes. Since the parents elect the duration of the leave, the effect of sample selection mediates the effect of parental leave and so induces biased estimations. To disentangle the effect of sample selection from the effect of parental leave, an endogenous switching regression model is a proper estimation method. This regression method can control for the effect of sample selection by approximating the parents' propensity to take a long parental leave. To do this, I construct a selection criterion using a binary variable that takes on the value 1 for the parents who stayed in home longer than six months after birth, otherwise takes on the value 0.

Let y_{1i} denote outcome of child i whose parents, the mother or the father, stayed in home between 7 to 12 months after birth, and y_{0i} denote her outcome if the parents stayed home 6 months or shorter. The average effect of a long parental leave for children whose parents took a leave between 7 to 12 months (TT) is given by

$$TT = E(y_{1i} \mid leave = 1) - E(y_{0i} \mid leave = 1) \quad , \tag{3.1}$$

and the average effect for all children in the population is

$$ATE = E(y_{1i} \mid leave = 1) - E(y_{0i} \mid leave = 0) \tag{3.2}$$

where $leave$ is a binary variable indicating the parents who stayed home longer than six months. Equation (3.1) measures the treatment effect on treated (TT). However, I do not observe a child in the two states simultaneously; moreover, I cannot use average outcomes of children in the other state, $E(y_{0i} \mid leave = 0)$, as a proxy for the second term on the right hand side of equation (3.1), because children and their parents are heterogeneously different across the states.

The causal effect of parental leave can be examined using an endogenous switching regression model. The model consists of one selection and two outcome equations,

$$leave_i^* = z_i\gamma + \mathrm{u}_i \text{ and } \begin{Bmatrix} leave_i = 1 \text{ if } z_i\gamma + \mathrm{u}_i > 6 \\ leave_i = 0 \text{ if } z_i\gamma + \mathrm{u}_i \le 6 \end{Bmatrix} \text{ (selection equation)} \tag{3.3}$$

$$\begin{aligned} y_{1i} &= x_{1i}\beta_1 + \varepsilon_{1i} \quad \text{if } leave_i = 1 \\ y_{0i} &= x_{0i}\beta_0 + \varepsilon_{0i} \quad \text{if } leave_i = 0 \end{aligned} \quad \text{(outcome equations)} \tag{3.4}$$

where z is a vector of explanatory variables in the selection equation. x_1 and x_0 are vectors of explanatory variables in the outcome equations for children whose parents stayed home longer or shorter than six months after birth, respectively. In the rest of paper, I assume explanatory variables in both regimes are the same, $x_1 = x_0 = x$. u_i denotes the idiosyncratic effect of returning to work in the selection equation, which is correlated with error terms in the outcome equations, ε_{1i} and ε_{0i}. For example, parents who are high skilled both at market activities and home responsibilities are less likely stay home for a long time after birth, and their children may experience higher development outcomes, $corr(u_i, \varepsilon_{0i}) < 0$. Notations γ, β_1 and β_0 are vectors of parameters to be estimated and the other notations are the same as the above.

In the econometrics textbooks, it is common to assume that the variance-covariance matrix of the error terms has a trivariate normal distribution which is given as

$$\Omega = \begin{bmatrix} \sigma_u^2 & . & . \\ \sigma_{u1} & \sigma_1^2 & . \\ \sigma_{u0} & . & \sigma_0^2 \end{bmatrix}$$

where σ_u^2 is a variance of the error term in the selection equation, which is set to 1 for the probit estimate, and σ_1^2 and σ_0^2 are variances of the error terms in the outcome equations. σ_{u1} is a covariance of u and ε_1, and σ_{u0} is a covariance of u and ε_0. The covariance between ε_0 and ε_1 is not identified as y_{1i} and y_{0i} are never observed simultaneously. A Heckman two-step method (Heckman, 1979) can be employed to estimate parameters of the switching regression model (3.3) to (3.4). To do that, one must first estimate the selection equation using a probit regression and then derive Mills ratios $\lambda_{1i} = f(z_i\gamma)/F(z_i\gamma)$ and $\lambda_{0i} = f(z_i\gamma)/(1-F(z_i\gamma))$, where $f(.)$ and $F(.)$ are the density and the cumulative density of the normal distribution functions, respectively. Second, the outcome equations can be estimated using an OLS regression with the corresponding inverse Mills ratio as a regressor correcting for selection bias in the effect of parental leave. Standard errors in the two-step method are wrong, but can be corrected using a bootstrap method.

Since the duration of parental leave is an endogenous variable, the model is parametrically identified if there exists at least one element of z which does not appear in x. Moreover, the excluded element is a valid instrument if, and only if it exogenously affects the parents' decision for returning to work with no direct effect on the children's outcomes. Baker and Milligan (2007, 2008) find the parental leave expansion in 2001 increased the duration that mothers stay home by three to three and half months for children born in 2001 and later, so a binary variable that indicates these children is an appropriate instrumental variable in the absence of any secular trend in children's outcomes over time. After correcting for selection bias, the average outcomes in each regime is given by

$$E(y_{1i} \mid leave = 1, z_i, x_{1i}) = x_{1i}\beta_1 + \sigma_1\rho_1\lambda_{1i} \tag{3.5}$$

and

$$E(y_{0i} \mid leave = 0, z_i, x_{0i}) = x_{0i}\beta_0 - \sigma_0\rho_0\lambda_{0i} \tag{3.6}$$

where $\rho_1 = \dfrac{\sigma_{u1}^2}{\sigma_u\sigma_1}$ and $\rho_0 = \dfrac{\sigma_{u0}^2}{\sigma_u\sigma_0}$ are the coefficients of correlation between u and ε_1 and ε_0, respectively.

3.5.2 A More Flexible Functional Form for the Endogenous Switching Regression

The normality assumption in the above model is criticized in that it suffers from a lack of robustness when the true distribution departures from normality (Goldberger, 1983; Paarsch, 1984). For this, I first test the normality assumption. Second, upon rejecting the assumption, I employ a more flexible functional form introduced by Heckman, Tobias, and Vytlacil (2000) that assumes the error terms follow a trivariate t-student distribution. This distribution is more flexible, as it has fat tails for low degrees of freedom and converges to a normal distribution as the degree of freedom increases.

3.5.2.1 Testing Normality Assumption

If u in the selection equation follows a normal distribution, the conditional mean of ε_0 or ε_1 given u is linear in u (Olsen, 1980). Thus, a simple test of normality assumption can be performed by adding non-linear transformed Mills ratios to the model (3.5) or (3.6), and then performing an F-test to get the significance level of the joint effects of Mills ratio and its non-linear transformed variables. For example, squared and cubic forms of Mills ratio can be used as non-linear transformations of Mills ratio. If the normality assumption holds, the coefficients of the quadratic and/or cubic Mills ratio are insignificant.

Moreover, to emphasize the robustness of the normality assumption test, I use a semi-parametric estimate of model (3.5) or (3.6) using a partial linear regression method to regress a child's outcome on the corresponding Mills ratio with an unknown functional form, while other explanatory variables are added to the regression model in a linear form. Using the semi-parametric method, I draw a scatter plot of the smoothed value of the outcome variable[10] against $z\hat{\gamma}$ and then compare it with a scatter plot of the Mills ratio against $z\hat{\gamma}$. Their graphical appearance should be similar if the normality assumption holds.

3.5.2.2 Endogenous Switching Regression With t-student Distribution

Following Heckman, Tobias, and Vytlacil (2000), in a latent variable framework, the selection equation assigns mothers to the treated state ($leave$=1) provided $6 - z_i\gamma \le u_i$; and if u has a univariate t-distribution, the truncated mean of ε_1 given u in the treated state is given by $E(\varepsilon_{1i} \mid u_i > 6 - z_i\gamma) = \left(\dfrac{\nu + u_i^2}{\nu - 1}\right)\dfrac{t_\nu(u_i)}{T_\nu(u_i)}$, ν>1 where t_ν and T_ν are the density and cumulative density of t-student distribution with ν degree of freedom, respectively.

In a case that a probit estimate is used for estimating the selection equation, u is distributed normally. Thus, the conditional mean of outcomes in models (3.5) and (3.6) under the t-student distribution assumption can be derived as described in Heckman *et al.* (2000). The selection biased terms with error terms that follow a trivariate t-student distribution, when u is normally distributed, are given by $\lambda_{1i}^t = \left(\dfrac{\nu + J_{T_\nu}(z\gamma)^2}{\nu - 1}\right)\dfrac{t_\nu(J_{T_\nu}(z\gamma))}{F(z\gamma)}$ and $\lambda_{0i}^t = \left(\dfrac{\nu + J_{T_\nu}(z\gamma)^2}{\nu - 1}\right)\dfrac{t_\nu(J_{T_\nu}(z\gamma))}{1 - F(z\gamma)}$ for ν>1 where $t_\nu(.)$ is the t-student density distribution with ν degree of freedom, $F(.)$ is the cumulative normal distribution and $J_{T_\nu}(z\gamma) = T_\nu^{-1}(F(z\gamma))$ where T_ν^{-1} is the inverse cumulative t-student distribution with ν

[10] Smoothed value of an outcome variable can be derived by Lowess method in Stata using "plreg" module (Lokshin, Stata Journal, Volume 6(3) PP. 377-83). Lowess method caries out a locally weighted regression of the outcome variable on explanatory variables when the data is sorted in the ascending order of the values of the Mills ratio.

degree of freedom. The optimum degree of freedom is derived by minimizing the sum of square of residuals, because the number of parameters of the models does not change with the degree of freedom. Given the above structural equations and selection bias terms, the average treatment effect (ATE) is estimated by

$$ATE(x) = E(y_1 - y_0 \mid x) = x(\beta_1 - \beta_0) \tag{3.7}$$

and the average treatment effect on treated (TT) is estimated by

$$TT(x, z, leave = 1) = x(\beta_1 - \beta_0) + (\rho_1\sigma_1 - \rho_0\sigma_0)\lambda_1^i \tag{3.8}$$

3.6 Results

In this section, I report the estimates of equation models (3.2) to (3.8). I estimate the models for children in four age groups, 7 to 12, 13 to 24, 25 to 48 and 49 to 72 months to find out the short and medium term effects of parental leave. I restrict the samples to mothers who returned to work within the first 12 months of their children's lives to approximate the effect of parental leave for the parents who were eligible for the leave, except for the estimates for children aged 7 to 12 months that include all mothers.

I report the estimated results in three sections. The first section includes results from normality assumption tests using an OLS and semi-parametric estimates of models (3.5) and (3.6). The second section displays the average treatment effect and the effect of treatment on treated for a variety of children's outcomes. In the last section, I illustrate the variation of the effect of parental leave with family income, the education of parents, and the age of the mother at childbirth.

3.6.1 Normality Assumption

Table 3.7 in Appendix 3.B displays the joint statistical significant levels of the sample selection coefficients, Mills ratio and the quadratic and cubic Mills ratios, as introduced in models (3.5) and (3.6). The results are derived using an OLS regression. Overall, the table displays that the selection bias term, Mills ratio variable, is not linearly correlated with children's outcome as the coefficients of the quadratic and cubic Mills ratios are statistically different from zero. This means that the normality assumption does not hold. In particular, I found the normality assumption is mainly violated for outcomes that are significantly affected by parental leave (the results are reported in Tables 3.2 to 3.5). Moreover, I use a semi-parametric estimate to test the nonlinearity between the selection bias term and child's outcome without imposing any particular functional form to the correlation function. The results are reported in Figure 3.2. The results are illustrated as pair of scatter plots. The first plot in each pair displays the correlation pattern between the weighted outcome and the predicted mother's propensity to return to work after birth ($z\hat{\gamma}$) using Lowess method by "plreg" command in Stata software[11]. Its pair plot displays the correlation pattern between the selection bias variable and the predicted propensity score ($z\hat{\gamma}$). Comparing the two plots, I find the normality assumption does not hold. In other words, since the two scatter plots represent different correlation functions between $z\hat{\gamma}$ and child's outcomes and the mother's propensity to return to work, I conclude that the normality assumption does not hold (due to space reserving, I do not report the plots for all outcomes).

3.6.2 Parental Leave and Child's Outcomes

Tables 3.2 to 3.5 display estimates of models (3.7) and (3.8), which display the average treatment effect and the effect of treatment on treated. I also report the correlation between unobservable factors that affect a mother's propensity to take a long

[11] Details about the command are available at http://www.stata-journal.com/article.html?article=st0109.

or short parental leave with child's outcomes, ρ_1 and ρ_0 in models (3.5) and (3.6). All estimates exclude children born in Quebec or live in single parent families. Moreover, estimates for children aged 13 months and older exclude mothers who did not return to work within the first 12 months of their children's lives.

3.6.2.1 Child's Health

I test the effect of parental leave on some aspects of a child's health that are more likely to be affected by the mother working, such as has the child ever been an inpatient in a hospital overnight, number of times the child had nose throat infections, ear infection, been seriously injured, and how many times been to see a doctor. The health outcomes are mostly available for all age groups of children. The estimated results are reported in Table 3.2. The table displays that parental leave has almost no contemporaneous effect, except it reduces the likelihood of having asthma, which falls by 33% for children 7 to 12 months old whose parents took a long parental leave. Some positive effects emerged for older children; in particular children of parents who took a long parental leave had ear infections two to three times less than other children did, and visited a medical doctor three to four times less. Moreover, I find the effects are more pronounced among children whose parents took a long parental leave compared with the average effects for all children. I find only one and two positive significant effects out of 12 tested health outcomes for children 7 to 12 and 13 to 24 months old, respectively. In addition, two positive significant effects out of 14 tested outcomes for children 25 to 48 months, one positive and two negative significant effects out of 11 outcomes for children 49 to 72 months old.

3.6.2.2 Child's Temperament

Temperament outcomes are parents' perception of a child's difficulty, such as how difficult the child is to be calmed or soothed, how fussy he is, how much he cries, how easily gets upset and in general what mood he is in. The outcomes are only available for children aged 0 to 35 months in the survey. Temperament outcomes are seven-measure variables with 1 as the best and 7 as the worst score. Thus, a negative estimated parameter can be interpreted as an improvement in the children's temperament. The results are summarized in Table 3.3. I find no contemporaneous effect for parental leave on a child's temperament, but more positive effects emerge as a child grows up. In which, 1 positive and 1 negative effect out of 12 tested outcomes emerged for children 13 to 24 months old, and 7 positive effects out of 12 tested outcomes emerged for children 25 to 35 months old. All positive effects are more pronounced among children whose parents took a long parental leave compared with average effects for all children. In general, one can infer that the expansion in parental leave in 2001 significantly improved the temperament of children aged two years old.

3.6.2.3 Development and Behavioural Outcomes

Development and behavioural outcomes are only available for children in the age groups 13 to 24 and 25 to 48 months. In this paper, development measures address a child's ability for performing certain tasks, such as washing hands, going to the toilet alone, walking, dressing, drawing pictures, and ability to count objects and numbers. The outcomes are used as binary dependent variables and the estimated results are reported in the left panel of Table 3.4. I find only 1 positive significant effect out of 13 tested outcomes for children aged 13 to 24 months old, and no significant effect is found for older children, 25 to 48 months old. Moreover, the most promising measures of a child's development in the NLSCY are the Motor Social Development score and the Peabody Picture Vocabulary Test (PPVT) for children aged zero to three and four to five years old, respectively. The results for these two measures are reported in Table 3.5. Since the development scores are age specific, I use the standard Motor Social Development score for children aged 7 to 47 months old and find longer parental leave has no effect on child development with respect to his Motor Social Development score. The PPVT score is the only measure in the NLSCY directly measured by interviewers. It measures the school readiness of a child and can be interpreted as a cognitive development measure. I find a long parental leave increases the PPVT score of the children four to five years old by 14 scores, which is equivalent to almost 1 standard deviation.

The behavioural outcomes in this study are binary variables indicating a child's behaviour, temper or attitude, such as getting into many fights, being defiant, easily distracted, worried, anxious and inattentive. The behavioural outcomes represent a child's stable personality in respond to stimulus from the environment. The outcomes are tested for children aged 25 to 48 months old, and the results are reported in the right panel of Table 3.4. I find 2 negative significant effects out of 14 tested outcomes, in which children of parents who took a parental leave longer than six months are 34% more likely to be defiant and 23% more likely to be nervous when aged two to three years. In addition, the NLSCY includes derived behavioural scores of the children, and their estimated results are reported in Table 3.5. The results are mixed, in which emotional and aggressive behaviours among children aged two to three years are worsened by the longer parental leave, and their hyperactivity behavioural score improves, however the adverse and the positive effects are found to be temporary and they disappear when children aged four to five years.

3.6.2.4 Milestone Achievements

The milestone achievements are used as development measures in very early childhood. They measure the age of child in months when the child succeed to perform a simple task for the first time, including sitting up, saying first word, taking first steps, eating solid food and feeding himself. The measures are available in the NLSCY for children aged 0 to 47 months. Since the outcomes are one time events over the course of life, I estimate them in a single age group 7 to 47 months, and the results are reported in the right panel of Table 3.5. I find parental leave has a mixed effect on the milestone achievements. The age at which a child starts to eat solid food is reduced by 2.8 months if the parents take a long parental leave, but the ages at which a child starts to feed himself or take his first steps are delayed by 2.7 and 1.8 months, respectively. Like the other outcomes, the TT is more pronounced than the ATE.

3.6.2.5 Literacy and Parenting

Literacy in early childhood is highly correlated with educational achievements in later life. Literacy questions in the NLSCY measure child's interest in reading and learning, and related activities that parents do with their child. The results are reported in the left panel of Table 3.5. The parents who take a longer leave is more likely to get involved in activities related to literacy, such as they teach new words to their child more often or take him out for a walk or to play on play ground when aged two to three years. Overall, I find two positive significant effects out of eight tested literacy outcomes.

Relationships between the family members of a child have a substantial and long lasting effect on the child's mental health. The family functioning score in the NLSCY measures a shortened version of the relationship score that focuses on problem solving, communication, roles, effective involvement, effective responsiveness and behavioural control among family members of a child. A higher score in this measure represents more dysfunctioning in the family. The results are reported for the three oldest age groups in Table 3.5. The table shows that a long parental leave is followed by a worsening in the family functioning for the middle age children, the score decreased by 4.6 out of scale 0 to 36 for children aged 25 to 48 months, but no effect is found for the older children. In addition, a similar pattern is found for ineffective or hostile parenting, in which the ineffective parenting score decreased by 2.6 out of scale 0 to 23 for children aged 25-48 months, but no effect is found for the older children.

3.6.2.6 Breastfeeding

The mother and her child both are beneficiaries of breastfeeding. A child benefits from breastfeeding by decreases in diarrhea, gastro-intestinal diseases, asthma, lower respiratory infections, sudden infant death syndrome, lymphoma and chronic digestive diseases. The mother benefits from breastfeeding by improved bone remineralization, reduced risk of ovarian and pre-menopausal breast cancer (American Academy of Pediatrics, 1997). The estimated effect of the policy change on breastfeeding is reported at the bottom of Table 3.5. This estimate excludes mothers who did not breastfeed at all. I find an increase in parental leave causes breastfeeding to increase by almost 4 weeks on average among all children, but increased by 25 weeks among children whose parents took a parental leave longer than six months.

For a better presentation of the estimated results reported in Tables 3.2 to 3.5, I illustrate some of selected outcomes in Figure 3.3. The figure includes four graphs, each comparing the Kernel densities of the effect of treatment on treated (TT) with the average treatment effect (ATE) for a selected outcome.

3.6.3 Parental Leave, Gender, Demographic and Socioeconomic Status of Family

This section discusses changes in the effect of parental leave by the socioeconomic status of the parents. To do this, I report TT and ATE, which are derived in a similar way for the whole population, by socioeconomic status of parents. Education and household income are used to address the socioeconomic status of a family. I dichotomize the education of parents to high-education, including some post secondary or university degree, and low-education, including high school diploma or less. I use a family's real income and the household size to identify low-income and high-income families. In this study, a family is a low-income if the family's real income is lower than \$20000, \$25000 or \$30000 for a household with 3, 4 or 5+ members, respectively. I also examine the effect of parental leave with age of the mother at birth. A dummy variable is generated to identify mothers who were teenagers at birth.

No difference is found for the effect of parental leave on a child's outcomes by age of the mother at birth and the education of parents (PMK), except for the duration of breastfeeding, where higher educated mothers breastfeed for a longer time if they take a longer parental leave. I find children of low-income families are worsened in respect to some of their behavioural outcomes, compared with those of high-income families if their parents take a long parental leave. Moreover, the gender of a child seems not to be an issue, except for temperament and behavioural outcomes, in which baby girls benefit more or loose less, respectively, from a longer parental leave. Moreover, I test the effect of parental leave by province and find no difference in the effect of parental leave across the provinces, despite the fact that changes in the duration of parental leave across the Canadian provinces, excluding Quebec, were unequal by the policy change (the results are not reported for space reserving).

Figure 3.4 illustrates the difference in the effects of parental leave by the socioeconomic status of the parents for some selected outcomes (other outcomes are not reported for space reserving).

3.7 Conclusion

Parental leave in Canada has expanded from roughly 6 to 12 months since the beginning of 2001. I use a sample of children born in a ten-year window surrounding the periods before and after the reform to examine the effect of parental leave on health and development outcomes for children aged 7 to 72 months old. I employ an endogenous switching regression model with a selection equation that indicates the mothers who returned to work within the first six months of their children's lives or stayed in home up to the first year. Using the switching regression model, I examined the effect of parental leave on the children's health, behaviours, milestone achievements, development, cognitive outcome, literacy, parenting style, family functioning and the duration of breastfeeding. In addition, I distinguished the effect of treatment on treated from the average treatment effects to recover heterogeneity in the effect of parental leave.

I found that parental leave has almost no contemporaneous effect. But, some positive and a few adverse effects emerge when the children were two to five years old. Overall, I tested 74 outcomes, of which 14 were health related, 12 temperaments, 12 developments, 13 behavioural, 8 literacy, 5 milestones, 1 breastfeeding and 9 derived scores. The majority of which are tested for four age groups: 7 to 12, 13 to 24, 25 to 48 and 49 to 72 months. I found 1 positive significant effect out of 19 tested outcomes for the 7-12 age group; 4 positive and 1 negative significant effects out of 46 tested outcomes for the 13-24 age group; 14 positive and 9 negative significant effects out of 71 tested outcomes for the 25-48 age group; and 2 positive and 2 negative significant effects out of 12 tested outcomes for the 49-72 age group.

In particular, from all tested outcomes, a child's temperament, cognitive development and breastfeeding appeared to be positively affected by parental leave, in which children of parents who took a long parental leave had better temperament outcomes, obtained 1 standard deviation higher in cognitive development measure, and were breastfed one to six months more. Moreover, I found that the effects were more pronounced among children of parents who took a long parental leave compared with the children in the whole population.

In contrast, I found that some behavioural outcomes were deteriorated for children aged two to three years old. I also found that the family functioning and ineffective parenting scores were reduced with the longer parental leave for these children. However, all negative effects appeared to be temporary, and they disappeared when the children aged four to five years old.

Overall, I found that there exists a substantial unobservable heterogeneity effect in parents' skills on child's outcomes, which cannot be examined by conventional estimate methods such as an OLS regression. Such a heterogeneity effect suggests that a targeted parental leave policy will be more effective than a universal policy.

Table 3.1: Summary statistics of the NLSCY cycles 2-7 (1996-2005) by age-group of children and the reform in parental leave in 2001

	7-12 months				13-24 months				25-48 months				49-72 months			
	Before		**After**		**Before**		**After**		**Before**		**After**		**Before**		**After**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Female	.49	.5	.48	.5	.5	.5	.49	.5	.5	.5	.49	.5	.49	.5	.49	.5
Younger siblings					.04	.2	.04	.19	.26	.44	.24	.43	.38	.49	.34	.47
Older siblings	.54	.5	.53	.5	.5	.5	.51	.5	.51	.5	.5	.5	.49	.5	.51	.5
Mother age at birth	29.7	4.9	29.2	5	29.7	4.8	29.7	4.9	29.6	4.8	29.7	4.8	29.3	4.8	29.6	4.8
Duration of leave	5.3	2.4	6.7	3.6	6	2.8	9	3.7	6.1	2.9	9	3.7	6	3	9	3.6
Leave > than 6 m.	.27	.45	.52	.5	.35	.48	.74	.44	.36	.48	.75	.44	.35	.48	.74	.44
Asthma	.03	.16	.02	.13	.05	.22	.03	.18	.09	.28	.07	.26	.12	.33	.12	.32
Wheezing	.26	.44	.23	.42	.28	.45	.26	.44	.22	.41	.22	.42	.18	.38	.21	.41
Take medication	.06	.23	.07	.25	.07	.26	.07	.26	.07	.26	.08	.28	.09	.29	.1	.3
Hospitalized	.12	.32	.09	.29	1	.3	.08	.27	.06	.24	.05	.23	.04	.2	.04	.21
Healthy	.91	.29	.94	.25	.91	.28	.92	.27	.91	.28	.93	.26	.92	.28	.93	.25
Injured	.03	.17	.03	.16	.08	.26	.07	.26	.12	.32	.1	.3	.1	.3	.1	.3
Times injured	.05	.26	.04	.21	.08	.32	.08	.31	.14	.44	.12	.37	.11	.34	.11	.36
Allergy	.04	.18	.06	.23	.08	.27	.07	.26	.09	.29	.1	.3	.13	.34	.12	.32
Nose throat infections	4.5	.79	4.5	.81	4.2	.87	4.3	.88	4.1	.83	4.1	.86	4.1	.78	4.3	.63
Ear infection	.33	.47	.27	.44	.56	.5	.45	.5	.66	.47	.58	.49	.69	.46	.73	.45
# of ear infections	1.8	2.5	.47	.97	1.5	1.9	1.1	1.9	1.6	1.8	2	3.3	1.8	1.6	2.5	2.9
Seen a MD	6.3	5.9	5.1	4.8	6.1	6.5	4.9	4.9	3.3	3.8	3	3.6	2.7	3.4	2.4	3
BMI									18.3	4.6	17.4	3.3	17.1	3.7	17.1	3.5
Obesity									.49	.5	.43	.5	.41	.49	.43	.49
Easy to calm	1.9	1.2	2	1.2	2.2	1.3	2.3	1.3	2.7	1.5	2.6	1.5				
Often get fussy	1.3	.81	1.4	.83	1.3	.81	1.5	.9	1.3	.84	1.5	.9				
Fuss intensively	2	1.2	2.1	1.1	2.2	1.2	2.2	1.2	2.4	1.3	2.3	1.3				
Get upset	2.7	1.5	2.8	1.5	3.1	1.5	3.2	1.5	3.4	1.5	3.4	1.5				
Cry vigorously	3.5	1.8	3.7	1.8	3.9	1.8	4	1.7	4	1.8	4	1.7				
Smile	1.4	.83	1.4	.75	1.5	.88	1.5	1	1.6	.97	1.5	.95				
General mood	1.5	.93	1.5	.96	1.5	.9	1.6	.9	1.6	.88	1.7	.92				
His mood changes fast	2.4	1.5	2.5	1.5	2.7	1.6	2.7	1.5	2.9	1.6	2.9	1.6				
Attention	3.5	1.7	3.4	1.6	3.4	1.6	3.3	1.5	3.3	1.6	3.2	1.5				
Respond to strangers	2.8	1.8	2.9	1.6	3.1	1.8	3.1	1.6	3.1	1.7	3.1	1.6				
Respond to new place	2.1	1.4	2.2	1.4	2.6	1.5	2.5	1.4	2.7	1.5	2.6	1.4				
Overall difficulty	2	1.2	1.9	1	2.3	1.3	2.1	1.2	2.4	1.3	2.3	1.2				
Age sat up first time	5.8	1.6	5.9	1.5												
Age started eating solid food (7-47 mth.)	6.2	2.7	6.0	2.4												
Age started feeding himself (7-47 mth.)	10.2	3.4	8.5	2.3												
Age took first steps	11.3	2.2	11.2	2.5												
Age said first words	10.5	3.8	9.5	3.1												
PPVT score (4-5 yrs)	102.9	13.9	102.0	13.4												
STD motor score (0-3 yrs)	101.5	14.6	100.2	14.2												
Number of obs.	1,300		606		4,340		2,576		5,905		3,399		6,351		1,516	

Table 3.1 (continued)

	13-24 months				25-48 months				49-72 months			
	Before		After		Before		After		Before		After	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Know his pants is wet	.62	.5	.6	.49	.85	.36	.86	.35				
Speaks a sentence	.4	.5	.35	.48	.97	.17	.94	.23				
Walk upstairs	.31	.46	.31	.46	.92	.27	.88	.33				
Wash hands	.4	.49	.42	.5	.91	.28	.9	.3				
Count 3 objects	.28	.45	.33	.47	.9	.29	.91	.29				
Has gone to toilet alone	.13	.33	.13	.34	.75	.44	.71	.46				
Knows his age and sex	.13	.34	.15	.36	.78	.42	.76	.42				
Knows name of at least 4 colours	.28	.45	.35	.48	.87	.33	.88	.32				
Ride a tricycle at least for 10 feet	.16	.37	.17	.38	.71	.45	.66	.47				
Dresses himself	.3	.46	.28	.45	.67	.47	.62	.49				
Says his first and last name together	.15	.36	.14	.34	.63	.48	.58	.49				
Counts up to 10	.21	.4	.24	.43	.69	.46	.71	.45				
Look at books	6.7	.85	4.8	.47	5.1	1	4.7	.62				
Talk about books					3.9	1.3	3.9	1.3				
Play with pencils	6.2	1.3	4.4	.88	4.9	1.1	4.5	.81				
Go to library	1.4	.78	1.4	.82	1.7	.97	1.7	.95				
Defiant	.69	.47	.75	.44	.8	.4	.81	.4				
Fight	.44	.5	.34	.48	.43	.5	.4	.49				
Easily distracted	.49	.5	.44	.5	.41	.49	.39	.49				
Anxious	.22	.41	.21	.41	.28	.45	.27	.44				
Impulsive	.72	.45	.71	.46	.66	.47	.67	.47				
Has hot temper	.72	.45	.81	.39	.71	.45	.71	.45				
Worried	.09	.29	.13	.34	.22	.42	.25	.43				
Has angry mood	.46	.5	.45	.5	.56	.5	.51	.5				
Cries a lot	.29	.46	.32	.47	.35	.48	.32	.47				
Constantly seek help	.42	.5	.41	.5	.41	.49	.39	.49				
Is nervous	.09	.28	.12	.33	.13	.33	.12	.32				
Does not sleep alone	.43	.5	.35	.48	.47	.5	.46	.5				
Inattentive	.38	.49	.36	.48	.37	.48	.36	.48				
Read story	6.5	1.5	4.7	.66	4.8	.56	4.82	.5				
Parents play action games with child	4.8	.54	4.8	.51	4.6	.77	4.7	.7				
Parents tell stories to the child	4.3	1.2	4.4	1.1	4.5	1	4.6	.86				
Parents sing song with the child	4.7	.72	4.7	.71	4.6	.8	4.6	.82				
Parents teach new words to child	4.8	.67	4.8	.58	4.7	.76	4.8	.62				
Parents take the child out	4.7	.54	4.5	.66	4.6	.57	4.5	.65				
Family dysfunctioning	8.5	4.8	8.2	5	8.5	4.9	7.9	5	8	4.9	7.8	4.9
Positive parenting	17.7	1.9	17.8	1.9	16.8	2.2	16.7	2.2	15.1	2.3	15.4	2.3
Ineffective parenting	2.6	1.4	2.7	1.6	8.9	3.3	8.5	3.2	8.5	3.3	8.2	3.2
Duration of breastfeeding in weeks (0-5 years old)	23.2	18.6	25.9	20.3								
Number of obs.	4,340		2,576		5,905		3,399		6,351		1,516	

Note: The summary statistics exclude children born in Quebec or living in single parent families. I restrict the samples to mothers who returned to work within the first 12 months of their children's lives for children aged 12 months and older. Other variables used in this study but not reported in the table are age of child in months (dummy variables), age of PMK (6 dummies), age of spouse of PMK (6 dummies), PMK's education (4 dummies), education of spouse of PMK (4 dummies), immigration status of both parents (2 dummies), household's income (9 dummies), province of residence (9 dummies), cycle of the data (6 dummies), city size (5 dummies), and child's behavioural scores.
Blank cells in the table indicate that there is no observation for that variable in the corresponding age group. Also, some variables are reported in a single age group, including breastfeeding (aged 7-72 months), PPVT score (aged 46-72 months), STD motor score (aged 7-47 months), and milestones achievements (aged 7-47 months old).

Table 3.2: Switching regression estimates on health outcomes for children aged 7-72 months old

	Age 7-12 months				Age 13-24 months				Age 25-48 months				Age 49-72 months			
	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT
Has asthma	.04	**.92**	-.01	**-.33**	.08	.96	-.00	-.22	.01	-.08	.00	.04	**.09**	.12	.01	-.07
	(.08)	**(.53)**	(1.4)	**(.20)**	(.07)	(.76)	(.01)	(.17)	(.04)	(.19)	(.01)	(.10)	**(.05)**	(.18)	(.01)	(.11)
Has had wheezing or whistling	-.02	-.39	-.02	.29	-.01	-.31	-.01	.22	-.04	.53	.01	-.19	-.01	.22	-.01	-.17
	(.10)	(.68)	(.98)	(.53)	(.05)	(.20)	(.01)	(.16)	(.06)	(.69)	(.01)	(.27)	(.06)	(.24)	(.01)	(.18)
Takes medication regularly	.07	.80	.01	-.39	-.01	-.53	.00	.18	-.07	-.42	.01	.17	.03	-.25	**.02**	.15
	(.09)	(.51)	(62)	(.26)	(.06)	(.75)	(.01)	(.24)	(.06)	(.30)	(.01)	(.11)	(.06)	(.20)	**(.01)**	(.11)
Has ever hospitalized overnight	.12	.08	.01	-.04	.02	.14	-.00	-.08	.04	-.23	.01	.09	-.02	.01	-.00	.01
	(.09)	(.59)	(6.3)	(.45)	(.05)	(.18)	(.01)	(.10)	(.06)	(.31)	(.01)	(.11)	(.06)	(.19)	(.01)	(.08)
Is healthy	-.01	.28	-.01	-.18	.01	.16	-.00	-.06	.04	.25	.01	-.08	-.00	.32	-.01	-.13
	(.08)	(.59)	(2.5)	(.72)	(.06)	(.88)	(.01)	(.31)	(.06)	(.78)	(.01)	(.28)	(.07)	(.92)	(.01)	(.37)
Has ever been Injured	.05	-.34	.00	.12	.00	.48	-.01	-.25	.04	-.09	-.00	.05	.04	-.65	-.00	.22
	(.08)	(.54)	(2.8)	(.20)	(.05)	(.23)	(.01)	(.11)	(.04)	(.19)	(.01)	(.05)	(.07)	(.75)	(.01)	(.27)
Times was injured	.08	-.91	.07	.83	.01	**.59**	-.01	**-.36**	.04	-.33	-.00	.20	.05	-.81	-.00	.32
	(.10)	(.87)	(.11)	(.90)	(.05)	**(.24)**	(.01)	**(.14)**	(.06)	(.31)	(.01)	(.20)	(.07)	(.77)	(.01)	(.33)
Has any kind of allergy	.03	.46	.00	-.14	.04	.05	-.00	-.03	-.01	-.22	.01	.12	-.00	**-.86**	.01	**.48**
	(.08)	(.70)	(.02)	(.22)	(.05)	(.18)	(.01)	(.09)	(.05)	(.19)	(.01)	(.10)	(.06)	**(.28)**	(.01)	**(.16)**
Time has had nose throat infections	-.00	-.66	-.05	.88	.11	-.44	.00	.41	.03	.09	-.03	-.17				
	(.08)	(.68)	(2.0)	(.94)	(.06)	(1.2)	(.03)	(1.2)	(.05)	(.21)	(.02)	(.31)				
Ever had ear infection	-.05	.75	.03	-.56	-.03	-.78	.00	.49	.04	-.23	.03	.22				
	(.10)	(.73)	(.03)	(.53)	(.06)	(.84)	(.01)	(.52)	(.06)	(.19)	(.01)	(.17)				
Times has had ear infections	NA	NA	NA	NA	**-.12**	**.83**	.00	**-2.7**	**-.26**	**.61**	**.11**	**-2.1**				
					(.06)	**(.34)**	(.04)	**(1.2)**	**(.06)**	**(.24)**	**(.06)**	**(.94)**				
Times has seen a MD	.07	.02	.02	-.10	**.10**	-.17	.12	2.1	**-.10**	**.54**	.01	**-3.5**	-.07	**.47**	.11	**-2.8**
	(.07)	(.67)	(.33)	(5.0)	**(.05)**	(.15)	(.14)	(1.7)	**(.06)**	**(.20)**	(.08)	**(1.4)**	(.07)	**(.28)**	(.31)	**(1.7)**
Body Mass Index									.14	-.25	.27	2.1	.08	.60	-.14	-4.3
									(.12)	(.32)	(1.2)	(4.5)	(.15)	(.47)	(.27)	(3.2)
Obesity									-.05	-.53	.02	.46	.01	-.71	.01	.28
									(.11)	(.38)	(.59)	(1.9)	(.18)	(1.4)	(.04)	(.51)

Note: NA denotes result is not reported because estimated correlation falls outside of the range (-1, 1). Blank cells indicate that there is no observation for that specific outcome in related age group. Number of observations for age groups 7-12, 13-24, 25-48 and 49-72 months old are 1900, 6900, 9200 and 7800, respectively. **Bold figures are significant at 0.10 or lower significance levels**, standard deviation is in parentheses. ρ_1 and ρ_0 are the correlations between a mother's propensity to take a long and a short leave, respectively with a child's outcome.

ATE and TT denote the average treatment effect and the effect of treatment on treated, respectively.

All estimates are robust and clustered at provinces and cycles of the data. Standard deviations were derived using a bootstrap method with 500 replications.

All estimates include following explanatory variables: gender of child, a dummy variable indicating households up to one younger sibling, a dummy variable indicating households up to two older siblings, mother's age at birth, immigration status of both parents, a dummy variable indicating the outcome is mother's report, 6 dummies indicating work intensity of the father (and 6 dummies for the mother), 10 dummies indicating type of childcare arrangement used, 3 dummies indicating highest level of education the father achieved (and 3 dummies for the mother), 8 dummies for province of residence, 5 dummies for city size and rural area, 8 dummies for real family income, 5 dummies for age group of the father (and 5 dummies for the mother), dummies for age of child in month and a time trend. All estimates exclude children born in Quebec, live in single parent families and those whose mothers did not return to work within the first twelve months of their lives, except for children in the age group 7-12 months.

Table 3.3: Switching regression estimates on temperament outcomes for children aged 7-35 months old

	Age 7-12 months				Age 13-24 months				Age 25-35 months			
	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT
How difficult is to calm or soothed	-.14	.62	.13	-1.1	-.06	**.44**	**.07**	**-1.0**	-.12	.61	.01	-1.3
	(.12)	(.82)	(.18)	(1.6)	(.06)	**(.18)**	**(.04)**	**(.46)**	(.08)	(.99)	(.06)	(2.1)
How often is fussy	NA	NA	NA	NA	**-.11**	.27	-.01	-.42	-.00	**.39**	-.01	**-.65**
					(.06)	(.25)	(.03)	(.37)	(.06)	**(.23)**	(.14)	**(.38)**
How much fusses when is upset	.02	.76	.09	-1.3	-.04	.37	**.05**	-.65	-.10	**.83**	.01	**-2.0**
	(1.0)	(.69)	(.08)	(1.2)	(.06)	(.45)	**(.03)**	(.84)	(.07)	**(.23)**	(.38)	**(.65)**
How easily gets upset	.10	.87	.10	-2.7	-.05	.22	.04	-.58	-.03	**.65**	-.00	**-1.8**
	(.09)	(.67)	(19)	(3.1)	(.05)	(.27)	(.05)	(.75)	(.07)	**(.20)**	(.29)	**(.61)**
How vigorously cries when is upset	.01	.63	-.08	-2.6	-.02	.30	-.02	-.97	.04	.10	.02	-.30
	(.09)	(.55)	(53)	(2.1)	(.05)	(.23)	(.04)	(.73)	(.07)	(.17)	(.39)	(.58)
How often smiles or make happy sounds					-.02	.08	.02	-.09	-.01	**.68**	.00	**-1.2**
					(.05)	(.26)	(.02)	(.43)	(.06)	**(.22)**	(.13)	**(.47)**
How often is in cheerful mood					**.12**	-.06	.03	.11	.03	.50	-.01	-.68
					(.06)	(.17)	(.03)	(.27)	(.06)	(.38)	(.03)	(.52)
How fast his mood changes	NA	NA	NA	NA	-.02	.06	.01	-.16	-.07	**.82**	-.01	**-2.4**
					(.06)	(.29)	(.04)	(.80)	(.07)	**(.23)**	(.36)	**(.72)**
How much attention requires other than care giving	.05	.86	.07	-3.0	.00	.46	.13	-.99	**-.15**	**.65**	.11	**-1.7**
	(.08)	(.55)	(50)	(1.9)	(.06)	(.46)	(.04)	(1.1)	**(.06)**	**(.22)**	(.72)	**(.62)**
How much unfavourably respond to new persons	**.23**	.07	-.07	-.27	-.02	.85	.02	-1.3	.06	-.15	.00	.41
	(.11)	(.68)	(.22)	(1.4)	(.07)	(.92)	(.04)	(1.4)	(.08)	(.24)	(.56)	(.72)
How much unfavourably respond to new places	.23	-.13	.02	.30	.03	**-.42**	.10	**1.2**	-.01	-.11	.01	.18
	(.12)	(.54)	(286)	(1.5)	(.05)	**(.22)**	(.07)	**(.63)**	(.09)	(1.0)	(.05)	(1.3)
Overall, how difficult child is	.03	-.07	.11	.22	**.12**	.01	.05	.01	-.06	**.67**	.07	**-1.5**
	(.09)	(.66)	(.10)	(1.1)	**(.05)**	(.25)	(.03)	(.57)	(.06)	**(.22)**	(.27)	**(.55)**

Note: NA denotes that results are not reported because estimated correlations fall outside of the range (-1, 1). Blank cells indicate that there is no observation for that specific outcome in related age group. Number of observations for age groups 7-12, 13-24 and 25-35 are 1900, 6900 and 5100, respectively. **Bold figures are significant at 0.10 or lower significance levels,** standard deviation is in parentheses. ρ_1 and ρ_0 are the correlations between a mother's propensity for taking a long and a short leave with her child's outcome, respectively. ATE and TT denote the average treatment effect and the effect of treatment on treated, respectively.

All estimates are robust and clustered at provinces and cycles of the data. Standard deviations were derived using a bootstrap method with 500 replications.

All estimates include following explanatory variables: gender of child, a dummy variable indicating households up to one younger sibling, a dummy variable indicating households up to two older siblings, mother's age at birth, immigration status of both parents, a dummy variable indicating the outcome is mother's report, 6 dummies indicating work intensity of the father (and 6 dummies for the mother), 10 dummies indicating type of childcare arrangement used, 3 dummies indicating highest level of education the father achieved (and 3 dummies for the mother), 8 dummies for province of residence, 5 dummies for city size and rural area, 8 dummies for real family income, 5 dummies for age group of the father (and 5 dummies for the mother), dummies for age of child in month and a time trend. All estimates exclude children born in Quebec, or living in single parent families and those whose mothers did not return to work within the first twelve months of their lives, except for children in the age group 7-12 months.

Table 3.4: Switching regression estimates on development and behavioural outcomes for children aged 13-48 months old

	Age 13-24 months				Age 25-48 months					Age 25-48 months			
Development	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT	**Behaviour**	ρ_1	ρ_0	ATE	TT
Let others know that he is wearing wet pants	.02	-.06	-.00	.04	-.08	-.96	.00	.34	Reacts to new food	-.05	-.24	.02	.16
	(.04)	(.18)	(.01)	(.15)	(.06)	(.81)	(.01)	(.28)		(.05)	(.37)	(.01)	(.21)
Has ever spoken partial of a sentence	.04	-.06	.02	.05	.03	-.14	.02	.07	Gets into many fights	.07	-.32	-.00	.26
	(.05)	(.22)	(.01)	(.18)	(.06)	(.24)	(.00)	(.09)		(.06)	(.24)	(.01)	(.22)
Has ever Walked upstairs without help	.02	-.50	-.01	.21	.07	-.15	-.01	.06	Is defiant	.06	**-.44**	**.02**	**.34**
	(.07)	(1.1)	(.01)	(.50)	(.05)	(.20)	(.01)	(.10)		(.06)	**(.19)**	**(.01)**	**(.14)**
Has ever washed his hands without help	-.01	-.61	-.02	.33	.07	-.26	.01	.15	Is easily distracted	-.01	-.09	.01	.06
	(.07)	(1.1)	(.02)	(.64)	(.05)	(.21)	(.01)	(.11)		(.06)	(.85)	(.01)	(.40)
Has ever counted 3 objects correctly	**-.12**	**-.44**	.01	**.35**	.01	.09	.01	-.04	Is anxious	.03	-.10	.02	.10
	(.06)	**(.27)**	(.01)	**(.21)**	(.05)	(.16)	(.01)	(.08)		(.05)	(.23)	(.01)	(.19)
Has ever gone to toilet alone	.03	.02	-.01	-.03	-.01	-.11	-.00	.08	Has hot temper	-.03	.06	.01	-.01
	(.06)	(.30)	(.01)	(.18)	(.05)	(.19)	(.01)	(.13)		(.07)	(.82)	(.01)	(.37)
Knows his own age and sex	-.09	.04	-.01	-.03	-.04	.25	.02	-.07	Has angry mood	.00	-.17	.01	.14
	(.07)	(.37)	(.81)	(3.4)	(.07)	(.86)	(.01)	(.33)		(.06)	(.41)	(.01)	(.31)
Knows name of at least four colours	-.08	-.49	.02	.40	.04	-.05	.00	.04	Is worried	-.03	.04	-.00	-.03
	(.08)	(.41)	(2.5)	(10)	(.05)	(.21)	(.01)	(.12)		(.05)	(.23)	(.01)	(.17)
Has ever ridden a tricycle at least for 10 feet	-.05	.14	-.04	-.14	.09	-.08	-.01	.04	Cries a lot	.02	-.24	.02	.22
	(.07)	(.39)	(1.7)	(7.4)	(.06)	(.19)	(.01)	(.15)		(.06)	(.24)	(.01)	(.21)
Has ever dressed himself without help	.05	-.19	-.01	.15	.08	-.07	-.01	.05	Is nervous	.03	**-.37**	**.02**	**.23**
	(.12)	(.50)	(6.7)	(25)	(.05)	(.22)	(.01)	(.17)		(.05)	**(.14)**	**(.01)**	**(.08)**
Has ever said his first and last name together	-.05	**-.75**	.00	.47	-.01	.18	.02	-.12	Is impulsive	-.04	-.19	.02	.19
	(.10)	**(.42)**	(6.2)	(23)	(.06)	(.21)	(.01)	(.17)		(.06)	(.23)	(.01)	(.19)
Has ever counted out loud up to 10	.05	-.34	.01	.13	.04	.04	.01	-.08	Is inattentive	.02	-.27	.00	.24
	(.13)	(1.2)	(.02)	(.42)	(.12)	(.16)	(.01)	(.12)		(.05)	(.19)	(.01)	(.17)
Has ever drawn picture of a human body	.14	-.02	.00	.00	-.08	.60	**-.03**	-.27	Constantly seeks help	-.09	-.87	.02	.44
	(.17)	(1.5)	(.01)	(.22)	(.07)	(.66)	**(.01)**	(.26)		(.07)	(.88)	(.01)	(.42)
									Does not want to sleep alone	-.07	-.11	.00	.10
										(.05)	(.23)	.01	(.20)

Note: Number of observations for age groups 13-24 and 25-48 are 6900 and 8700, respectively. **Bold figures are significant at 0.10 or lower significance levels**, standard deviation is in parentheses. ρ_1 and ρ_0 are the correlations between a mother's propensity to take a long and a short leave with her child's outcome, respectively. ATE and TT denote the average treatment effect and the effect of treatment on treated, respectively.

All estimates are robust and clustered at provinces and cycles of the data. Standard deviations were derived using a bootstrap method with 500 replications.

All estimates include following explanatory variables: gender of child, a dummy variable indicating households up to one younger sibling, a dummy variable indicating households up to two older siblings, mother's age at birth, immigration status of both parents, a dummy variable indicating the outcome is mother's report, 6 dummies indicating work intensity of the father (and 6 dummies for the mother), 10 dummies indicating type of childcare arrangement used, 3 dummies indicating highest level of education the father achieved (and 3 dummies for the mother), 8 dummies for province of residence, 5 dummies for city size and rural area, 8 dummies for real family income, 5 dummies for age group of the father (and 5 dummies for the mother), dummies for age of child in month and a time trend. All estimates exclude children born in Quebec, live in single parent families and those whose mothers did not return to work within the first twelve months of their lives.

Table 3.5: Switching regression estimates on literacy and milestone achievement outcomes, child behavioural and development scores, and family functioning and parenting scores for children aged 13-72 months old

	Age 13-24 months				Age 25-48 months			
Literacy	ρ_1	ρ_0	ATE	TT	ρ_1	ρ_0	ATE	TT
How often parents play action	.01	.47	.03	-.24	**.12**	-.65	.00	.60
games with child	(.08)	(1.3)	(.03)	(.77)	**(.07)**	(1.0)	(.03)	(.96)
How often parents tell stories	-.01	.54	-.03	-.83	**.12**	.59	.05	-.43
to the child	(.54)	(.52)	(.06)	(.29)	**(.08)**	(.88)	(.03)	(.68)
How often parents read	.25	.47	.05	-.99	.06	.36	**.04**	-.15
stories to the child	(.14)	(.64)	(.04)	(1.4)	(.09)	(1.2)	**(.02)**	(.64)
How often parents sing songs	.04	.21	.01	-.26	.03	-.20	.00	-.01
with the child	(.11)	(.30)	(.03)	(.38)	(.07)	(.82)	(.03)	(.54)
How often parents teach new	.10	-.08	.02	.11	-.11	**-.95**	-.02	**.98**
words to the child	(.08)	(.29)	(.04)	(.33)	(.09)	**(.06)**	(.03)	**(.54)**
How often parents take out	**.27**	-.35	-.02	.31	**.23**	**-.53**	.02	**.57**
the child	**(.08)**	(.35)	(.03)	(.35)	**(.07)**	**(.25)**	(.02)	**(.26)**
How often the child talks					-.08	.76	.03	-.91
about books					(.10)	(1.2)	(.05)	(1.4)
How often the child goes to					.07	-.31	-.01	.48
The library					(.05)	(.36)	(.03)	(.60)
Parenting Scores								
Family dysfunctioning	-.01	-.29	.23	2.7	**.12**	**-.52**	.07	**4.6**
	(.04)	(.28)	(.16)	(2.5)	**(.05)**	**(.28)**	(.11)	**(2.5)**
Positive Parenting	-.01	-.79	.04	1.9	NA	NA	NA	NA
	(.07)	(.84)	(.08)	(2.1)				
Ineffective parenting	-.04	-.31	.07	.86	.03	-.40	**.21**	**2.6**
	(.06)	(.23)	(.05)	(.59)	(.05)	(.25)	**(.10)**	**(1.5)**
Behavioural Scores								
Hyper activity					-.06	**.47**	.03	**-1.9**
					(.05)	**(.25)**	(.06)	**(1.1)**
Emotional disorder					.07	**-.45**	**.05**	**1.2**
					(.06)	**(.27)**	**(.03)**	**(.71)**
Separation anxiety					-.04	-.89	.06	2.2
					(.06)	(.98)	(.04)	(2.4)
Aggression					.00	**-.78**	.10	**3.3**
					(.07)	**(.43)**	(.07)	**(1.8)**

	Age 25-48 months			
Milestone Achievement	ρ_1	ρ_0	ATE	TT
Age sat up for first time	**.17**	-.44	.08	.66
(N=11,729)	**(.06)**	(.55)	(.03)	(.73)
Age started eating solid food	-.05	**.73**	**-.12**	**-2.8**
(N=11,984)	(.04)	**(.37)**	**(.05)**	**(1.4)**
Age fed himself for the first time	**.18**	**-.53**	.05	**2.7**
(N=11,732)	**(.05)**	**(.17)**	(.06)	**(.85)**
Age took first steps	**.13**	**-.42**	**.10**	**1.8**
(N=11,619)	**(.04)**	**(.21)**	**(.05)**	**(.87)**
Age said first words	**.15**	.12	-.13	-.90
(N=11,149)	**(.06)**	(.22)	(.09)	(1.3)
Development Scores				
Motor Social Development score	-.04	-.06	.24	1.8
(age 7-47 months, N=17258)	(.04)	(.16)	(.30)	(4.0)
PPVT_R score	**.13**	**-.52**	**.58**	**14**
(age 46-72 months, N=7028)	**(.07)**	**(.30)**	**(.33)**	**(7.9)**
Parenting Scores	Age 49-72 month			
Family dysfunctioning	-.09	.33	.02	-2.7
	(.06)	(.55)	(.14)	(4.5)
Positive Parenting	.03	-.52	.12	2.1
	(.07)	(.46)	(.06)	(1.8)
Ineffective parenting	.01	.96	-.03	-3.8
	(.07)	(1.3)	(.08)	(5.1)
Behavioural Scores				
Hyper activity	.03	.16	**-.11**	-.60
	(.07)	(.92)	**(.06)**	(2.8)
Aggression	.07	-.01	**-.11**	-.08
	(.06)	(.26)	**(.05)**	(.93)
Breastfeeding				
Duration of breastfeeding in weeks	.08	**-.73**	**3.8**	**25**
(age 7-72 months, N=12298)	(.05)	**(.27)**	**(.45)**	**(8.2)**

Note: NA denotes results are not reported because estimated correlations fall outside of the range (-1, 1). Blank cells indicate that there is no observation for that specific outcome in related age group. Number of observations for age groups 13-24, 25-48 and 49-72 are 6900, 9200 and 7800, respectively, otherwise reported inside the table. **Bold figures are significant at 0.10 or lower significance levels**, standard deviation is in parentheses. ρ_1 and ρ_0 are the correlations between a mother's propensity to take a long and a short leave with her child's outcome, respectively. ATE and TT denote the average treatment effect and the effect of treatment on treated, respectively.
All estimates are robust and clustered at provinces and cycles of the data. Standard deviations were derived using a bootstrap method with 500 replications.
All estimates include following explanatory variables: gender of child, a dummy variable indicating households up to one younger sibling, a dummy variable indicating households up to two older siblings, mother's age at birth, immigration status of both parents, a dummy variable indicating the outcome is mother's report, 6 dummies indicating work intensity of the father (and 6 dummies for the mother), 10 dummies indicating type of childcare arrangement used, 3 dummies indicating highest level of education the father achieved (and 3 dummies for the mother), 8 dummies for province of residence, 5 dummies for city size and rural area, 8 dummies for real family income, 5 dummies for age group of the father (and 5 dummies for the mother), dummies for age of child in month and a time trend. All estimates exclude children born in Quebec, live in single parent families and those whose mothers did not return to work within the first twelve months of their lives.

Figure 3.1: Parental leave in Canada by the policy change

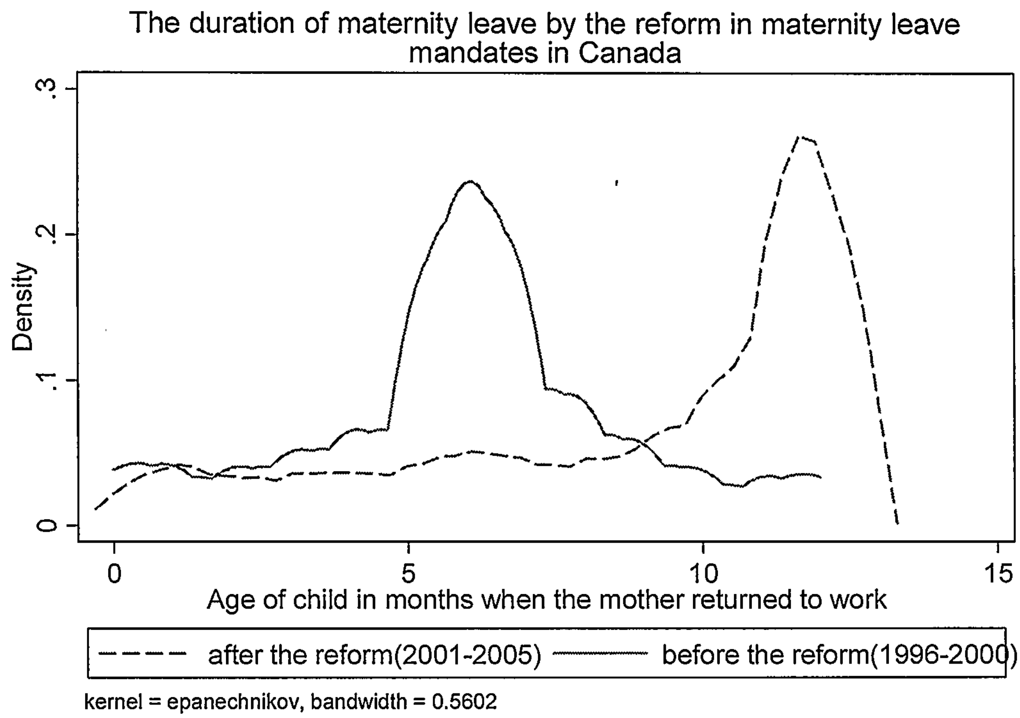


Figure 3.2: Normality assumption test using a semi-parametric estimate

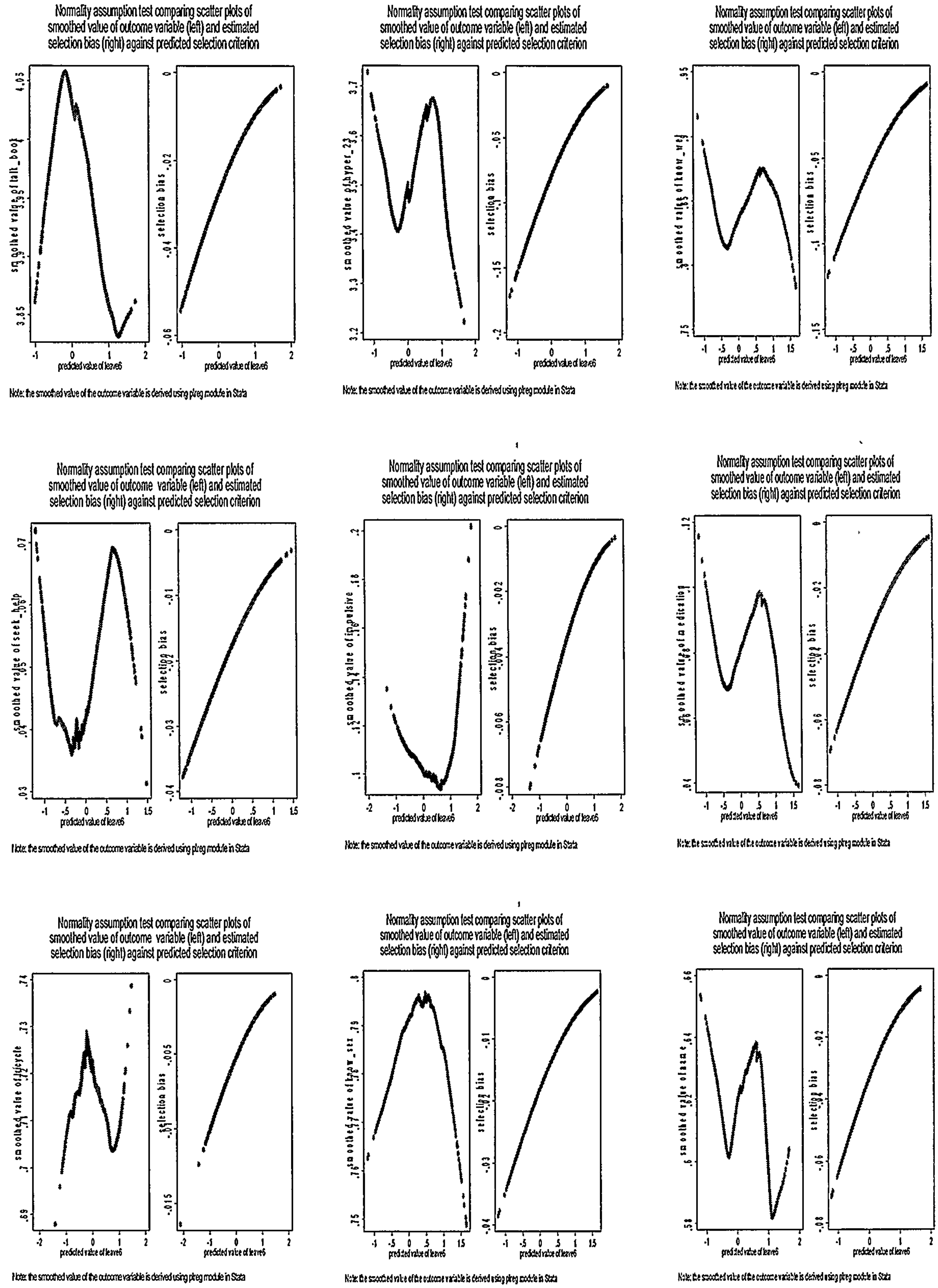


Figure 3.3: Treatment effect on treated (TT) and the average treatment effect (ATE) for some selected outcomes

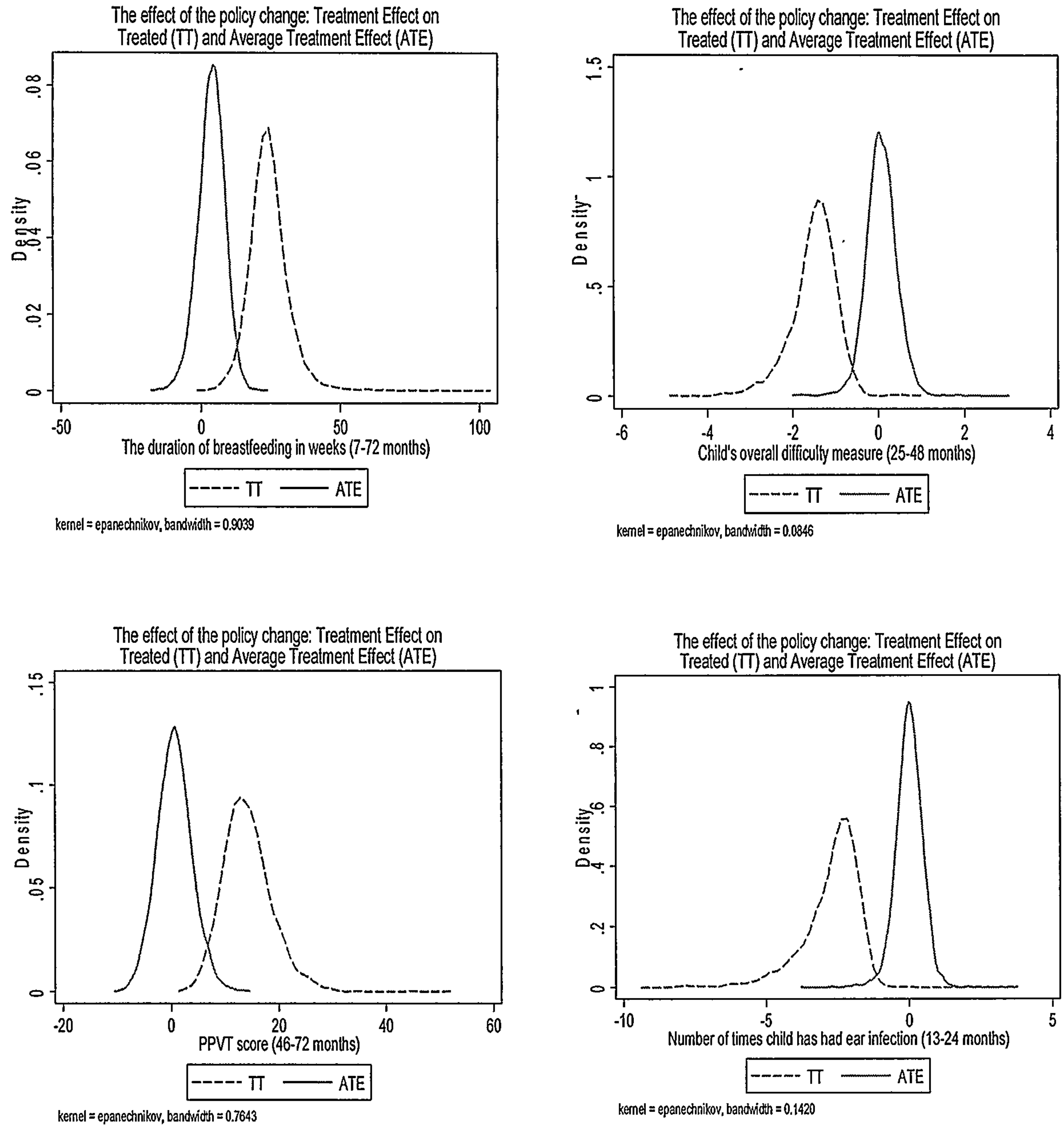
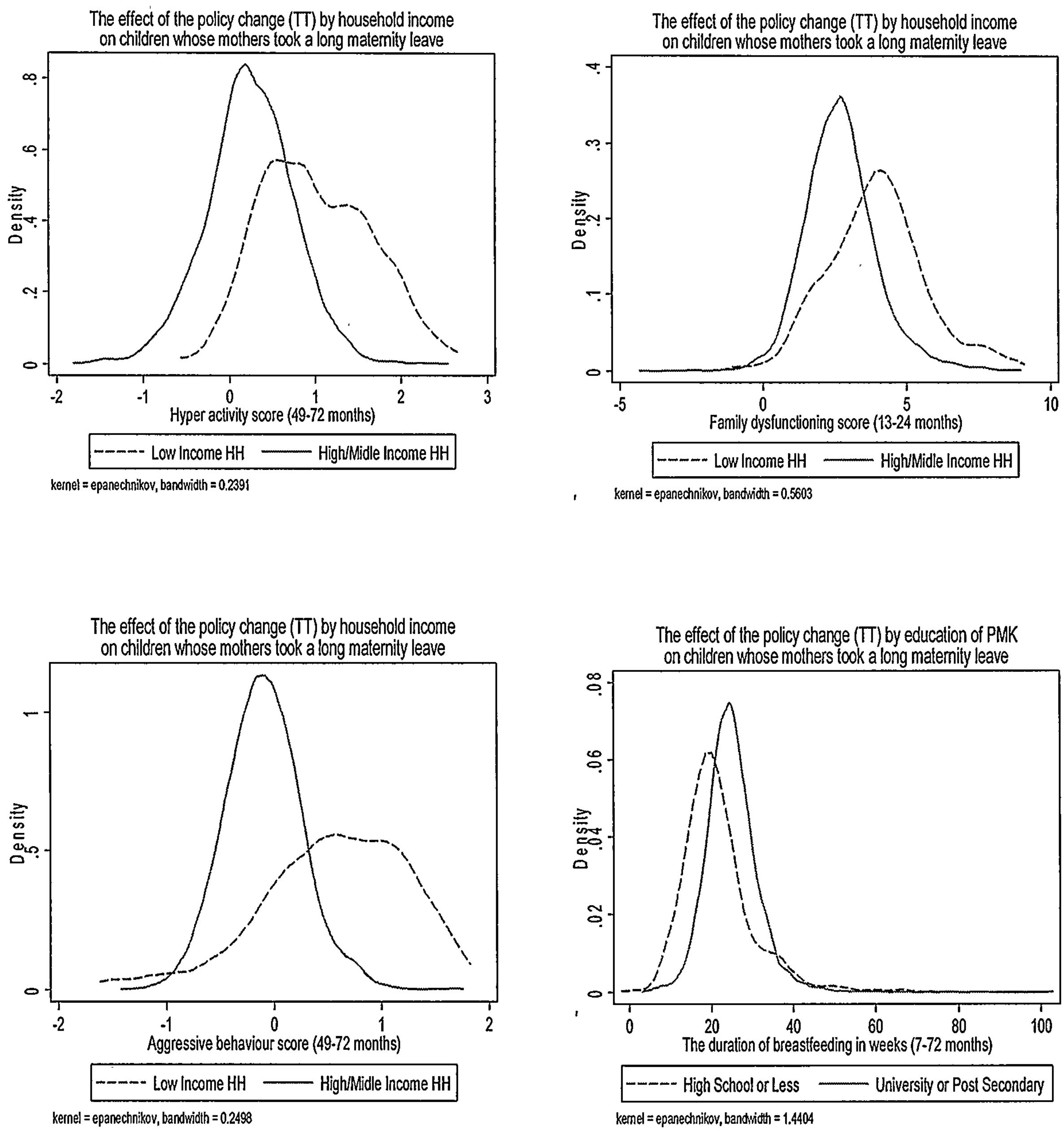


Figure 3.4: Treatment effect on treated (TT) by household's income and parent's education for some selected outcomes



Appendix 3.A: Description of Variables

Table 3.6: Description of variables used in this study

Variable	Description	Variable	Description
Health:		**Development:**	
Wheezing*	Child has had wheezing or whistling in past 12 months	Know his pants is wet*	Child lets someone know, w/o crying, he is wearing wet pants
Take medication*	Child takes medication regularly	Speak a sentence*	Child has ever spoken partial of a sentence of three words and more.
Hospitalized*	Child ever been inpatient in a hospital overnight in past 12 m.	Walk upstairs*	Child has ever walked upstairs without help
Healthy*	Child is healthy	Wash hands*	Child has ever washed his hand without help
Injured*	Child has ever been seriously injured in past 12 months	Count 3 objects*	Child has ever counted three objects correctly
Times injured	Number of times child has been injured in past 12 months	Use toilet alone*	Child has ever gone to the toilet alone
Allergy*	Child has Allergy	Know his age*	Child knows his own age and sex
Nose throat infection	# of times child has had nose throat infections in past 12 mth	Know colours*	Child has ever said name of at least four colours
Ear infection*	Child has ever had an ear infection since birth	Does tricycle*	Child ever pedaled a tricycle at least 10 feet.
Seen a MD	# of times child has seen a medical doctor in past 12 m.	Dress himself*	Child ever dressed himself w/o any help except for tying shoes
Temperament:		Say his name*	Says his first and last name together
Easy to calm	Child is easy to calm	Count up to 10*	Child ever counted loud out up to 10
Often get fussy	How often child gets fussy	Draw*	Child ever drawn at least two parts of picture of man or woman
Fuss intensively	How intensively child is fussy when he is upset	STD motor	STD score - Motor Social Development
Get upset	How easily child gets upset	PPVT	Peabody Picture Vocabulary Test
Cry vigorously	How much child cries when is upset	**Behavior:**	
Smile	How much child smiles	Cry a lot*	Child often cries a lot
General mood	General mood of child	difficult to feed*	Child is difficult to feed
His mood changes fast	How fast child's mood changes	Defiant*	Child is defiant
Attention	How much attention other than care giving child requires	Easily distracted*	Child often is distracted
Respond to new persons	Child responds to new persons unfavourably	Anxious*	Child often is anxious
Respond to new places	Child responds to new places unfavourably	Hot temper*	Child often has hot temper
Milestone Achievement:		Angry mood*	Child has angry mood
Age sat up	Age when child first sat up	Worried*	Child often is worried
Age ate solid food	Age when child started eating solid food	React to new food*	Child reacts to new foods
Age fed himself	Age when child started feeding himself	Nervous*	Child often is nervous
Age first steps	Age when child took first steps	Impulsive*	Child often is impulsive
Age first words	Age when child said first words	Fight*	Child gets into many fights

Table 3.6 (continued)

Variable	Description	Variable	Description
Breastfeed:		Inattentive*	Child often is inattentive
Breastfeeding	# of weeks that the child was breastfed	Seek help*	Child often constantly seeks help
Literacy:		Sleep alone*	Child doesn't want to sleep alone
Play game	How often parents play action games with child.	**Behavioral Score:**	
Read story	How often parents read stories or show pictures to the child.	Hyper activity	Hyper activity score
Tel story	How often parents tell story to the child.	Emotional disorder	Emotional disorder score
Sing song	How often parents sing song with the child.	Separation anxiety	Separation anxiety score
Teach words	How often parents teach new words to the child.	Aggression	Aggression score
Take out	How often parents take the child outside for a walk or to play on playground.	**Parenting Score:**	
Look at books	How often the child looks at books or magazines	Family functioning	Family functioning
Talk about books	How often the child talks about books that he has been read to	Positive Parenting	Positive Parenting
Play with pencils	How often the child plays with pencils	Ineffective Parenting	Ineffective Parenting
Go to library	How often the child goes to library		

Symbol * denotes that the variable is a binary

Appendix 3.B: Test of Normality Assumption

Table 3.7: Test of normality assumption using joint significance levels of the effects Mills ratio and the quadratic and cubic Mills ratios on the child's outcomes

	Children Aged 7-12 Months						Children Aged 13-24 Months						Children Aged 25-48 Months					
	Model (2-5)			Model (2-6)			Model (2-5)			Model (2-6)			Model (2-5)			Model (2-6)		
	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$
Asthma																		
Wheezing													**	***	**			
Take medication												*						
Hospitalized												**				*	**	**
Healthy																		
Injured										*		***				**	***	**
Times injured										**	*	***				*	**	**
Allergy	**	**	**															
Nose throat infections			***						**									
Ear infection																		*
Ever had ear infection							**		**									
Seen a MD	***	***	***				**											
BMI													**	**	*			
Obesity																		
Easy to calm																		
Often get fussy				**	**	***				**	*					**	*	*
Fuss intensively				*														
Get upset											*							
Cry vigorously																		
Smile																		
General mood	*	*						*										
His mood changes fast																		
Attention							*											
Respond to strangers																		
Respond to new places														*				
Overall difficulty																		
Age first sat up									***									
Age started eating solid food																		
Age started feeding himself							***	*		***		**						
Age took first steps												***						
Age said first words									**									
Know his pants is wet																		
Speaks a sentence										**	**	**						
Walk upstairs										**	*			*				
Wash hands							**	**	**									
Count 3 objects																		
Has gone to toilet alone																		

Table 3.7 (continued)

	Children Aged 13-24 Months						Children Aged 25-48 Months						Children Aged 49-72 Months					
	Model (5)			Model (6)			Model (5)			Model (6)			Model (5)			Model (6)		
	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$	λ_1	λ_1^2	$\lambda_1^2+\lambda_1^3$	λ_0	λ_0^2	$\lambda_0^2+\lambda_0^3$
Knows his age and sex							**	*										
Knows name of 4 colours	*	*																
Ride a tricycle	***	***	**															
Dresses himself																		
Says his first and last name							*											
Counts up to 10																		
Look at books																		
Talk about books																		
Play with pencils																		
Go to library											*							
Defiant							**	**	**									
Fight																		
Easily distracted																		
Anxious																		
Impulsive								*										
Has hot temper																		
Worried																		
Has angry mood																		
Cries a lot							**	*	*									
Constantly seek help							***	***	***									
Is nervous																		
Does not sleep alone																		
Inattentive																		
Read story																		
Play action game with child																		
Tel story to the child																		
Sing song with the child																		
Teach new words to child																		
Take the child out																		
Family dysfunctioning																		
Positive parenting																***	***	**
Ineffective parenting	*	*											**	**	*	*	**	
Duration of breastfeeding																		
PPVT score																		
STD motor score																		

Models (2-5) and (2-6) estimates the effect of Mills ratio, quadratic and cubic Mills ratios, using an OLS regression in two different regimes separated by a long and a short maternity leave, respectively. λ_1 and λ_0 are Mills ratios in corresponding models. Symbols *, **, *** denote the coefficient of Mills ratio, quadratic and cubic ratio are significant at 10%, 5% and 1% significance level, respectively (some outcomes are excluded for space reserving).

Chapter 4

Short and long term effects of income on health expenditures in Canada: Wagner's Law revisited

4.1 Introduction

The proportionate share of health expenditure (HE), as a percentage of national income (GDP), has been increasing for several decades in almost all developed countries. The extent of the increase, however, varies across countries. For instance, in the Unites States for the period 1960-2003, it increased by 10%, from 4.8% to 14.8%, while in Canada for the period 1960-2006, it increased by 5%, from 5.5% to 10.5%. The increases are major policy concerns as they raise the issue of cost contamination in healthcare services. In other words, a publicly financed healthcare system reduces the net price for consumers, thereby increases demand for healthcare services beyond the optimal amount. This phenomenon has been at the core of an enormous number of studies that have attempted to explain the reasons behind the increases in health expenditure. Wagner's Law predicts that health is a luxury good and so health expenditure as a percentage of national income is predicted to grow over time. In the context of public finance, Wagner's Law inherently suggests that government expenditures (health expenditures) tend to increase more than proportionately as the economy expands but declines less than proportionately when the economy slows. In this study, I examine Wagner's Law using two specification methods. First, using a dynamic panel, I test the long-term income elasticity of health expenditures in Canada for the period 1975-2006. Second, using an error correction estimate model, I test for the relationship between economic growth and the growth rate of health expenditure in periods of economic strength and weakness. I test the effect of income by subcategories of health expenditure, including by source of finance and use of funds, to find which parts of the expenditure drive the estimates.

The literature suggests several reasons for the increases in health expenditure over the last century, including health being a luxury good with an income elasticity greater than 1[12], the spread of democracy and women franchise for voting (Aidt *et al.*, 2006), advances in technology (Rabinovich *et al.*, 2007; Fogel, 2008), shifts in healthcare finances from individual budgets to insurance companies (Getzen, 2000), demographic changes (Zweifel *et al.*, 1999; Seshamani and Gray, 2004a, 2004b) and Baumol's unbalanced growth approach (Hartwing, 2008).

To the best of my knowledge, no effort has been taken to examine the relationship between economic growth and the growth rate of health expenditure in periods of economic weakness and strength. Wagner's Law suggests an asymmetrically-positive response of health expenditure to economic growth, which is another expression for saying health is a luxury good. However, asymmetrically-inverse or no-relation, on the other hand, between health expenditure and economic growth, is also an expectable outcome. An asymmetrically-inverse response of health expenditure to economic growth would contradict the sprit of Wagner's Law and support a 'fiscal stimulus' argument that suggests government expenditure exogenously and counter-cyclically changes over time in order to stabilize the economy, while based on Wagner's Law growth of health expenditure is an endogenous response to changes in economic growth. However, the no-

---

[12] Parkin *et al.*, (1987); Gerthham *et al.*, (1992); Hitris and Posnett, (1992); Hansen and King, (1996); Bolmqvist and Carter, (1997); Di Matteo and Di Matteo, (1998) and Freeman, (2003); Greger and Reimers (2005)

relation argument suggests that public expenditure programs are set by politicians to deliberately target those expenditures on the projected long-term growth path of the economy and that, once set, these budget become too sticky to short-term economic fluctuations and/or government revenues (Wahab, 2004). The latter outcome is known as "budget stickiness" argument, which suggests that health expenditure behaviour over time is likely to be influenced by factors affecting the steady state of the economy, such as demographic changes and technological advances.

In this study, I employ a panel of the ten Canadian provinces to test for stationarity and cointegration between health expenditures and income, controlling for cross dependence and serial correlation between the time series in the panel. Previous studies that used panel of Canadian provinces did not attempt to control for cross dependence between the units of the panel when testing for stationairty and cointegration between the time series. Lack of control on cross dependence can result in a severe size distortion in estimates, and could even result in a spurious regression if the null of no cointegration hypothesis is falsely rejected. I examine unit roots and cointegration between income and health expenditure by subcategories, including by sources of finance and use of funds.

I find health expenditure and income variables are non-stationary and not cointegrated, but are cointegrated when controlled for the lagged health expenditure. Moreover, I find health is not a luxury good with an income elasticity between 0.47 and 0.61 when controlled for other covariates, including the proportion of the elderly population, a time trend, and the relative price of healthcare services to overall prices in the economy, but has an income elasticity greater than 1 when there was no control for the covariates. Examining health expenditure by subcategories, I find hospital and physician expenditures, which account for 50% of total health expenditure, are more income sensitive than the other types of expenditures, including capital and public health expenditures. The long-term income elasticity of hospital expenditure is between 1.32 and 0.90 and for physician expenditure is between 1.54 and 1.19, which vary with model specification.

If there is an asymmetry in the effect of income on health expenditure, then the dynamic panel specification provides limited information as it pools the two states of health expenditure responses to economic growth. To overcome this short coming, I use an error correction model (ECM) to test for the asymmetry effect of income. Using this method, I find that there is no-relation between the economic growth and the growth rate of health expenditure in periods of economic strength and weakness. This finding contradicts Wagner's Law and suggests that health is not a luxury good. The results stand even when different estimation methods are employed, i.e. fixed effects estimate versus a GMM estimate approach. A no-relation between the economic growth and the growth rate of health expenditure suggests that health expenditures are likely set by policy makers on the projected long-term growth path of the economy and does not change by economic fluctuations. Moreover, such an effect suggests that an ageing population and technological advances will have stronger effect on the growth of health expenditure over time than the growth rate of the economy.

4.2 Previous Studies

The existence of insurance financing for healthcare costs has diminished the role of individual budget constraints in determining household healthcare expenditures. The diminished effect is more pronounced in countries that have adopted publicly financed healthcare systems, such as Canada. Nonetheless, the population still faces an aggregate income constraint. Thus, the literature has employed aggregate data and pooled cross-section or time series methods to test income elasticity of healthcare expenditures. However, as Hansen and King (1996) point out, the time series approach raises statistical and methodological issues not previously relevant to the cross-sectional studies. For example, if healthcare expenditures and income both contain a unit root and are not cointegrated, then the problem of spurious regressions arises. In other words, if health expenditure and income are cointegrated, then there is a linear relationship between the time series variables in the long-term, so called equilibrium, which can be estimated using a conventional statistical method, such as an OLS regression if the other classical assumptions of regression are held. The cointegration relationship between time series variables implies that the variables cannot move independently of each other. However, if time series variables are not cointegrated, then the estimated relationship between the variables is subject to a random shock and does not represent their long-term relationship.

Many studies have tried to test properties of demand for health in order to determine whether health is a luxury good[13]. If health is a luxury good, then one should expect healthcare expenditure to increase by more than one dollar for each dollar increase in income. Given the fact that income increases over time, it is predicted that healthcare expenditure will grow continuously over time. The literature has employed a variety of econometric techniques to test the income elasticity of healthcare expenditure. Newhouse (1977) uses time series data within countries and finds income elasticity is around unity,

[13] Newhouse, (1977); Culyer, (1988); Gerdtham *et al.*, (1992); Gbesemate and Gedtham, (1992); Freeman, (2003); Greger and Reimers (2005); Wang and Rettenmaier, (2007); Tosetti and Moscone, (2007)

but he argues that in the presence of a time trend variable the elasticity is expected to be lower than unity. Culyer (1988) argues that the estimate of income elasticity using individual level data tends to be lower than an elasticity that is derived using aggregate level data. Other studies find elasticity in studies that use pooled cross-section or time series data of OECD countries tend to change with the conversion factor (Purchasing Power Parity (PPP) versus exchange rate, and with the functional form such as linear, semi-log or exponential) (Newhouse, 1977; Parkin *et al.*, 1987; Gerdtham *et al.*, 1992). Di Matteo (2003) uses parametric and nonparametric econometric techniques for the Unites States, Canada, and 16 OECD countries and finds income elasticities vary with income level and with the level of analysis, in which the elasticity in international studies is higher than the elasticity in national or regional studies.

Individual time series studies that use aggregate level data usually suffer from the small sample size problem. Recent research has employed panel data to obtain higher statistical power. Moreover, panel data is more attractive to researchers because it enables them to take into account the heterogeneity effect of panel's units using fixed or random effects estimates. Gerdtham *et al.* (1992) use static and dynamic panels for 22 OECD countries for the period 1972-1987 and find the magnitude of income elasticity is sensitive to inclusion of time-period effects and whether these were treated as fixed or random variables. Income elasticity in their findings changed from 0.18 to 0.74. Hitiris and Posnet (1992) use a panel for 20 OECD countries for the period 1960-1987. They allow the intercept to differ across groups of countries and assume error terms are cross-sectionally heteroscedastic but serially correlated, and find income elasticity is around unity. Roberts (1999) considers heterogeneous relationships across countries using a panel for 20 OECD countries for the period 1960-1993. She uses four approaches, including heterogeneous fixed country effects, mean group (country) estimator, time series with the data averaged across countries and cross-sections with the data averaged over time. She finds the long-term and short-term income elasticity differ considerably by model specification (dynamic versus static and mean group versus pooled time series). In addition, she finds the long-term mean group elasticities are extremely sensitive to exclusion of a time trend. Di Matteo and Di Matteo (1998) use panel data of the ten Canadian provinces for the period 1965-1991, a pooled time-series cross-section regression and find income elasticity is less than unity (0.77).

One of the most significant advances in empirical studies over the last two decades was the identification of spurious results from regressions involving non-stationary variables (Phillips, 1986; Engle and Granger, 1987). Recent studies have tried to test stationarity and cointegration between health expenditures and income using econometric techniques, which vary from a simple OLS regression to homogenous or heterogeneous panels with/without serial and cross-sectional dependence. Hansen and King (1996) use data for 20 OECD countries and find HE and GDP are individually non-stationary but non-cointegrated. Blomqvist and Carter (1997) use data for 18 OECD countries for the period 1960-1991 and find cointegration test results based on residuals in static and dynamic models are different from that of Hansen and King (1996) when a Phillips or Perron (1989) test is used instead of an augmented Dickey-Fuller test. In addition, Roberts (1999, 2000) and McCoskey and Selden (1998) test for unit root and cointegration; they find the results are sensitive to test type, that is an augmented Dickey-

Fuller test versus Johansen (1991) test or country by country versus panel data tests of Levin and Lin (1993) or Im *et al.* (2003), and inclusion versus exclusion of a time trend.

Until recently, studies using panel data did not attempt to consider statistical implications in the panel data, mainly due to a lack of theoretical support. However, recently, new advanced techniques have been introduced for unit root and cointegration tests in panels under a complex structure of errors. Unit root tests in panels encompass more complicated situations, as panel data induces both cointegration between the variables across the groups (cross-section cointegration) as well as within group cointegration. Moreover, the asymptotic theory is more complicated due to the fact that the sampling design involves a time as well as a cross-section dimension (Breitung and Pesaran, 2008). Using recent advances in panel unit root tests, Gertham and Lothgren (2000) find HE and GDP across OECD countries are non-stationary and cointegrated. Jewell *et al.* (2003) and Carrion-i-Silvestre (2005) conclude HE and GDP are stationary around one or two breaks. Freeman (2003) find HE and GDP in the Unites States at the state level in the period 1966-1998 are non-stationary and cointegrated, and income elasticity is less than 1. Wang and Rettenmaier (2007) use a panel of the US states, employ a unit root test that allows structural breaks, and find HE and GDP are non-stationary and cointegrated. They find income elasticity for some states are greater than 1 and for others are less than 1.

4.3 Unit Root and Cointegration

Heterogeneity across individual series in a panel implies some complications in unit root and cointegration hypotheses in the panel data. The heterogeneity characteristic of a panel implies that the cointegration vector differs across individual series or groups of the panel. Given the assumption of homogeneity versus heterogeneity in panels, Breitung and Pesaran (2008) categorize unit root and cointegration tests in panels as first and second generation tests, respectively. The first generation unit root tests assume no cross-correlation across units of a panel[14]. The null hypothesis in these tests is that all units in the panel contain unit roots. The alternative hypothesis is that either all units are stationary or a significant proportion of the series is stationary. The second generation of unit root tests has emerged to deal with the issue of cross-correlation between the individual series in a panel. Cross-correlation likely exists if the panel contains macroeconomic time series of a country or a sector of an economy, such as regional or provincial data. For instance, cross-correlations arise due to several factors such as omitted observed common factors, spatial spill over effects or unobserved common factors. In the context of healthcare, cross dependence in panels can arise from unobservable shocks that can originate in technological advances, health shocks and/or implementation of new health policies that affect all units of a panel simultaneously. Nonetheless, not all units necessarily respond to a common shock in a homogeneous way. However, if the cross-correlation is due to a single observed/unobserved common factor that homogenously affects cross units of a panel, then the effect can be removed simply by subtracting the cross-section means from the data (Pesaran, 2007; Moon and Perron, 2004; Phillips and Sul, 2003).

Conventional estimators for panels, such as an OLS regression, are inefficient when units of the panel are contemporaneously correlated (Phillips and Sul, 2003). In addition, stationarity tests that assume no cross dependence might have substantial size

[14] Levin and Lin, (1993); Maddala and Wu, (1999); Choi, (2001); Levin, Lin and Chu, (2002); Chang, (2002); Im, Pesaran and Shin, 2003)

distortion if the assumption does not hold (Phillips and Sul, 2003; Maddala and Wu, 1999).

There are two major approaches to test conintegration in panels. First, one can perform panel unit root test on residuals from the regression of one time series on another one (Kao, 1999; Pedroni, 1999, 2004; Westerlund, 2005; Pesaran, 2007). Second, one can use system approach tests that allow for more than one cointegration relationship between the time series (Larsson, Lyhegen and Lothgren, 2001; Groen and Kleigergen, 2003; Breitung, 2005).

To overcome nuisance in the asymptotic properties of unit root and cointegration tests in presence of cross dependence, the literature either relies on techniques for correcting the standard error of estimators, including the GLS estimator of seemingly unrelated regression system (Groen and Kleigergen, 2003), the panel corrected standard error approach (Jonsson, 2005), and a bootstrap method (Maddalla and Wu, 1999; Chang and Park, 2003), or it uses an unobserved common effects approach for controlling heterogeneity in response to common shocks (Pesaran, 2007). However, the cross-correlation is a more serious problem if it leads to cross-cointegration, and so a different test is required to deal with this difficulty in panels (Bai and Ng, 2004).

In this study, I employ a bootstrap method and some recent advances in time series econometric techniques to test the unit root and cointegration between HE and GDP in a panel of the ten Canadian provinces. I apply the tests for total health expenditure and by the following subcategories: source of finance, including public, federal direct, and provincial expenditure, and use of funds, including hospital, physician, capital, and public health expenditures. Then, upon finding a cointegration vector between the time series, I test the long-term and short-term effects of income on health expenditures.

I also test stationarity and cointegration using the common factor effects approach advanced by Pesaran (2007) and find cross-dependence persists even after incorporating cross-mean group variables to approximate the cross-dependence effect into the regression model (the results are reported in Appendix 4.A). The seemingly unrelated regression approach proposed by Groen and Kleigergen (2003) works when the number of units in the panel is very low relative to the time periods, which is not the case in this study. The panel corrected standard error approach gives an efficient estimate when $T \to \infty$ followed by $N \to \infty$, where T and N are time and cross-unit dimensions of a panel, respectively (Breitung and Pesaran, 2008). Thus, I rely on the bootstrap method recommended by Maddala and Wu (1999) and Chang and Park (2003) to get the empirical distribution of the test statistics. This is also the strategy adopted by Carrion-i-Silvestre *et al.* (2005), Wang and Rettenmaier (2007) and Westerlund (2007).

To test for stationarity of real per capita GDP and real per capita HE by subcategories, I employ an augmented-Dickey-Fuller test given as

$$\Delta y_{it} = \alpha_i + \phi_i y_{i,t-1} + \beta TT + \sum_{j=1}^{p} \gamma_j \Delta y_{i,t-j} + \varepsilon_{it} \tag{4.1}$$

where y_{it} is either real per capita GDP or HE in the i^{th} province at time t, $i = 1,...,N$ and $t = 1,...,T$. Δ denotes the first difference operator, TT is a time trend variable, ε is an error term, p denotes maximum allowed lags in the model and other notations are parameters to be estimated. The null hypothesis of the unit root in the panel is that all N

time series containing unit roots, i.e. $\phi_1 = \phi_2 = ... = \phi_N = 0$. The alternative hypothesis is that at least one of the series is stationary, $\phi_i < 0$ and $i = 1,...,n_1$ where $n_1 \leq N$.

In general, it is recommended that researcher should construct a test statistic based on the average of individual t-statistics from the augmented-Dickey-Fuller model to test stationarity in a panel (Im, Pesaran and Shin, 2003). However, if cross-units are correlated, then t-statistics are correlated and the critical values of the Dickey-Fuller test are not valid. I employ the Fisher statistic proposed by Maddala and Wu (1999) which is defined as

$$\lambda = -2\sum_{i}^{N} \ln(P_i) \tag{4.2}$$

where P_i is the observed significance level (p-value) of the null hypothesis for the i^{th} province. The sum of the p-values has χ^2 distribution with $2N$ degree of freedom. For the cross dependence problem, Maddala and Wu (1999) propose a bootstrap method to get the empirical distribution of the test statistic. The test has an excellent performance and is very simple to implement, as it is very flexible to apply to different null hypotheses, stationarity versus non-stationarity null, and to adopt different maximum lags in individual augmented-Dickey-Fuller tests.

Moreover, Choi (2001) proposes a series of tests based on combination tests, but the most promising is the inverse normal test defined by

$$Z = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Phi^{-1}(P_i) \tag{4.3}$$

where $\Phi(.)$ is the standard normal cumulative distribution function and the other notations are similar to the Fisher test. For fixed N, Z has a standard normal distribution with a mean of zero and a variance of 1.

A cointegration test is very simple to perform using the Fisher or Choi statistic, in which one can perform the stationarity test on residuals from regression HE on GDP. In addition, I employ an error correction model approach to test the cointegration relationship between HE and GDP. I employ a new method advanced by Westerlund (2007) which is based on structural rather than residual. Westerlund (2007) argues that residual based tests require the long-term cointegrating vector for the variables in their level to be equal to the short-term adjustment process for the variables in their first difference. However, an error correction based test allows the short-term and the long-term adjustment processes to differ. Using a simulation method, he shows that his new test has more power than the popular residual based test by Pedroni (2004). He also suggests a bootstrap method for getting the empirical distribution of his statistics in the presence of cross-dependence. The results of this supplementary cointegration test are reported in Appendix 4.B.

4.4 Income Elasticity of Health Expenditure

The results from testing the cointegration relationship between HE and GDP using the Fisher and Choi statistics show that the time series are cointegrated after control for the lagged HE (the results are reported in section 4.6 and in Table 4.3). Thus, to estimate the long-term income elasticity, an estimate model should include the lagged HE as an explanatory variable. To do this, I use the following dynamic panel

$$HE_{it} = \alpha_i + \delta GDP_{it} + \lambda X_{it} + \beta HE_{i,t-1} + u_{it} \qquad (4.4)$$

where HE denotes log of real per capita health expenditure, GDP denotes log of real per capita gross domestic product, X is a vector of explanatory variables, α_i is a provincial specific intercept and u_{it} is an error term. In model (4.4), δ measures the short-term income elasticity, and the long-term income elasticity can be estimated by the ratio $\frac{\delta}{1-\beta}$.

The presence of the lagged dependent variable in model (4.4) induces error terms to be correlated with it, even if the error terms are not serially correlated. Thus, the problem of endogeneity arises, which induces biased and inconsistent estimates. To overcome the difficulty, I employ an instrumental variable technique (IV). The literature suggests using the second lag of the dependent variable as an excluded instrument; alternatively, the second and third lags of the dependent variable together, in their level or in their first difference, can be used if a model employs more than one instrument (Anderson and Hsiao, 1981; Arellano, 1989). The instruments are valid if they are uncorrelated with the error terms and are highly correlated with the endogenous variable. The instrumental variable technique leads to a consistent estimate but it is not necessarily efficient, because it does not make use of all the available moment conditions (Ahn and Schmidt, 1995). Arellano and Bond (1991) propose a Generalized Method of Moments (GMM) procedure that is more efficient than the IV estimate.

In this study, I use two different sets of instruments. The first set includes the second and third lags of the dependent variable. The second set includes the second lag of the dependent variable and the lag of federal cash transfers to provinces. Federal cash transfer to provinces is highly correlated with health expenditure, because it has been one of the major sources of funding for provincial health expenditures, and its first lag is not correlated to current shocks (error terms) if provinces do not carry forward the transferred money to subsequent years. I test validity of the instrumental variables using a Sargan test (Sargan, 1958) and then choose whichever set that its validity is not rejected by the test.

To run the Saran test, one needs to examine the correlation between all instruments and the residuals from IV/GMM estimate. Although the residuals are unobservable, they can be predicted after running the IV/GMM estimate. However, if the estimates are exactly identified, i.e. the number of excluded instruments is equal to the number of endogenous variables, the orthogonality condition implies that the correlation between predicted residuals and the lagged dependent variable, the endogenous variable, will be zero by construction. Thus, to get a higher moment condition, one needs to utilize an over identified model, in which the number of excluded instruments should at least be equal to the number of endogenous variables plus 1. I use the following steps, as described by Davison and Mackinnon (2004), for testing the validity of instruments. First, I derive the residuals from GMM estimation of model (4.4). Second, I run an OLS regression of the predicted residuals on the full set of instruments, both included and excluded instruments, and then calculate N times R^2, where N is the number of observations and R^2 is the overall goodness of fit from the latter regression. The result is a test statistic. If the model is correctly specified, then the test statistic will have χ^2 distribution with $l-k$ degree of freedom, where l denotes the number of excluded instruments and k denotes the number of endogenous variables. The critical value of the test statistic can be derived from the χ^2 distribution table at $l-k$ degree of freedom, which can be used as a threshold point for rejecting the null hypothesis.

The dynamic panel estimate provides limited information, as it estimates the relationship between HE and GDP in a single state, asymmetric regime. An error correction model is a popular approach for testing asymmetry in time series analyses. In this study, I use an ECM model to test asymmetrically-response of health expenditure to economic growth in periods of economic strength and weakness. The model is given as

$$\Delta HE_{it} = \alpha + \eta Z_{i,t-1} + \sum_{j=1}^{K} \delta_j \Delta HE_{i,t-j} + \sum_{s=0}^{L} \pi_s \Delta GDP_{i,t-s} + \varepsilon_{it} \tag{4.5}$$

where, once again, HE is real per capita health expenditure, GDP is real per capita gross domestic product, Δ denotes the first difference operator, and K and L are maximum lags on the first difference of the variables. Z is the predicted residual from regression HE on GDP, in which

$$HE_{it} = \gamma + \varphi GDP_{it} + \nu_{it} \tag{4.6}$$

and

$$Z_{i,t-1} = HE_{i,t-1} - \hat{\gamma} - \hat{\varphi} GDP_{i,t-1} \tag{4.7}$$

where $\hat{\gamma}$ and $\hat{\varphi}$ are estimated parameters in model (4.6). In model (4.5), η is the coefficient of the error correction term which measures the effect of deviation from the long-term equilibrium on growth of health expenditure. In other words, it measures the speed of adjustment of health expenditure toward the long-term equilibrium level, while δ and π measure the short-term adjustments of the dependent variable respectively to changes in the lagged dependent variable and economic growth. Other notations are parameters to be estimated. Lag numbers in model (4.5) are estimated using an information criterion, such as Akaiki's Information Criterion (AIC). If HE and GDP

contain unit roots and their first differences are stationary and if they are cointegrated, then all right hand side variables in model (4.5), including Z and ε , are stationary.

The asymmetric feature of income effect can be incorporated into model (4.5) as follows

$$\Delta HE_{it} = \alpha + \eta_1 Z^{+}_{i,t-1} + \eta_2 Z^{-}_{i,t-1} + \sum_{j=1}^{K_1} \delta_{1j} \Delta HE^{+}_{i,t-j} + \sum_{j=1}^{K_2} \delta_{2j} \Delta HE^{-}_{i,t-j} + \sum_{s=0}^{L_1} \pi_{1s} \Delta GDP^{+}_{i,t-s} + \sum_{s=0}^{L_2} \pi_{2s} \Delta GDP^{-}_{i,t-s} + \varepsilon_{it}$$

(4.8)

where $Z^{+}_{i,t-1} = \max\{\tau, Z_{i,t-1}\}$, $Z^{-}_{i,t-1} = \min\{\tau, Z_{i,t-1}\}$, where τ is a threshold point, which can be equal to 0 or its value can be estimated by maximizing the likelihood function.

In model (4.8), δ_{1j} measures the effect of $\Delta HE_{i,t-j} > \tau$, δ_{2j} measures the effect of $\Delta HE_{i,t-j} \leq \tau$, π_{1s} measures the effect of $\Delta GDP_{i,t-s} > \tau$ and π_{2s} measures the effect of $\Delta GDP_{i,t-s} \leq \tau$.

To find the threshold point, one can use the average historical growth rate of the economy. However, I estimate the optimum threshold point using an approach proposed by Chan (1993), which is used in estimate of threshold autoregressive models. This method works by maximizing the likelihood function of the asymmetric specification, thus provides a more robust estimate for rejecting the null of health is a luxury good. To do this, first I calculate the cross averages of the income variable (ΔGDP) in order to get a single value of the variable for all the provinces in a given year and then sort them in ascending order. Second, I set the threshold point equal to each value of the cross-averaged ΔGDP and run model (4.8). The optimum threshold point is derived by minimizing the AIC criterion. To ensure that there are enough observations in each regime split by the threshold point, I use the values of ΔGDP that fall between the 20th and 80th percentiles of its distribution.

To estimate model (4.8), I first estimate model (4.6) using a fixed effect estimate to derive residuals Z, then model (4.8) is estimated using a fixed effects or a GMM approach. Asymmetry in the effect of income can be tested using the following F-test:

$$\begin{cases} H_0 : \eta_1 + \sum \pi_{1s} = \eta_2 + \sum \pi_{2s} \\ H_A : \eta_1 + \sum \pi_{1s} \neq \eta_2 + \sum \pi_{2s} \end{cases}$$

The lagged dependent variable in model (4.8) is an endogenous variable, so to use a GMM approach I employ four sets of instruments. The first set includes the second and third lags of the dependent variable. The second set includes the second lag of the dependent variable and the first lag of federal cash transfers to provinces. The last two sets consist of the first two sets of instruments in their levels rather than in their first differences. Having run the estimates, I choose a set of instruments which passes the validity test (Sargan test).

Finally, if the variables of model (4.8), which are expressed in their first differences, are stationary, then excluding the lagged dependent variable from model (4.8) does not affect the stationarity of the residuals. Such an exclusion restriction removes the short-term adjustments of current health expenditure to changes in the lagged health expenditures, but it improves consistency of the estimates in the case that the instruments are weak or invalid. Thus, I also estimate model (4.8) where it excludes the lagged dependent variable and compare the results with the GMM estimate.

4.5 Data

This study uses aggregate data for income and health expenditure in Canada at the provincial level for the period 1975 to 2006. Health expenditure can be broken down into a number of subcategories by sources of finance: including public, federal direct and provincial expenditures; and by uses of funds: including hospital, physician, capital and public health expenditures. The data for health expenditures are obtained from the Canadian Institute for Health Information (CIHI). Other data, including GDP, federal cash transfers to provinces, total population, the population aged 65 and older, consumer price index and health services price index are collected from Statistics Canada (CANSIM).

Federal cash transfers to provinces include transfers to health, postsecondary education and equalization. The share of total federal cash and tax point transfers as a percentage of total provincial health expenditure has decreased from 41.3% to 29.3% during the period 1975 to 2000 (Lazar and St-Hilaire, 2004), but these transfers have remained as one of the major determinants of provincial health expenditures. The proportion of the provincial population aged 65 and older for the period 1975 to 2006 has increased from a low of 4%, from 10% to 14%, for Manitoba to a high of 8%, from 6% to 14%, for Newfound Land. Since the elderly have a higher demand for healthcare services, the aging population should have a substantial effect on growth of health expenditure.

In this study, the estimated models use the log value of GDP and HE per capita at the provincial level, adjusted for inflation using the provincial consumer price index. Figure 4.1 displays the increases in the proportion of real health expenditure per capita to real GDP per capita for the period 1975-2006 by region in Canada: including Atlantic, Central Canada, the Prairies and BC. Table 4.1 includes summary statistics of the data.

4.6 Results

This section discusses estimates of stationarity, cointegration, income elasticity, and asymmetric response of health expenditure to economic growth using estimate models (4.1), (4.4), and (4.8). The results are reported in Tables 4.2 to 4.8.

4.6.1 Unit Root and Cointegration

Table 4.2 displays panel unit root tests for real per capita GDP and health expenditures, by source of finance and use of funds. The unit root tests are performed by the Fisher and Choi statistics derived from an augmented Dickey-Fuller test, with the lag number that varies from 0 to 4 and with/without a time trend. Then, a bootstrap method (Maddala and Wu, 1999) with 1000 replications is used to get the empirical distribution of the test statistics. I find GDP is not stationary. The only exception is when the Fisher statistic is employed with a time trend and 0 lag. Moreover, total health expenditure and its subcategories are not stationary when the time trend is excluded, except for federal direct, hospital and capital expenditures, which are trend stationary when the augmented Dickey-Fuller test includes a higher number of lags. However, all variables are stationary in their first difference (not reported).

Table 4.3 displays the residual based cointegration tests using the Fisher and Choi statistics. Residuals from regression of HE on GDP are estimated by a fixed effects estimate method in three models. The first model includes GDP as the only regressor. The second model includes GDP and a time trend; and the third model includes GDP and the lagged dependent variable. Then, the Fisher and Choi's tests are performed on predicted residuals. Table 4.3 displays that health expenditure per capita and its subcategories are cointegrated with GDP per capita when the lagged dependent variable is included in the model (refer to the last set of three columns of Table 4.3). In addition, the existence of a cointegraton vector is independent of the number of lags included in the model, except for capital and public health expenditures when four lags are included and the Fisher's test is employed. Moreover, exclusion of the lagged dependent variable and inclusion of a time trend in the model make health expenditures by use of funds are not cointegrate with GDP, but health expenditures by source of finance are conintegrated only at a higher lag number except for capital and public expenditures (refer to the middle set of three columns in Table 4.3). Health expenditure and GDP are not cointegrated when both time trend and the lagged dependent variable are excluded from the test, except for public, provincial, hospital and physician expenditures at a higher lag number (refer to the first set of three columns in Table 4.3). The cointegraton results suggest that a conservative regression model should include the lagged dependent variable to avoid spurious regression.

Furthermore, Table 4.10 in Appendix 4.B displays a supplementary cointegration test, using an error correction model approach advanced by Westerlund (2007), which was performed to find sensitivity of the cointegration test results to model specification. The results show that GDP and health expenditure, by subcategories, are conintegrated (the results for the subcategories of health expenditure are not reported for space reserving). The results are consistent with the Fisher and Choi's tests when the lagged dependent variable is included.

4.6.2 Long-term Income Elasticity

Tables 4.4 and 4.5 display the estimated short-term and long-term income elasticities of health expenditures using estimate model (4.4), without and with control for other covariates, respectively. I first test for appropriateness of the random effects estimate versus the fixed effects estimate using the Hausman test (Hausman, 1978). The test results are reported in the column of Hausman Test in both tables. Appropriateness of the random effects estimate was rejected for all outcomes. Thus, a fixed effects estimate method is used for estimating the income elasticities.

Table 4.4 shows that the short-term income elasticity is small and varies in the range 0.11 to 0.39 when fixed effects estimate is used, and in the range 0.11 to 0.51 when GMM estimate approach is used. However, the long-term income elasticity for all outcomes is above unity using both fixed effects and GMM estimates, except for the elasticity of hospital expenses which is below but close to unity. However, I find the estimated income elasticities greater than 1 are due to a lack of control for other covariates. Table 4.5 displays estimates of the short-term and long-term income elasticities, when controlled for other covariates, including a time trend, the proportion of the population aged 65 and older, and the relative price of healthcare services to overall prices in the economy. The table displays the decreasing magnitude of both the short-term and long-term income elasticities after control for the other covariates, in which the long-term income elasticities fall bellow unity, except for physician and hospital expenditures which remain above unity at 1.32 and 1.54, respectively when fixed effect estimate is used and control for the other covariates. The results persist when GMM method is used, except for hospital expenditure that falls below unity at 0.90. Overall, the results are consistent with the recent literature that finds health is not a luxury good. In particular, I find an income elasticity of total health expenditure that is lower than what was found by Di Matteo and Di Matteo (1998), in which in my estimates the income elasticity is 0.61 and 0.47, respectively when fixed effects and GMM estimates are used compared with Di Matteo and Di Matteo's (1998) estimation of 0.77 when a pooled time-series cross-section regression is used.

4.6.3 Asymmetric Response of Health Expenditure to Economics Growth

This section discusses the estimate of model (4.8) using a fixed effects estimate method and a GMM estimate approach. The results are reported in Tables 4.6 to 4.8. All variables, including the dependent and explanatory variables, are entered in the form of the first difference of their log values, except for the error correction term which is entered in the form of log value in its level. Each model is estimated with 0 to 5 lags and then the optimum lag number is derived using the AIC criterion.

The left panel of Table 4.6 displays the symmetric and asymmetric effects of the growth rate of income on the growth rate of health expenditure. The table reports the sum of the long-term and short-term adjustments of health expenditure to economic growth. For the asymmetric regime, the effects of income in periods of economic strength and weakness are estimated when a zero threshold point is used to identify the states of economic growth. I used an F-test to identify whether the effect of income on health expenditure in the state of economic strength is different from the effect in the state of economic weakness. The left panel of the table shows that there is no asymmetry in the effect of income with the states of economic growth, that is, I cannot reject the equality in the effects of income by the states of economic growth using the F-test. The results sustain when a GMM estimate approach or a fixed effects estimate that excludes the lagged dependent variable is used (the results are reported in the middle and the right panel of Table 4.6, respectively).

A Sargen test (not reported) is used for testing the validity of instruments in the GMM estimates. Among the four sets of available instruments, I find the second lag of the dependent variable and the first lag of federal cash transfers to provinces in their level are valid instruments for all outcomes, which are used in the estimates.

4.6.4 Asymmetric Response of Health Expenditure to Economics Growth, Where the Threshold Point is Optimized

I use an approach proposed by Chan (1993) to get the optimum value of the threshold point. The estimated results are reported in Table 4.7. The optimum threshold points are reported in the second column of the table only for the models that includes the lagged dependent variable. Overall, using both fixed effects and GMM estimates, I find there is no asymmetry in the effect of income when the threshold point is optimized.

4.6.5 Asymmetric Response of Health Expenditure to Economics Growth When Control for Other Covariates and the Threshold Point is Optimized

This section discusses the results of the symmetric versus asymmetric effects of income using model (4.8), when control for other covariates. I add a time trend, the proportion of the population aged 65 years and older, and the relative price of healthcare services to overall prices in the economy. The time trend variable recovers the effect of technological advances on health expenditures. The proportion of the population aged 65 years and older recovers the effect of demographic changes. The relative price index recovers the effect of rapid growth of prices in the health sector compared with growth of prices in the whole economy. The results are reported in Table 4.8, where the threshold point is optimized by minimizing the AIC criterion.

Once again, I find no asymmetry in the effect of changes in the growth rate of income on the growth rate of health expenditure. The results stand with employment of different model specification, a fixed effect estimate versus a GMM estimate procedure, and also when the lagged dependent variable is excluded from the fixed effect estimates.

The absence of asymmetric response pattern of health expenditure to economic growth suggests that health is not a luxury good. The no-relation suggests that health expenditure is likely to be determined by policy makers based on the long-term growth path of the economy, and is more likely affected by policies and social programs that have long lasting effects rather than by growth of income itself. For instance, demographic changes, the ageing population, and technological advances should have stronger effect on health expenditure than the growth rate of the economy.

4.7 Conclusion

This study used a panel data of the ten Canadian provinces on health expenditure and income to test Wagner's Law for explaining the increases in health expenditure over time. This study contributed to the literature by using a new specification method and advances in time series econometrics. I tested Wagner's Law by estimating the long-term income elasticity of health expenditure and identifying asymmetry in the effect of income on health expenditure. I estimated the effects by subcategories of health expenditure, including by source of finance and use of funds.

I first tested for stationarity and cointegration between the time series using the Fisher and Choi statistics, which are combined tests based on p-values of individual augmented Dickey-Fuller unit root tests. Moreover, I overcame the difficulty of cross-section dependence among individual units of the panel using a bootstrap method introduced by Maddala and Wu (1999). Using the combined tests, I found GDP and health expenditure series contain unit root. I tested for cointegration between the time series using residuals derived from a fixed effects estimate of health expenditure on GDP. I found that the time series are conintegrated when controlled for the lagged dependent variable.

I tested the short-term and long-term income elasticity of health expenditures and found that health is not a luxury good and has an income elasticity between 0.47 and 0.61, which varies with model specification, but hospital and physician expenditures with income elasticities above or close to unity are highly income sensitive.

I tested asymmetry in the effect of income on health expenditure using an error correction model. To identify an asymmetric regime, I distinguished periods of economic strength from economic weakness using a threshold point that maximized the likelihood of having an asymmetric regime. That is, to reject the null of existence of an asymmetric regime, I maximized the likelihood function under the assumption of existence of an asymmetric regime in order to get a higher power test. I found that there is no asymmetry in the effect of income and concluded that health is not a luxury good. The absence of an asymmetrically response pattern of growth of health expenditure to economic growth suggests that health expenditure is more likely determined by policy makers based on the long-term growth path of the economy rather than by short-term fluctuations in the economy. Moreover, it suggests that the long-term growth of health expenditure is determined by factors affecting the steady state of the economy such as technological advances and demographic changes.

Figure 4.1: The share of health expenditure as a proportion of GDP in Canada

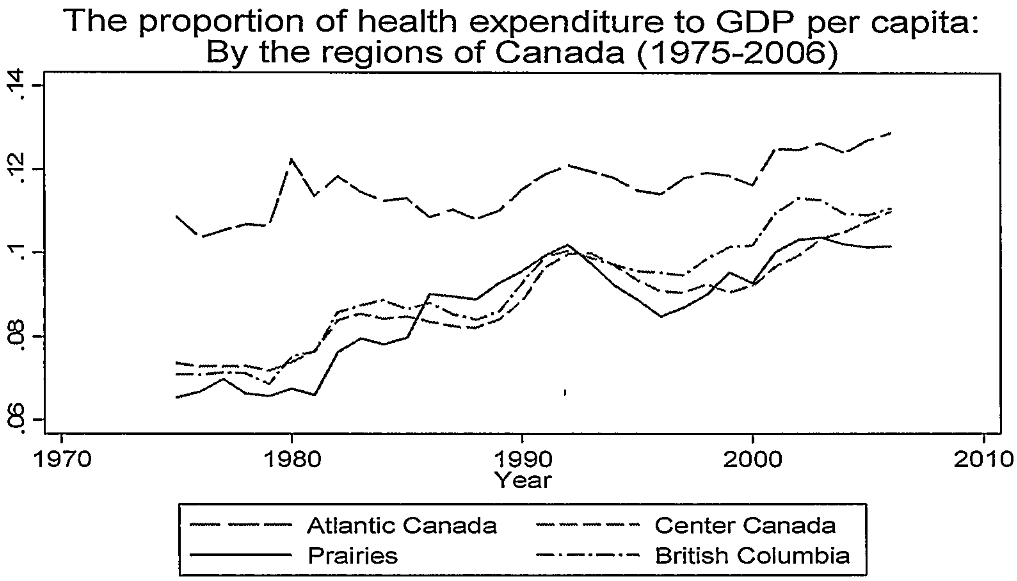


Table 4.1: Summary Statistics of health expenditures and GDP per capita in Canada: The period 1965-2006

	Mean	SD	Min.	Max.
Gross Domestic Product	28359.63	8085.43	13503.39	63411.71
Total Expenditure	2737.44	708.05	1413.33	4585.69
Public Expenditure	2015.01	495.18	1094.12	3446.55
Provincial Expenditure	1862.19	438.68	1037.48	3049.56
Federal Expenditure	110.21	65.21	27.96	347.10
Hospital Expenses	1004.04	167.06	697.15	1588.51
Physician Expenses	351.96	107.99	157.86	628.82
Capital Expenses	111.15	60.69	18.97	380.48
Public Health Expenses	123.24	70.62	37.18	414.85
Federal Cash Transfers	1855.52	1034.79	466.30	7165.32
Population aged 65 and older (%)	.12	.02	06	.15
Price index of healthcare to CPI	.96	.08	.68	1.10
Number of Observations	320			

Note: Health expenditure and GDP are in real per capita values.

Table 4.2: The Fisher and Chio panel unit root tests for GDP and health expenditure per capita by sources of finance and use of funds.

		Without Trend			With Trend		
Variable	Lags	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)
GDP	0	0.99	1.0	0.99	0.47	0.09	0.79
	1	0.99	1.0	0.99	0.82	0.95	0.99
	2	0.99	1.0	0.99	0.71	0.69	0.83
	3	0.99	1.0	0.99	0.81	0.76	0.85
	4	0.99	1.0	0.99	0.97	0.94	0.92
Total Expenditure	0	1.0	1.0	1.0	0.99	0.96	0.90
	1	1.0	1.0	1.0	0.99	0.99	0.97
	2	1.0	1.0	0.99	0.96	0.94	0.72
	3	1.0	1.0	0.99	0.78	0.71	0.51
	4	1.0	1.0	0.99	0.51	0.45	0.16
Public Expenditure	0	1.0	1.0	0.99	0.99	0.78	0.43
	1	1.0	1.0	1.0	0.96	0.98	0.93
	2	0.99	1.0	0.99	0.81	0.79	0.65
	3	0.99	1.0	0.99	0.60	0.47	0.17
	4	0.99	1.0	0.99	0.19	0.20	0.12
Province Expenditure	0	1.0	1.0	0.97	0.99	0.95	0.86
	1	1.0	1.0	1.0	0.94	0.94	0.81
	2	0.99	1.0	1.0	0.81	0.89	0.81
	3	0.99	1.0	0.99	0.57	0.71	0.63
	4	0.99	1.0	0.99	0.29	0.26	0.15
Federal Expenditure	0	1.0	1.0	0.99	0.17	0.07	0.00
	1	1.0	1.0	1.0	0.00	0.05	0.02
	2	1.0	1.0	0.99	0.00	0.01	0.00
	3	0.99	1.0	0.99	0.02	0.05	0.01
	4	0.99	0.96	0.98	0.08	0.10	0.10
Hospital expenses	0	0.99	0.99	0.98	0.99	0.81	0.50
	1	0.99	1.0	0.99	0.77	0.89	0.75
	2	0.93	0.85	0.65	0.58	0.54	0.24
	3	0.80	0.81	0.78	0.53	0.65	0.32
	4	0.54	0.53	0.33	0.44	0.08	0.01
Physician expenses	0	0.99	1.0	0.99	0.99	0.88	0.78
	1	0.99	1.0	0.99	0.77	0.91	0.85
	2	0.99	0.99	0.99	0.84	0.69	0.48
	3	0.99	1.0	1.0	0.87	0.88	0.78
	4	0.99	0.95	0.99	0.27	0.17	0.11
Capital expenses	0	0.25	0.10	0.06	0.15	0.04	0.02
	1	0.02	0.08	0.04	0.11	0.62	0.35
	2	0.10	0.30	0.40	0.54	0.65	0.38
	3	0.24	0.27	0.50	0.87	0.89	0.70
	4	0.55	0.46	0.54	0.99	0.98	0.92
Public Health expenses	0	0.99	0.99	0.98	0.97	0.52	0.22
	1	0.99	1.0	0.99	0.84	0.92	0.95
	2	0.99	1.0	1.0	0.93	0.95	0.90
	3	1.0	1.0	0.99	0.98	0.95	0.93
	4	1.0	1.0	0.99	0.99	0.93	0.74

* Bootstrap results were derived using 1000 replications.

Table 4.3: Cointegration between GDP and health expenditures per capita, by source of finance and use of funds

		Model (1)			Model (2)			Model (3)		
Variable	Lags	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)
Total Exp.	0	0.58	0.17	0.20	0.72	0.29	0.11	0.00	0.00	0.00
	1	0.64	0.70	0.66	0.26	0.49	0.21	0.00	0.00	0.00
	2	0.67	0.62	0.77	0.09	0.18	0.02	0.00	0.01	0.00
	3	0.46	0.63	0.75	0.03	0.28	0.07	0.00	0.01	0.00
	4	0.02	0.08	0.30	0.01	0.09	0.02	0.00	0.00	0.00
Public Exp.	0	0.24	0.29	0.21	0.61	0.26	0.07	0.00	0.00	0.00
	1	0.23	0.53	0.62	0.18	0.54	0.23	0.00	0.00	0.00
	2	0.20	0.30	0.33	0.01	0.05	0.01	0.00	0.01	0.00
	3	0.05	0.19	0.23	0.01	0.08	0.02	0.00	0.01	0.00
	4	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.02	0.00
Province Exp.	0	0.34	0.05	0.01	0.69	0.15	0.05	0.00	0.00	0.00
	1	0.21	0.38	0.40	0.18	0.30	0.08	0.00	0.00	0.00
	2	0.18	0.24	0.21	0.03	0.06	0.01	0.00	0.01	0.00
	3	0.06	0.08	0.04	0.02	0.09	0.02	0.00	0.01	0.00
	4	0.00	0.03	0.03	0.00	0.02	0.01	0.01	0.02	0.00
Federal Exp.	0	0.40	0.37	0.30	0.00	0.03	0.00	0.00	0.00	0.00
	1	0.40	0.78	0.82	0.00	0.00	0.00	0.00	0.00	0.00
	2	0.20	0.36	0.55	0.00	0.00	0.00	0.00	0.00	0.00
	3	0.59	0.66	0.66	0.00	0.01	0.00	0.00	0.00	0.00
	4	0.58	0.67	0.65	0.00	0.00	0.01	0.00	0.00	0.00
Hospital Exp.	0	0.76	0.42	0.17	0.80	0.63	0.31	0.00	0.00	0.00
	1	0.09	0.35	0.14	0.05	0.12	0.04	0.00	0.00	0.00
	2	0.02	0.05	0.01	0.01	0.09	0.01	0.00	0.00	0.00
	3	0.01	0.04	0.00	0.01	0.06	0.01	0.00	0.00	0.00
	4	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00

Table 4.3 (continued)

		Model (1)			Model (2)			Model (3)		
Variable	Lags	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)	Observed P-value	Bootstrap P-value (Fisher)	Bootstrap P-value (Choi)
Physician Exp.	0	0.73	0.62	0.31	0.97	0.75	0.35	0.00	0.00	0.00
	1	0.27	0.59	0.58	0.72	0.92	0.80	0.00	0.00	0.00
	2	0.32	0.57	0.66	0.75	0.81	0.59	0.00	0.00	0.00
	3	0.22	0.31	0.29	0.40	0.59	0.50	0.00	0.00	0.00
	4	0.01	0.05	0.07	0.13	0.13	0.07	0.01	0.01	0.00
Capital Exp.	0	0.03	0.08	0.01	0.02	0.09	0.01	0.00	0.00	0.00
	1	0.00	0.21	0.22	0.00	0.20	0.19	0.00	0.00	0.00
	2	0.06	0.15	0.15	0.05	0.15	0.17	0.00	0.00	0.00
	3	0.16	0.40	0.61	0.18	0.22	0.24	0.00	0.01	0.01
	4	0.46	0.53	0.54	0.45	0.30	0.28	0.11	0.20	0.09
Public Exp.	0	0.59	0.24	0.15	0.56	0.33	0.08	0.00	0.00	0.00
	1	0.70	0.87	0.85	0.20	0.54	0.70	0.00	0.00	0.00
	2	0.67	0.65	0.74	0.18	0.29	0.62	0.00	0.01	0.00
	3	0.80	0.58	0.42	0.50	0.27	0.33	0.01	0.03	0.01
	4	0.89	0.72	0.49	0.73	0.81	0.79	0.13	0.14	0.07

- The cointegration tests are based on residuals predicted from a fixed effects estimate of health expenditure on GDP.
    1. Model 1 includes GDP as only regressor
    2. Model 2 includes GDP and a time trend as regressors.
    3. Model 3 Includes GDP and the lagged dependent variable as regressors.
- Bootstrap results were derived using 1000 replications.

Table 4.4: The short-term and long-term income elasticity of health expenditures

	Fixed Effects Estimate			Generalized Method of Moment (GMM) Estimate			
	S.R. Elasticity	L.R. Elasticity	Hausman test* (FE vs. RE)	S.R. Elasticity	L.R. Elasticity	Sargan Test**	First Stage F-Test***
Total Expenditure	0.11	1.80	28.08	0.11	1.91	0.36	2172
	(.02)	(.22)	[.00]	(.02)	(.25)	[.55]	[.00]
Public Expenditure	0.15	1.62	44.1	0.16	1.66	0.53	1536
	(.02)	(.15)	[.00]	(.03)	(.17)	[.47]	[.00]
Provincial Expenditure	0.15	1.53	44.4	0.16	1.54	0.47	1470
	(.02)	(.14)	[.00]	(.03)	(.14)	[.51]	[.00]
Federal Expenditure	0.14	3.49	11.84	0.19	2.86	0.01	1079
	(.05)	(.99)	[.00]	(.06)	(.56)	[.94]	[.00]
Hospital Expenses	0.13	0.94	41.73	0.17	0.87	0.20	706
	(.02)	(.12)	[.00]	(.02)	(.07)	[.66]	[.00]
Physician Expenses	0.13	2.04	34.82	0.17	1.82	0.17	1524
	(.02)	(.28)	[.00]	(.03)	(.20)	[.68]	[.00]
Capital Expenses	0.39	1.31	7.79	0.51	1.09	0.06	144
	(.11)	(.36)	[.02]	(.11)	(.25)	[.80]	[.00]
Public Health Expenses	0.16	2.48	5.64	0.19	2.74	0.15	875
	(.05)	(.54)	[.06]	(.06)	(.52)	[.70]	[.00]

Note: A fixed effects estimate was used to regress health expenditure per capita on GDP per capita and the first lagged dependent variable. The GMM estimates use the first lag of federal cash transfers to provinces and the second lag of the dependent variable as excluded instruments. Standard errors are robust and reported inside parentheses, and p-value of the test statistics are reported inside brackets. Number of observations is equal 310.

* The Hausman test was used to examine appropriateness of random effects (RE) estimate versus fixed effects (FE) estimate. Random effects estimate was rejected for all outocmes.

** The Sargan test was used to test validity of the instruments. I used the first lag of federal cash transfers and the second lag of the dependent variable as excluded instruments.

*** F-Test displays power of excluded instruments.

Table 4.5: The short-term and long-term income elasticity of health expenditures, controlling for other covariates

	Fixed Effects Estimate			Generalized Method of Moment (GMM) Estimate			
	S.R. Elasticity	L.R. Elasticity	Hausman test* (FE vs. RE)	S.R. Elasticity	L.R. Elasticity	Sargan Test**	First Stage F-Test***
Total Expenditure	0.09	0.61	28.7	0.09	0.47	0.01	428
	(.02)	(.15)	[.00]	(.03)	(.12)	[.91]	[.00]
Public Expenditure	0.13	0.89	24.5	0.13	0.66	0.02	492
	(.02)	(.18)	[.00]	(.03)	(.12)	[.90]	[.00]
Provincial Expenditure	0.13	0.95	13.0	0.13	0.67	0.01	549
	(.02)	(.20)	[.02]	(.03)	(.12)	[.91]	[.00]
Federal Expenditure	0.10	0.56	23.0	0.14	0.49	0.24	334
	(.05)	(.27)	[.00]	(.06)	(.18)	[.62]	[.00]
Hospital Expenses	0.16	1.32	43.4	0.18	0.90	0.10	585
	(.03)	(.29)	[.00]	(.03)	(.15)	[.75]	[.00]
Physician Expenses	0.11	1.54	158.9	0.14	1.19	0.40	792
	(.03)	(.45)	[.00]	(.03)	(.27)	[.53]	[.00]
Capital Expenses	0.09	0.31	9.25	0.14	0.31	0.03	145
	(.20)	(.66)	[.06]	(.23)	(.52)	[.86]	[.00]
Public Health Expenses	0.03	0.21	27.5	0.01	0.06	0.31	419
	(.06)	(.38)	[.00]	(.06)	(.28)	[.58]	[.00]

Note: A fixed effects estimate was used to regress health expenditure per capita on GDP per capita, the first lagged dependent variable, a time trend, the proportion of population aged 65 and older, and relative price of healthcare services to overall prices in the economy. The GMM estimates use the first lag of federal cash transfers to provinces and the second lag of the dependent variable as excluded instruments. Standard errors are robust and reported inside parentheses, and p-value of the test statistics are reported inside brackets. Number of observations is equal 310.

* The Hausman test was used to examine appropriateness of random effects (RE) estimate versus fixed effects (FE) estimate. Random effects estimate was rejected for all outocmes.

** The Sargan test was used to test validity of the instruments. I used the first lag of federal cash transfers and the second lag of the dependent variable as excluded instruments.

*** F-Test displays power of excluded instruments.

Table 4.6: Error correction model estimate of asymmetric effect of income on health expenditures where the threshold point is set at zero

	Symmetric vs. Asymmetric Error Correction Model*				GMM Estimate of Asymmetric Error Correction Model**			Asymmetric Error Correction Model without lagged dependent variable***		
	$\eta+\pi$	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test
Total Expenditure	1.09	1.08	1.05	.19	1.09	1.15	.79	1.09	1.08	.03
	(.02)	(.03)	(.05)	[.67]	(.06)	(.09)	[.38]	(.03)	(.04)	[.87]
Public Expenditure	1.11	1.11	1.11	.00	1.07	1.26	2.37	1.13	1.15	.04
	(.02)	(.03)	(.06)	[.97]	(.08)	(.08)	[.12]	(.05)	(.05)	[.84]
Provincial Expend.	1.11	1.11	1.12	.03	1.07	1.27	2.13	1.13	1.16	.09
	(.03)	(.03)	(.06)	[.86]	(.10)	(.10)	[.14]	(.05)	(.06)	[.77]
Federal Expend.	1.12	1.16	1.09	.40	1.09	1.09	.00	1.12	1.10	.04
	(.03)	(.06)	(.06)	[.54]	(.09)	(.09)	[1.0]	(.06)	(.06)	[.84]
Hospital Expenses	1.16	1.19	1.14	.26	1.32	1.23	.09	1.17	1.15	.04
	(.04)	(.05)	(.04)	[.62]	(.19)	(.13)	[.76]	(.06)	(.05)	[.85]
Physician Expenses	1.12	1.12	1.13	.01	1.04	1.12	.46	1.11	1.15	.21
	(.03)	(.05)	(.03)	[.92]	(.09)	(.05)	[.50]	(.06)	(.03)	[.66]
Capital Expenses	1.41	1.22	1.66	.62	7.95	.77	.28	.86	1.77	3.04
	(.18)	(.35)	(.29)	[.45]	(11.6)	(2.7)	[.60]	(.33)	(.29)	[.12]
Public Health Exp.	1.10	1.11	1.10	.01	1.26	1.12	.42	1.08	1.11	.05
	(.02)	(.06)	(.08)	[.92]	(.14)	(.09)	[.51]	(.06)	(.08)	[.83]

Note: $\eta+\pi$ denotes the sum of the short-term and long-term adjustments of health expenditure to changes in income in a symmetric regime as introduced in model (3.5), while $\eta_1+\pi_1$ and $\eta_2+\pi_2$ denotes the corresponding effects in an asymmetric regime during periods of economic strength and weakness, respectively as introduced in model (3.8). Standard deviation of estimate is inside parentheses and p-value of test statistic is inside brackets. Standard errors are robust. Total number of observations in the symmetric model is 310, and in the asymmetric model are 221 and 89 for periods of economic strength and weakness, respectively.

F-test performs the test of equality of the effects of income in the asymmetric regime: $H_0: \eta_1+\pi_1 = \eta_2+\pi_2$

* The model includes an error adjustment term and the first lag of the dependent variable, one for each regime, as covariates.

** The GMM estimate model includes an error adjustment term and the first lag of the dependent variable, one for each regime, as covariates. Four excluded instruments are used, including the second lag of the dependent variable and the first lag of federal cash transfers to provinces, both in their levels and one for each regime. The Sargan test did not reject validity of the instrument for any outcome (not reported for space reserving).

*** The model excludes the lagged dependent variables.

Table 4.7: Error correction model estimate of asymmetric effect of income on health expenditures where the threshold point is optimized by maximizing the likelihood function.

	Symmetric vs. Asymmetric Error Correction Model*					GMM Estimate of Asymmetric Error Correction Model**			Asymmetric Error Correction Model without lagged dependent variable***		
	$\eta+\pi$	τ	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test
Total Expenditure	1.09	.036	1.03	1.07	1.48	.90	1.25	1.84	1.05	1.09	1.23
(+N = 90 & –N =220)	(.02)		(.02)	(.02)	[.25]	(.12)	(.19)	[.17]	(.03)	(.02)	[.29]
Public Expenditure	1.11	.019	1.11	1.09	.16	.97	1.27	1.46	1.10	1.13	.33
(+N = 159 & –N =151)	(.02)		(.02)	(.04)	[.70]	(.11)	(.21)	[.23]	(.05)	(.03)	[.58]
Provincial expendiure	1.11	.036	1.08	1.11	.22	1.28	1.01	.08	1.12	1.14	.11
(+N = 90 & –N =220)	(.03)		(.04)	(.03)	[.65]	(.78)	(.47)	[.77]	(.05)	(.04)	[.75]
Federal Expenditure	1.12	.036	1.09	1.11	.11	1.11	1.23	.94	1.11	1.09	.04
(+N =.90 & –N =220)	(.03)		(.04)	(.04)	[.75]	(.15)	(.18)	[.33]	(.05)	(.05)	[.84]
Hospital Expenses	1.16	.036	1.17	1.13	.24	1.08	1.00	.39	1.15	1.14	.07
(+N = 90 & –N = 220)	(.04)		(.06)	(.03)	[.64]	(.24)	(.32)	[.53]	(.05)	(.03)	[.79]
Physician Expenses	1.12	.036	1.12	1.11	.02	.86	1.19	1.10	1.09	1.13	.40
(+N = 90 & –N =220)	(.03)		(.05)	(.02)	[.89]	(.19)	(.15)	[.29]	(.05)	(.03)	[.54]
Capital Expenses	1.42	.011	1.22	1.67	.82	4.22	1.56	.57	**.87**	**1.77**	**3.51**
(+N = 187 & –N =123)	(.18)		(.32)	(.26)	[.39]	(3.02)	(1.0)	[.45]	**(.31)**	**(.28)**	**[.09]**
Public Health Expen.	1.10	.019	1.11	1.10	.00	1.23	1.13	.42	1.08	1.10	.04
(+N =160 & –N =150)	(.02)		(.05)	(.06)	[.97]	(.11)	(.07)	[.52]	(.05)	(.06)	[.84]

Note: $\eta+\pi$ denotes the sum of the short-term and long-term adjustments of health expenditure to changes in income in a symmetric regime as introduced in model (3-5), while $\eta_1+\pi_1$ and $\eta_2+\pi_2$ denotes the corresponding effects in an asymmetric regime during periods of economic strength and weakness, respectively as introduced in model (3-8). Standard deviation of estimate is inside parentheses and p-value of test statistic is inside brackets. Standard errors are robust. τ denotes the threshold point, $+N$ and $-N$ denote number of observations in periods of economic strength and weakness, respectively. The number of observations is only reported for the estimates that include the lagged dependent variable.

F-test performs the test of equality of the effects of income in the asymmetric regime: $H_0: \eta_1+\pi_1 = \eta_2+\pi_2$

* The model includes an error adjustment term and the first lag of the dependent variable, one for each regime, as covariates.

** The GMM estimate model includes an error adjustment term and the first lag of the dependent variable, one for each regime, as covariates. Four excluded instruments are used, including the second lag of the dependent variable and the first lag of federal cash transfers to provinces, both in their levels and one for each regime. The Sargan test did not reject validity of the instrument for any outcome (not reported for space reserving).

*** The model excludes the lagged dependent variables.

Table 4.8: Error correction model estimate of asymmetric effect of income on health expenditures, when control for other covariates and the threshold point is optimized by maximizing the likelihood function.

	Symmetric vs. Asymmetric Error Correction Model*					IV Estimate of Asymmetric Error Correction Model**			Asymmetric Error Correction Model without lagged dependent variable***		
	$\eta+\pi$	τ	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test	$\eta_1+\pi_1$	$\eta_2+\pi_2$	F-test
Total Expenditure	1.09	.036	1.05	1.06	.09	.92	1.34	.96	1.07	1.08	.15
(+N = 90 & –N =220)	(.02)		(.03)	(.03)	[.76]	(.19)	(.37)	[.33]	(.03)	(.03)	[.70]
Public Expenditure	1.11	.019	1.12	1.08	.77	.99	1.29	1.79	1.12	1.13	.01
(+N = 159 & –N =151)	(.02)		(.03)	(.05)	[.40]	(.13)	(.18)	[.18]	(.05)	(.04)	[.93]
Provincial expenditure	1.11	.019	1.13	1.08	1.31	1.01	1.35	1.54	1.14	1.13	.02
(+N = 159 & –N =151)	(.03)		(.03)	(.05)	[.28]	(.18)	(.23)	[.22]	(.05)	(.04)	[.88]
Federal Expenditure	1.12	.019	1.13	1.06	1.29	1.14	1.16	.01	1.10	1.05	.69
(+N = 159 & –N =151)	(.03)		(.04)	(.03)	[.29]	(.05)	(.14)	[.91]	(.04)	(.04)	[.42]
Hospital Expenses	1.16	.036	1.20	1.11	2.33	1.06	.87	1.12	1.19	1.12	1.69
(+N = 90 & –N =220)	(.04)		(.05)	(.03)	[.16]	(.23)	(.29)	[.29]	(.05)	(.03)	[.23]
Physician Expenses	1.12	.036	1.12	1.07	1.80	.87	1.19	1.94	1.11	1.09	.31
(+N = 90 &–N =220)	(.03)		(.03)	(.02)	[.21]	(.16)	(.13)	[.16]	(.04)	(.03)	[.59]
Capital Expenses	1.42	.011	1.25	1.44	.22	3.25	1.13	.50	.96	1.55	2.17
(+N = 187 & –N =123)	(.18)		(.30)	(.19)	[.65]	(2.5)	(.79)	[.48]	(.28)	(.25)	[.17]
Public Health Expense	1.10	.019	1.10	1.08	.05	1.25	1.08	1.44	1.09	1.09	.01
(+N = 160 & –N =150)	(.02)		(.05)	(.04)	[.82]	(.12)	(.07)	[.23]	(.05)	(.04)	[.93]

Note: $\eta+\pi$ denotes the sum of the short-term and long-term adjustments of health expenditure to changes in income in a symmetric regime as introduced in model (3-5), while $\eta_1+\pi_1$ and $\eta_2+\pi_2$ denotes the corresponding effects in an asymmetric regime during periods of economic strength and weakness, respectively as introduced in model (3-8). Standard deviation of estimate is inside parentheses and p-value of test statistic is inside brackets. Standard errors are robust. τ denotes the threshold point, *+N* and *–N* denote number of observations in periods of economic strength and weakness, respectively. The number of observations is only reported for the estimates that include the lagged dependent variable.

F-test performs the test of equality of the effects of income in the asymmetric regime: $H_0: \eta_1+\pi_1 = \eta_2+\pi_2$

* The model includes an error adjustment term, the first lag of the dependent variable, a time trend, the proportion of the population aged 65 and older, and relative price index of healthcare services to CPI as covariates.

** The GMM estimate model includes an error adjustment term and the first lag of the dependent variable, one for each regime, as covariates. Four excluded instruments are used, including the second lag of the dependent variable and the first lag of federal cash transfers to provinces, both in their levels and one for each regime. The Sargan test did not reject validity of the instrument for any outcome (not reported for space reserving).

*** The model excludes the lagged dependent variables.

Appendix 4.A: Cross Section Dependence and Unit Root Test

Cross-section dependence in panels arises from different sources, such as health shocks or implementation of public policies. A recent approach to approximate the effect of cross-section dependence in unit root and cointegration tests is the Common Correlated Effects (CCE) approach advanced by Pesaran (2007).

Consider following estimate model

$$h_{it} = \alpha_i + \beta_i y_{it} + u_{it}, \qquad i = 1,...N; t = 1,...T \tag{4.A1}$$

where h_{it} is health expenditure per capita and y_{it} is GDP per capita in province i at time t. α_i is province-specific intercept and u_{it} is error term. Assume cross-section dependence arises from global shocks, in which

$$u_{it} = \gamma_i' f_t + v_{it} \tag{4.A2}$$

where f_t is the $m \times 1$ vector of unobserved common effects and v_{it} is province-specific effect. Coefficient γ_{ij} allows the effect of global shocks to be heterogeneous across provinces and shocks, $j = 1,...m$. Pesaran (2007) introduces a method to approximate effects f_t by the cross section average of the dependent and explanatory variables. In which model (4.A1) can be estimated using

$$h_{it} = \alpha_i + \beta_i y_{it} + g_i \overline{z}_t + e_{it} \tag{4.A3}$$

where $\overline{z}_t = (\overline{h}_t, \overline{y}_t)$ (the reader is referred to Pesaran (2007) for more details). Using this method, unit root test of health expenditure or GDP can be performed by an augmented Dickey-Fuller test where the cross section averages are added to the unit root test model

$$\Delta q_{it} = a_i + b_i q_{i,t-1} + c_i TT + \sum_{j=1}^{p} d_{ij} \Delta q_{i,t-j} + g_i' \overline{z}_t + e_{it} \tag{4.A4}$$

where q_{it} is either health expenditure, GDP or estimated residuals from model (4.A1), $\overline{z}_t = (\overline{q}_{t-1}, \Delta\overline{q}, \Delta\overline{q}_{t-1}, ..., \Delta\overline{q}_{t-p})'$, and TT is a time trend.

I used model (4.A4) to test for unit roots of health expenditure and GDP, with and without cross averages, and then calculated cross-section dependence to find out whether cross section averages were able to capture the effect of cross-dependence. I used two diagnostic tests for cross-dependence advanced by Pesaran (2004) and frees (1995), respectively given by

$$CP_p = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{ij}$$

and

$$CP_F = \sqrt{\frac{1}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (T \rho_{ij}^2 - 1)$$

where

$$\rho_{ij} = \frac{\sum_{i=1}^{T} q_{it} q_{jt}}{\left(\sum_{t=1}^{T} q_{it}^2\right)^{1/2} \left(\sum_{t=1}^{T} q_{jt}^2\right)^{1/2}},$$

ρ_{ij} is the pairwise correlation, N and T are cross unit and time dimensions, respectively.

Table 4.9 displays calculated cross-dependence for total health expenditure and GDP per capita time series, using residuals from estimate of model (4.A4). I estimated the model with the lag number that varies from 0 to 4, and found the cross-dependence persists even after incorporating the cross-section averages into the augmented Dickey-Fuller unit root test.

Table 4.9: Cross-section dependence and panel unit root test of Pesaran (2007)

		Without Cross Averages				With Cross Averages			
	Lags	Frees		Pesaran		Frees		Pesaran	
GDP	0	1.35	(.00)	13.9	(.00)	1.4	(.00)	13.8	(.00)
	1	1.19	(.00)	12.0	(.00)	1.26	(.00)	11.9	(.00)
	2	1.24	(.00)	12.1	(.00)	1.26	(.00)	11.6	(.00)
	3	1.36	(.00)	11.9	(.00)	1.37	(.00)	11.5	(.00)
	4	1.42	(.00)	11.8	(.00)	1.29	(.00)	10.8	(.00)
HE	0	1.15	(.00)	14.6	(.00)	1.15	(.00)	14.6	(.00)
	1	.70	(.00)	11.1	(.00)	.77	(.00)	10.5	(.00)
	2	.57	(.00)	10.3	(.00)	.66	(.00)	9.5	(.00)
	3	.66	(.00)	10.1	(.00)	.72	(.00)	9.5	(.00)
	4	.67	(.00)	9.7	(.00)	.78	(.00)	9.1	(.00)

Note: Figure in parentheses is p-value of the test statistic. The null hypothesis is that there is no cross dependence.

Appendix 4.B: Cointegration Test Using an Error Correction Model

I used an error correction approach to test the cointegration vector between health expenditure per capita and GDP per capita in a panel of the ten Canadian provinces. I used a test advanced by Westerlund (2007), which is based on structural rather than residual dynamics. Using an error correction model approach for unit root test, I would reject null of no cointegration if null of no error correction was rejected. Advantage of structural based tests, such as error correction approach, compared with residual based tests, is that they release assumption of equality between the long-term cointegrating vector and the shor-term adjustments that are implicitly assumed in residual based tests.

Consider following error correction model as introduced by Westerlund (2007)

$$h_{it} = \phi_{1i} + \phi_{2i}t + z_{it},$$
$$y_{it} = y_{i,t-1} + v_{it}, \qquad i = 1,...N; t = 1,...T \tag{4.B1}$$

For simplicity, y_{it} is modeled as a pure random walk and h_{it} consists of both deterministic term $\phi_{1i} + \phi_{2i}t$ and a stochastic term z_{it}. Conditional error correction model for h_{it} is given as

$$\Delta h_{it} = \delta_{1i} + \delta_{it}TT + \alpha_i(h_{i,t-1} - \beta_i y_{i,t-1}) + e_{it} \tag{4.B2}$$

If $\alpha_i < 0$, then there is an error correction, which implies that h_{it} and y_{it} are cointegrated. Westerlund (2007) proposed four panel test statistics that are based on this idea. Each pair of tests refers to different alternative hypothesis in the panel. That is, when $H_0 : \alpha_i = 0$ for all i is tested versus $H_1 : \alpha_i = \alpha < 0$ for all i, the first pair of test statistics are based on pooling information regarding the error correction along the cross-sectional dimension of the panel. When the null hypothesis is tested against $H_1^g : \alpha_i < 0$ for at least some i, the second pair of statistics refer to group mean statistics. To construct these test statistics, Westerlund proposes a multi-step method starting with estimate of following model

$$\Delta h_{it} = \delta_i d_t + \alpha_i h_{i,t-1} + \lambda_i y_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{ij}\Delta h_{i,t-j} + \sum_{j=0}^{p_i} \gamma_{ij}\Delta y_{i,t-j} + e_{it} \tag{4.B3}$$

where $d_t = (1, TT)'$. Following the multi-step estimate method (the reader is referred to Westerlund (2007) for more details), group mean statistics are given by

$$G_t = \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)} \text{ and } G_\alpha = \frac{1}{N}\sum_{i=1}^{N}\frac{T\hat{\alpha}_i}{\hat{\alpha}_i(1)}, \text{ where } \hat{\alpha}_i(1) = 1 - \sum_{j=1}^{p_i}\alpha_{ij}$$

and $SE(.)$ denotes conventional standard deviation.

The panel statistics are complicated by the fact that the both parameters and dimension of estimate model (4.B3) are allowed to differ between the cross-sectional units. A three-step method is proposed to estimate the panel statistics. The statistics are given by

$$P_t = \frac{\hat{\alpha}}{SE(\hat{\alpha})} \text{ and } P_\alpha = T\hat{\alpha} \text{ where } \hat{\alpha} = \left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{h}_{it}^2\right)^{-1}\sum\sum\frac{1}{\hat{\alpha}_i(1)}\tilde{h}_{i,t-1}\Delta\tilde{h}_{it}$$

For cross section dependence, Westerlund (2007) suggests a bootstrap method.

Table 4.10 displays cointegration test results between total health expenditure per capita and GDP per capita using Westerlund error correction model (estimate of cointegration between GDP and subcategories of health expenditure are not reported for space reserving). The estimates were derived using "xtwest" command in Stata software. "xtwest" command is not an official command in Stata. It is developed by Damiann and Westerlund in 2008. It can be downloaded using following web access: http://ideas.repec.org/c/boc/bocode/s456941.html.

The estimated results identify that one can not reject null of cointegration in all four tests, with and without a time trend and with the number of lags that varies from 0 to 4.

Table 4.10: Coingration between health expenditure and GDP using Westerlund (2007) error correction approach.

	Without Time Trend				With Time Trend			
Lags	G_t	G_α	P_t	P_α	G_t	G_α	P_t	P_α
0	-.46	-1.0	-1.6	-.95	-1.8	-6.1	-5.3	-5.6
	(.79)	(.88)	(.47)	(.50)	(.91)	(.94)	(.81)	(.87)
1	-.32	-.85	-.99	-.62	-2.2	-8.5	-6.1	-6.8
	(.84)	(.91)	(.61)	(.54)	(.58)	(.70)	(.65)	(.70)
2	-.53	-1.3	-1.9	-1.2	-2.3	-8.8	-7.1	-7.2
	(.63)	(.71)	(.34)	(.26)	(.32)	(.69)	(.21)	(.65)
3	-.60	-1.3	-2.3	-1.4	-2.3	-8.6	-6.7	-6.9
	(.56)	(.73)	(.31)	(.24)	(.28)	(.69)	(.23)	(.54)
4	-.66	-1.1	-2.2	-1.2	-2.5	-7.3	-6.1	-6.0
	(.52)	(.78)	(.31)	(.28)	(.11)	(.68)	(.18)	(.50)

Note: Figure in parentheses is bootstrap p-value, derived with 500 replications.
The null hypothesis is that the time series are cointegrated.

Bibliography

Ahn, S. C., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics*, 68, 5-27.

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76, 598-606.

Arellano, M. (1989). A note on the Anderson-Hsiao estimator for panel data. *Economics Letters*, 31, 337-341.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and application to employment equations. *Review of Economic Studies*, 58, 277-97.

Auld, M. C. (2005). Causal effect of early initiation on adolescent smoking patterns, *Canadian Journal of Economics*, 38, 709-734.

Bai, J., & Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, 72, 1127-1177.

Baker, M., & Milligan, K. (2005). How does job-protected maternal leave affect mothers' employment and infant health. *National Bureau of Economic Research*, Working Paper No. 11135.

Baker, M. & Milligan, K. (2007). Maternal employment, breastfeeding, and health evidence from maternity leave mandates. *National Bureau of Economic Research*, Working Paper No. 13188.

Baker, M., & Milligan, K. (2008). Evidence from maternity leave expansions of the impact of maternal care on early child development. *National Bureau of Economic Research*, Working Paper No. 13826.

Baker, M., Gruber, J., & Milligan, K. (2005). Universal childcare, maternal labor supply and family well-being. *National Bureau of Economic Research*, Working Paper No. 11832.

Bayder, N., & Brooks-Gunn, J. (1991). Effects of maternal employment and child care arrangement in infancy on preschoolers' cognitive development outcomes: evidence from the children of the NLSY. *Developmental Psychology*, 27, 918-31.

Becker, G.S., & Murphy, K. (1988). A theory of rational addiction. *Journal of Political Economy*, 96, 675–700.

Belsky, J. (1988). The effect of infant day care reconsidered. *Early Childhood Research Quarterly*, 3, 235-272.

Belsky, J., & Eggebeen, D. (1991). Early and extensive maternal employment/child care and 4-6 year old socioemotional development: children of the National Longitudinal Survey of Youth. *Journal of Marriage and Family*, 53, 1083-99.

Belsky, J. and Fearon, R. M. P. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: Does continuity in development depend on caregiving? *Attachment and Human Development*, 4, 361-387.

Berger, L. M., Hill, J., & Waldfogel, J. (2005). Maternal leave, early maternal employment and child health and development in the US. *The Economic Journal*, 115, 29-47.

Blau, R. H., & Grossberg, A. J. (1992). Maternal labor supply and children's cognitive development. *Review of Economics and Statistics*, 74, 474-81.

Blaylock, J.R., & Blisard, W. N. (1992). US cigarette consumption: The care of low-income women. *American Journal of Agricultural Economics*, 74, 689-705.

Blomqvist, A., & Carter, R. A. L. (1997). Is health care really a luxury. *Journal of Health Economics*, 16, 207-229.

Breitung, J. (2005). A parametric approach to the estimation of cointegration vectors in panel data. *Econometric Reviews*, 24, 151-174.

Breitung, J., & Pesaran, M. H. (2008). Unit root and cointegration in panels, in The Econometrics of Panels Data. Third Ed. Springer-Verlag Berlin Heidelberg.

Brook-Gunn, J., Han, W. J., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: the NICHD study of early child care. *Child Development,* 73, 1051-1072.

Caliendo, M., & Koneinig, S. (2005). Some practical guidance for the implementation of propensity score matching. Discussion Paper No. 1588, DIW Berlin and IZA Bonn, and University of Cologne.

Carrion-i-Silvestre, J. L. (2005). Health care expenditure and GDP: are they broken stationary. *Journal of Health Economics*, 24, 839-854.

Carrion-i-Silvestre, J. L., Barrio-Castro, T. D., & Lopez-Bazo, E. (2005). Breaking panels: An application of the GDP per capita. *Econometrics Journal*, 8, 159-175.

Cawley, J., Markowitz, S., & Tauras, J. (2003). Lighting up and slimming down: The effect of body weight and cigarette prices on adolescent smoking initiation. *National Bureau of Economic Research*, Working Paper No. 9561.

Cawley, J., Markowitz, S., & Tauras, J. (2004). Body weight, cigarette prices, youth access laws and adolescent smoking initiation. *Manuscript.*

Chaloupka, F. J., & Warner, K. E. (2000). The economics of smoking. In Handbook of Health Economics, Vol. 1B, edited by Joseph P. Newhouse and Anthony J. Culyer. Amsterdam: North-Holland.

Chaloupka, F. J., & Wechsler, H. (1997). Price, tobacco control policies and smoking among young adults. *Journal of Health Economics*, 16, 359–373.

Chan, K. S. (1993). Consistency and limiting distribution of the Least Square Estimator of threshold autoregressive model. *The Annals of Statistics,* 21, 520-533.

Chang, Y. (2002). Nonlinear IV unit root tests in panels with cross-sectional dependency. *Journal of Econometrics,* 110, 261-292.

Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics,* 120, 263-293.

Chang, Y. (2005). Residual based tests for cointegration in dependent panels. *Manuscript*, Rice University.

Chang, Y., & Park, J. (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis,* 24, 379-400.

Chia, Y. F. (2008). Maternal labour supply and childhood obesity in Canada: evidence from the NLSCY. *Canadian Journal of Economics,* 41, 217-42.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance,* 20, 249-272.

Coleman, J. S. (1988). Social capital and the creation of human capital. *American Journal of Psychology,* 94, s95-s120.

Culyer, A. J. (1988). Health expenditures in Canada: Myth and reality; past and future. *Canadian Tax Paper*, No. 82 (Canadian Tax Foundation, Toronto).

Davidson, R., & MacKinnon, J. G. (2004). Econometrics theory and methods. Oxford University Press.

DeCicca, P., Kenkel, D. S., & Mathios, A. D. (2002). Putting out the fires: Will higher taxes reduce the onset of youth smoking. *Journal of Political Economy,* 110, 144-169

Desai, S., Chase-Lansdale, P. L., & Michael, R. (1989). Mother or market? effects of maternal employment on cognitive development of four year old children. *Demography,* 26, 454-61.

Di Matteo, L. (2003). The income elasticity of health expend care spending. *European Journal of Health Economics*, 4, 20-29.

Di Matteo, L. (2009). Policy choice or economic fundamentals: what drives the public-private health expenditure balance in Canada. *Health Economics, Policy and Law*, 4, 29-53.

Di Matteo, L., & Di Matteo, R. (1998). Evidence on the determinants of Canadian provincial government health expenditures: 1965-1991. *Journal of Health Economics*, 17, 211-228.

Donohue, J. J., & Levitt, D. S. (2001). The impact of legalized abortion on crime. *The Quarterly Journal of Economics,* Vol. CXVI, Issue 2.

Dorlet, M. (2002). Wives, mothers, and wages: does timing matter. Ottawa, Ontario: Analytical Studies, Statistics Canada.

Douglas, S. (1998). The duration of the smoking habit, *Economic Inquiry*, 36, 49–64.

Douglas, S., & Hariharan, G. (1994). The hazard of starting smoking: estimates from a split population model. *Journal of Health Economics*, 13, 213–230.

Dreger,C., & Riemers, H. E. (2005). Health care expenditure in OECD countries: a panel unit root and cointegration analysis. *International Journal of Applied Economics and Quantitative Studies*, 2, 5-20.

Engle, R., & Granger, C. (1987). Co-integration and error correction: Representation, Estimation and Testing. *Econometrica*, 35, 251-276.

Evans, W., & Huang, L. (1998). Cigarette taxes and teen smoking: New evidence from panels of repeated cross-sections. *Manuscript*, University of Maryland.

Evans, W. N., Ringel, J. S., & Stech, D. (1999). Tobacco taxes and public policy to discourage smoking. *Tax Policy and the Economy*, 13, 1-55. Cambridge MIT Press.

Fogel, R. W. (2008). Forecasting the cost of US health care in 2040. *National Bureau of Economic Research*, Working Paper No. 14361.

Forster, M., & Jones, A. (2001). The role of tobacco taxes in starting and quitting smoking: duration analysis of British data. *Journal of the Royal Statistical Society*, 164, 517-547.

Forster, M. & Jones, A. M. (1999). The role of tobacco taxes in starting and quitting smoking: duration analysis of British data. *Manuscript*, University of York.

Freeman, D. G. (2003). Is health care a necessity or a luxury? Pooled estimates of income elasticity from US state-level data. *Applied Economics*, 35, 495-502.

Gagne, L. (2003). Parental work, child-care use and young children's cognitive outcomes. *Research Data Center Program*, Statistics Canada, Ottawa, Catalogue No. 9-594-XIE.

Garcia, J., & Labeaga, J.M. (1996). Alternative approaches to modeling zero expenditure: An application to Spanish demand for tobacco. *Oxford Bulletin of Economics and Statistics*, 58, 489-506.

Gbsemete, K., & Gerdtham, U. G. (1992). The determinants of health expenditure in Africa: a cross-sectional study. *World Development*, 20, 303-308.

Gerdtham, U.-G. & Lothgren, M. (2000). On stationarity and cointegration on international health expenditure and GDP. *Journal of Health Economics*, 19, 461-475.

Gerdtham, U. G., Sogaard, J. Anderson, F., & Jonsson, B. (1992). Econometric analysis of health expenditure: a cross-section study of the OECD countries. *Journal of Health Economics*, 11, 63-84.

Getzen, T. E. (2000). Health care is an individual necessity and a national luxury: Applying multilevel decision models to the analysis of health care expenditures. *Journal of Health Economics*, 19, 259-270.

Glied, S. (2002). Youth tobacco control: Reconciling theory and empirical evidence. *Journal of Health Economics*, 21, 117–135.

Goldberg, A. (1983). Abnormal selection bias. In Karlin, Samuel, Takeshi Amemiya and Leo Goldman. In Studies in Econometrics, Time Series, and Multivariate Statistics, New York: Academic Press.

Goldin, C. D. (1990). Understanding the gender gap: an economic history of American women. New York, NY: Oxford University Press.

Gottfried, A. E. and Gottfried, A. W. (1988). Maternal employment and children's development. New York: Plenum Press.

Gould, E. (2004). Decomposing the effects of children's health on mother's labor supply: is it time or money. *Health Economics,* 13, 525-541.

Greenstein, T. N. (1995). Are the 'most advantaged' children truly disadvantaged by early maternal employment? Effects on child cognitive outcomes. *Journal of Family Issues,* 16, 149-169.

Groen, J. J., & Kleibergen, F. (2003). Likelihood-Baed Cointegration analysis in panels of vector error-correction models. *Journal of Business and Economic Statistics*, 21, 295-318.

Gruber, J., & Koszegi, B. (2000). Is addiction rational? Theory and evidence. *National Bureau of Economic Research.* Working No. 7507.

Hamilton, V. H., Levinton, C., St-Pierre, Y., & Grimard, F. (1997). The effect of tobacco tax cuts on cigarette smoking in Canada. *Canadian Medical Association Journal.* 156, 187-191.

Han, W. J., Waldfogel, J., & Brooks-Gunn, J. (2001). The effects of maternal employment on later cognitive and behavioral outcomes. *Journal of Marriage and the Family,* 63, 336-54.

Hanratty, M., & Trzcinski, E. (2006). Who benefits from expanded paid family leave? impact of family leave expansion in Canada on mothers' employment after birth. *Manuscript*, University of Minneapolis and Wayne Sate University.

Hansen, P., & King, A. (1996). The determinants of health care expenditure: a cointegration approach. *Journal of Health Economics*, 15, 127-137.

Harris, J., & Chan, S. (1999). The continuum of addiction: cigarette smoking in relation to price among Americans aged 15–29. *Health Economics*, 8, 81–86.

Hartwing, J. (2008). What derives health care expenditure? – Baumol's model of 'unbalanced growth' revisited. *Journal of Health Economics*, 27, 603-623.

Harvey, E. (1999). Short-term and long-term effects of early maternal employment on children of the national longitudinal survey of youth. *Developmental Psychology*, 35, 445-59.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46, 1251-1238.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47, 153-161.

Heckman, J. J. (2007). The economics, technology and neuroscience of human capability formation. *National Bureau of Economic Research*, Working Paper No. 13195.

Heckman, J. J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market program. In Handbook of Labor Economics, vol. (3), ed. By O. Ashenfelter and D. Card, PP. 1865-2097: Elsevier, Ameterdam.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behavior. *National Bureau of Economic Research*, Working Paper No. 12006.

Heckman, J. J., Tobias, J. L., & Vytlacil, E. (2000). Simple estimators for treatment parameters in a latent variable framework with an application to estimating the returns to schooling. *National Bureau of Economic Research*, Working Paper No. 7950.

Hitiris, T., & Posnett, J. (1992). The determinant and effects of health expenditure in developed countries. *Journal of Health Economics*, 11, 173-181.

Hoffman, L. W. (1980). The effects of maternal employment on the academic attitudes and performance of school-aged children. *School Psychology Review*, 9, 319-335.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115, 53-74.

Jewell, T., Lee, J., Tieslau, M. & Strazicich, M. C. (2003). Stationarity and of health expenditure and GDP: Evidence from panel unit root tests with heterogeneous structural breaks. *Journal of Health Economics*, 22, 313-323.

Johansen, S. (1991). Estimation and hypothesis testing for cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59, 1551-1580.

Jones, A. M. (1989). The UK demand for cigarettes 1954-1986, a double-hurdle approach. *Journal of Health Economics*, 8, 133-141.

Jones, A. M. (1989b). A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics*, 4, 23-39.

Jonsson, K. (2005). Cross-sectional dependency and size distortion in a small-sample homogenous panel-data unit root test. *Oxford Bulletin of Economics and Statistics*, 63, 369-392.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, 90, 1-44.

Larsson, R., Lyhagen, L., & Lothgren, M. (2001). Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal*, 4, 109-142.

Laux, F.L. (2000). Addiction as a market failure: using rational addiction results to justify tobacco regulation. *Journal of Health Economics*, 19, 421–437.

Lazar, H., & St-Hilaire, F. (2004). Money, politics and health care. The Institute for Research on Public Policies (IRPP): Printed in Canada.

Lefebvre, P., & Merrigan, P. (2000). The effects of child-care arrangement on development outcomes of young children. *Paper prepared for the Conference on Families, Labour Market, and Well-being of Children*, University of British Columbia, Vancouver, Canada, June 1-2.

Lehrer, E. L. (1992). The impact of children on married women's labor supply: black-white differential revisited. *Journal of Human Resources,* 27, 422-444.

Levin, A., & Lin, C. F. (1993). Unit root test in panel data: asymptotic and finite-sample properties. *Manuscript* , University of California, San Diego.

Levin, A., Lin, C., & Chu, C. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108, 1-24.

Lewit, E., Coate, D., & Grossman, M. (1981). The effects of government regulation on teenage smoking. *Journal of Law and Economics*, 24, 545–573.

Li, H., & Maddala, G. S. (1996). Bootstrapping time series models. *Econometric Review,* 15, 115-158.

Lipps, G., & Yiptong-Avila, J. (1999). From home to school – how Canadian children cope. *Cultural, Tourism and the Center for Education Statistics,* Ottawa, Statistics Canada, Catalogue No. 89-F0117-XIE.

Lokshin, M. (2006). Semi-parametric difference –based estimation of partial linear regression models. *The Stata Journal*, 69, 377-383.

Lyness, K. S., & Judiesch, M. K. (2001). Are female mangers quitters? The relationship of gender, promotion, and family leaves of absence to voluntary turnover. *Journal of Applied Psychology,* 86, 1167-1178.

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, Special Issue 631-652.

McCoskey, S. K., & Sleden, T. M. (1998). Health care expenditure and GDP: Panel data unit root test results. *Journal of Health Economics*, 17, 369-376.

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, 305-26.

Moon, H. R., & Perron, B. (2005). Efficient estimation of the seemingly unrelated regression cointegration model and testing for purchasing power parity. *Econometric Review*, 23, 293-323.

Moon, R., & Perron, B. (2004). Testing for unit root in panels with dynamic factors. *Journal of Econometrics*, 122, 81-126.

Mroz, A. T., & Savage, H. T. (2001). The long-term effect of youth unemployment. *Manuscript*, Employment Policies Institute.

Neidell, M. (2000). Early parental time investment in children's human capital development: effects of time in the first year on cognitive and non-cognitive outcomes. Department of Economics, *University of California, Los Angeles*, Working Paper No. 806.

Neumark, D., & Nizalova, O. (2004). Minimum wage effect in the long-run. *National Bureau of Economic Research*, Working paper 10656.

Newhouse, J. P. (1977). Medical care expenditure: a cross-national survey. *Journal of Human Resources*, 12, 115-125.

Non-Smoker's Rights Association. (2003). Tobacco taxes and prices in Canada. Web access: http://www.nsra-adnf.ca/cms/index.cfm?group_id=1199.

O'Connell, P. (1998). The overvaluation of purchasing power parity. *Journal of International Economics*, 44, 1-19.

Olsen, R. J. (1980). A least square correction for selectivity bias. *Econometrica,* 48, 1815-1820.

Orphanides, A., & Zervos, D. (1995). Rational addiction with learning and regret. *Journal of Political Economy*, 103, 739–758.

Paarsch, H. J. (1984). A Monte-Carlo comparison of estimators for censored regression models. *Journal of Econometrics,* 67, 306-310.

Pagan, A., & Ullah, A. (1999). Nonparametric Econometrics. *Manuscript*, Cambridge University Press.

Parcel, T. L., & Menaghan, E.G. (1994). Early parental work, family social capital, and early childhood outcomes. *American Journal of Psychology,* 99, 972-1009.

Parkin, D., A. McGuire, & Yule, B. (1987). Aggregate health expenditure and national income: is health care a luxury good. *Journal of Health Economics*, 6, 109-127.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, 61, 653-670.

Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory*, 20, 597-625.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. CESifo, Working Paper No. 3032.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74, 967-1012.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics*, 22, 265-312.

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33, 311-340.

Phillips, P. C. B., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, 6, 217-259.

Phillips, P. C. B., & Perron, P. (1989). Testing for a unit root in time series regression. *Biometrika*, 75, 335-346.

Ram, B., Abada, T. & Hou, F. (2004). The effect of early maternal employment on children's cognitive outcomes: the Canadian experience. *Manuscript*, Statistics Canada and University of Western Ontario.

Roberts, J. (1999). Sensitivity of elasticity estimates for OECD health care spending: analysis of dynamic heterogenous data field. *Health Economics*, 8, 459-472.

Roberts, J. (2000). Spurious regression problems in the determinants of health care expenditure: a comment on Hitiris. *Applied Economics Letters*, 5, 279-283.

Robinovich, M., Wood, F., & Shemer, J. (2007). Impact of new medical technologies on health expenditures in Israel 2000-07. *International Journal of Technology Assessment in Health Care*, 23, 443-448.

Ruhm, C. (2000). Parental leave and child health. *Journal of Health Economics,* 19, 931-960.

Ruhm, C. (2004). Parental employment and child cognitive development. *The Journal of Human Resources,* XXXIX, 155-92.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26, 393-415.

Seshamani, M., & Gray, A. (2004a). Time to death and health expenditure: an improved model for the impact of demographic change on health care costs. *Age and Ageing*, 33, 556-561.

Seshamani, M., & Gray, A. (2004b). Aging and health care expenditure: a red herring argument revisited. *Health Economics*, 13, 302-314.

Studlar, D.T. (2002). Tobacco Control, comparative politics in the Unites States and Canada. *Broadview Press LTD.*

Suranovic, S. M., Goldfarb, R. S., & Leonard, T.C. (1999). An economic theory of cigarette addiction. *Journal of Health Economics*, 18, 1–29.

Tauras, J. A. (1999). The Transition to Smoking Cessation: Evidence from Multiple Failure Duration Analysis. *National Bureau of Economic Research*, No. 7412.

Tauras, J. A., & Chaloupka, F. J. (1999). Determinants of smoking cessation: an analysis of young adult men and women. *National Bureau of Economic Research*, Working Paper No. 7262.

Thomas, P. M. (2005). Comparing theories of child development. Belmont: Thomson Wadsworth press.

Tosetti, E., & Moscone, F. (2007). Health Expenditure and income in the United States. Working Paper No. 07/14, Department of Economics, University of Leicester.

Vandell, D. L., & Ramannan, J. (1992). Effect of early and recent maternal employment on children from low-income families. *Child Development,* 63, 938-49.

Wahab, M. (2004). Economic growth and governenment expenditure: evidence from a new test specification. *Applied Economics*, 36, 2125-2135.

Waldfogel, J., Han, W. J., & Brooks-Gunn, J. (2002). The effect of early maternal employment on child cognitive development. *Demography,* 39, 369-92.

Wang, Z., & Rettenmaier, A. J. (2007). A note on cointegration of health expenditures and income. *Health Economics*, 16, 559-578.

Warner, K. E., Chaloupka, F. J., Cook, P. J., Manning, W. G., Newhouse, J. P., Novotny, T. E., Schelling, T. C., & Townsend, J. (1995). Criteria for determining an optimal cigarette tax: The economist's perspective. *Tobacco Control*, 4, 380–386.

Westerlund, J. (2005). New simple tests for panel cointegration. *Econometric Reviews*, 24, 297-316.

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, 69, 709-748.

World Health Organization. (2000). Health aspects of maternity leave and maternity protection, statement to the International Labor Conference. June 2, 2000. Web access: http://www.who.int/reproductive-health/publications/maternal_mortality_2000/Health_aspects_of_maternity_leave.en.html.

Yatchew, A. (2003). Semiparametric regression for applied econometrician, *Cambridge University Press.*

Zick, C. D., Bryant, W. K., & Osterbacka, E. (2001). Mother's employment, parental involvement, and the implications for intermediate child outcomes. *Social Science Research,* 30, 25-49.

Zweifel, P., Felder, S., & Meiers, M. (1999). Ageing of population and health care expenditure: a red herring?, *Health Economics*, 8, 485-496.