Three Essays on Health and Labour Economics

by

Mina Alizadehsadrdaneshpour

A Thesis Submitted to the Faculty of Graduate Studies of
The University of Manitoba
in Partial Fulfillment of the Requirements of the Degree of

Doctor of Philosophy

Department of Economics
University of Manitoba
Winnipeg, Manitoba

DOCTOR OF PHILOSOPHY (2020)	University of Manitoba
(Economics)	Winnipeg, Manitoba
TITLE:	Three Essays on Health and Labour Economics
AUTHOR:	Mina Alizadehsadrdaneshpour M.A. Azad University, Science and Research Branch, Tehran, Iran
SUPERVISORS:	Dr. Julia Witt (Advisor)
	Dr. Umut Oguzoglu (Co-Advisor)
	Dr. Evelyn Forget (Internal Examiner)
	Dr. Mahmoud Torabi (Internal Examiner)
	Dr. Mohammad Hajizadeh (External Examiner)
NUMBER OF PAGES:	xii, 205

Abstract

The dissertation consists of three essays. The first essay examines a dynamic effect of diabetes on employment. The second essay uses a broader measure of health and investigates its interaction with employment. The third essay explores a dynamic effect of education on hourly wage.

The first essay investigates the diabetes effect on employment in Canada. The data is taken from the National Population Health Survey and men and women between age 25 and 64 are analysed separately. In contrast to the previous static studies on the effect of diabetes on labour market outcomes, this essay uses a dynamic model to identify the impacts of diabetes on employment in Canada. Results show that diabetes has a positive but insignificant effect on employment for men. The effect of diabetes on employment for women is negative and significant. The results confirm the signs and significance of diabetes coefficients estimated by static studies; however, the numbers are much smaller. Particularly, precise estimates of diabetes effect on employment would be helpful for policy makers to know the economic burden and design the appropriate policies.

The second essay uses a broader measure of health to explore the relationship with employment. In contrast to previous static Canadian studies on the impact of health on labour market outcomes, this essay estimates a dynamic model using simultaneous equations to obtain more precise model specification for the interaction of health and labour market outcome. Results show that there is a high state dependency in employment and health for both men and women. Moreover, there is a highly significant and positive effect of health on employment for both men and women. As a result, health policies that have positive and direct effects on health can have positive and indirect effects on employment.

The third essay investigates the return to education using a dynamic approach in Canada. This essay uses the longitudinal Survey of Labour and Income Dynamics and estimates the dynamic model for men and women between the age of 25 and 64 separately. In contrast to the previous static Canadian studies on the return to education, this essay estimates a dynamic Mincer model through a system-GMM method to obtain more precise model specification for the return to education. Results demonstrate that the hourly wage is highly persistent for both men and women between age 25 and 64. The results also show that the return to schooling is increasing at the beginning of the working life for both men and women compared to a constant return to schooling by static Mincer function. Identifying the return to education can be useful for policy makers to decide on education expenditures and finance schooling programs.

Acknowledgment

I would like to express my deep and sincere gratitude to my thesis advisors Dr. Julia Witt and Dr. Umut Oguzoglu for the continuous support of my PhD study and research, for their patience, motivation, and immense knowledge. I always keep in my heart their support and patience during the recent years I spent undertaking my research in Toronto and also the time I spent at Winnipeg. Without their guidance and constant feedback this PhD would not have been achievable. I would like to thank my internal examiners, Professor Evelyn Forget, and Professor Mahmoud Torabi for their insightful comments and valuable suggestions which have contributed greatly to the improvement of the thesis.

I would like to thank my external examiner, Dr. Hajizadeh, for his constructive suggestions and valuable feedback. I am also grateful to Betty McGregor for her continuous administrative support and assistance. Many thanks to the staff and faculty of the Department of Economics at the University of Manitoba for their various support during my program.

The Statistics Canada Research Data Centers at the University of Manitoba and University of Toronto are acknowledged. I was permitted the access to confidential longitudinal micro data- the Longitudinal National Population Health Survey and the Survey of Labor and Income Dynamics. Financial support from my advisor through Graduate Enhancement of Tri-Council Stipends is also thankfully acknowledged. I would also thank to Manitoba RDC analyst, Dr. Ian Clara and Toronto RDC analyst, Dr. Angela Prencipe. All of you have been there to support me when I collected data for my PhD thesis. Without their precious support it would not be possible to conduct this research.

Last but not the least, I would like to say a special thanks to my family. Words cannot express how grateful I am to my parents for all of the sacrifices that you've made on my behalf. I

would also like to thank my brother and sister who incented me to strive towards my goal and supported me spiritually throughout writing this thesis and my life in general.

DEDICATION

To my parents, for their endless love, support and encouragement, and
to my brother and sister who have never left my side.

Table of Contents

List of Tables and Figures

Chapter 1. Introduction

The main focus of this dissertation is to show the importance of dynamic investigation in analyzing the effect of health outcomes/education on labour outcomes (employment/ hourly wage). In this regard, the first essay examines a dynamic effect of diabetes on employment by a dynamic probit model in Canada. The second essay extends the dynamic analysis and uses a broader measure of health (self-reported health) and investigates its interaction with employment. However, this time a simultaneous equation model has been used instead of an instrumental variable approach which gives more efficient estimates. The third essay explores a dynamic effect of education on hourly wage through a system generalized method of moments (GMM) approach in order to control for endogeneity of education along with the persistency of hourly wage.

The first essay is focused on diabetes effect on employment in Canada. In this regard, the data is taken from 9 cycles of NPHS (National Population Health Survey) and men and women between the age 25 and 64 are analysed separately. According to the International Diabetes Federation, about 463 million adults (20-79 years) were diagnosed with diabetes in 2019 and this will increase to 700 million by 2045. In 2017, 7.3% of Canadians aged 12 and older (about 2.3 million) reported being diagnosed with diabetes (Statistics Canada, 2018). Diabetes and its complications can lead to direct and indirect costs. Total health care costs related to diabetes over a 10-year period (2011/12 to 2021/22) for the Canadian aged 20 years and over were $15.36 billion (Anja and Laura, 2017). Apart from the direct costs of diabetes, such as hospitalization costs, specialist costs, and medication costs (Canadian Diabetes Association, 2009), there are indirect costs related to diabetes as well. Individuals with diabetes may not work as effectively as

they did before they had diabetes-related complications. There are studies which investigate the effect of diabetes on labour market outcomes (Kahn, 1998; Ng et al., 2001; Kraut et al., 2001; Bastida and Pagan, 2002; Tunceli et al., 2005; Brown et. al. 2005; Latif, 2009; Harris, 2009; Lin, 2011; Minor, 2011; Seuring et al., 2015). However, the previous studies are all static or cross-sectional and they have not considered the dynamic effects of diabetes on labour market outcomes. Since employment status at previous period has an important role on employment status at current time, this essay considers the dynamic nature of employment and uses a dynamic panel probit model to identify the impacts of diabetes on employment in Canada. As the prevalence and incidence of diabetes is increasing, it is important to have accurate estimates of the labour market costs of diabetes. Particularly, a precise estimates of diabetes effect on employment would be helpful for policy makers, government or employers to know the economic burden and design the appropriate policies to reduce related costs.

The second essay have extended the dynamic analysis in the first essay but using a broader measure of health (self-reported health). Particularly, this time, the relation between health and employment has been investigated. In this regard, health and employment are specified jointly through a simultaneous equation model to account for unobserved characteristics which might jointly affect them. Health status is considered an important factor in determining labor supply decisions (Stern, 1989; Bound et al., 1999; Currie and Madrian, 1999; Au et al., 2005; Kalwij and Vermeulen, 2008). Health status can influence individuals' income. Individuals can not apply for paid employment or lose their job once they become ill or retire early (Martikainen et al., 2003). On the other hand, income can affect health and longevity through different mechanism such as

clinical, behavioral, social, and environmental factors (Khullar and Chokshi, 2018). Health may affect individual's educational attainments or outcomes. For instance, individual's health status measured by the probability of sickness, can significantly affect academic success (Gan and Gong, 2007). On the other hand, education may determine health as well. Particularly, higher educational levels positively affect self-reported health and reduce the number of chronic conditions (Lundborg, 2008). Improvement in well-being can result in improved workplace performance such as profitability, labor productivity and the quality of outputs or services (Bryson, Forth and Stokes, 2014). Although health is fundamental to well-being, work acts as an important determinant of health as well. Particularly, individuals who are employed are healthier than those who are not employed (Yelin and Trupin, 2003; Thomas and Ellis, 2013; McKee-Ryan et al., 2005) and employment can have social, psychological, and financial benefits that improve health (Goodman, 2015).

Canadian studies that have investigated the impact of health on labour market outcomes are cross-sectional or static (Breslaw and Stelener, 1987; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have not examined the dynamic interaction between health and labour market outcomes. Moreover, they have solved the endogeneity of health through a two-stage method or instrumental variable (IV) approach which does not give an efficient estimate for health effect. Thus, the second essay, contributes to the Canadian literature by estimating a dynamic panel data model through a simultaneous equation model to obtain more precise model specification for the interaction of health and labour market outcome. Understanding the relationship between health and labor outcomes is even more important as more individuals reach the life stage where their

health status gradually depreciates to the point where it can influence their labor market behavior (Currie and Madrian, 1999). Furthermore, a good understanding of how health influences labor force participation is also important for policy makers to understand the relationship between health and labor outcomes and apply more efficient policies to improve health status of labor market participants.

The third essay investigates the return to education using a dynamic approach in Canada. In this regard, it uses the longitudinal Survey of Labour and Income Dynamics (SLID) and estimates the dynamic model for men and women between the age of 25 and 64 separately. Education accounts as an important determinant of labour outcomes. Individuals with higher degrees have higher chance to be employed. For instance, post-secondary graduates are more likely to be employed and they earn more than those who do not continue their studies past high school (Berger and Parkin, 2009). In 2010, about one in four (24.1%) of individuals with university degree were in top 10% of income recipients compared to one in ten of those with a postsecondary certificate or diploma below bachelor's degree (Statistics Canada, 2016).

Canada has high completion rates in postsecondary education. In 2012, 57.3 per cent of Canadian aged 25-34 had either a college or university certificate which was the third highest among OECD members (Parliamentary Budget Officer, 2016). Tuition costs is increasing rapidly across most of the country. On average, undergraduate pay 40 percent more in tuition than they did 10 years ago (Statistics Canada, 2016). According to a study by Van Pelt and Emes in 2015, during the period 2001/2 to 2011/12, the overall nominal public spending in Canada on government schools increased from $38.9 billion to $59.6 billion. In addition, education spending per pupil in government schools also had

a growth of 63.2 percentage points ($7,250 to $11,835) per pupil. In 2004-5, total federal spending in postsecondary education was about $8.7 billion and it increased steadily to more than $12.8 billion in 2010-11 (PBO, 2016).

The human capital theory proposed by Becker (1964) has been widely used to analyze the wage differentials among individuals with differences in their human capital such as education and experience. According to this theory, education is as an investment of knowledge or skills which can increase productivity of individuals and result in higher incomes. There is an extensive literature on using the Mincer earning function as a framework for measuring return to education (Card, 1999; Heckman et al. 2003; Trostel, 2005; Lemieux, 2006; García-Suaza et al. 2014; Hartog and Gerritsen, 2016; Mazza et al. 2018). According to the standard Mincer equation proposed by Mincer (1974), the observed wage is explained by schooling years, potential labour market experience and potential labour-market experience squared. The Canadian studies that have investigated the return to education are static or cross-sectional (Murphy, Riddell & Romer; 1998; Burbidge, Magee & Robb, 2002; Lemieux, 2006; Oreopoulos, 2006; Boudarbat, Lemieux & Riddell, 2010; Bourbeau, Lefebvre & Merrigan, 2012; Lemieux, 2014; Ren and Shannon, 2017; Aydede and Dar, 2017) and they have not estimated the return to education through a dynamic approach. In other words, they have not taken into account earnings persistence in estimating the return to education. Moreover, they have used an instrumental variable approach (IV) to solve the endogeneity of education. This study contributes to the Canadian literature by estimating a dynamic Mincer model through a system GMM method to obtain more precise model specification for the return to education. As individuals consider education as an investment for their future, thus they

decide to continue attending colleges or universities if the benefits of higher education exceed of their tuition costs. Therefore, identifying the return to education can help individuals on investment decision purposes. Furthermore, the return to education can be useful for policy makers who decide on education expenditures and prioritize schooling levels and finance schooling programs.

In brief, the above-mentioned essays all measuring the effects on labour outcomes using a dynamic investigation to show the importance of persistency of labour market outcomes.

References

Au, D. W. H., Crossley, T. F., & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health economics*, *14*(10), 999-1018.

Aydede, Y., & Dar, A. (2017). Is the lower return to immigrants' foreign schooling a postarrival problem in Canada?. *IZA Journal of Migration*, *6*(1), 4.

Bastida, E., & Pagán, J. A. (2002). The impact of diabetes on adult employment and earnings of Mexican Americans: findings from a community based study. *Health economics*, *11*(5), 403-413.

Becker, G. (1964). Human Capital, Nueva York. *National Bureau of Economic Research*, *8*(299), 493-517.

Berger, J., & Parkin, A. (2009). The value of a degree: Education, employment and earnings in Canada. *The price of knowledge: Access and student finance in Canada*, 7-22.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics*, *6*(2), 179-202.

Bourbeau, E., Lefebvre, P., & Merrigan, P. (2012). The Evolution of the Returns to Education for 21-to 35-Year-Olds in Canada and across Provinces: Results from the 1991–2006 Analytical Census Files. *Canadian Public Policy*, *38*(4), 531-549.

Boudarbat, B., Lemieux, T., & Riddell, W. C. (2010). The evolution of the returns to human capital in Canada, 1980–2005. *Canadian Public Policy*, *36*(1), 63-89.

Breslaw, J. A., & Stelcner, M. (1987). The effect of health on the labor force behavior of elderly men in Canada. *Journal of Human Resources*, 490-517.

Bryson, A., Forth, J., & Stokes, L. (2014). Does worker wellbeing affect workplace performance. *Department of Business Innovation and Skills, London*.

Burbidge, J. B., Magee, L., & Robb, A. L. (2002). The education premium in Canada and the United States. *Canadian Public Policy/Analyse de politiques*, 203-217.

Canadian Diabetes Association. (2009). An economic tsunami: The cost of diabetes in Canada. *Toronto, ON: Canadian Diabetes Association*.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, *3*, 3309-3416.

Gan, L., & Gong, G. (2007). *Estimating interdependence between health and education in a dynamic model* (No. w12830). National Bureau of Economic Research.

García-Suaza, A. F., Guataquí, J. C., Guerra, J. A., & Maldonado, D. (2014). Beyond the Mincer equation: the internal rate of return to higher education in Colombia. *Education Economics*, *22*(3), 328-344.

Goodman, N. (2015). The impact of employment on the health status and health care costs of working-age people with disabilities. *Lead Center Policy Brief*.

Harris, A. (2009). Diabetes, cardiovascular disease and labour force participation in Australia: An endogenous multivariate probit analysis of clinical prevalence data. *Economic Record*, *85*(271), 472-484.

Hartog, J., & Gerritsen, S. (2016). Mincer Earnings Functions for the Netherlands 1962–2012. *De Economist*, *164*(3), 235-253.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2003). *Fifty years of Mincer earnings regressions* (No. w9732). National Bureau of Economic Research.

Hum, D., Simpson, W., & Fissuh, E. (2008). The Impact of Health on Labour Supply over the Life Cycle. *Canada: Department of Economics, The University of Manitoba*.

International Diabetes Federation. (2019). Diabetes facts and figures. Retrieved from https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html.

Kahn, M. E. (1998). Health and labor market performance: the case of diabetes. *Journal of Labor Economics*, *16*(4), 878-899.

Kalwij, A., & Vermeulen, F. (2008). Health and labour force participation of older people in Europe: what do objective health indicators add to the analysis?. *Health economics*, *17*(5), 619-638.

Khullar, D., Chokshi, D.A. (2018) .*Health, Income, & Poverty: Where We Are & What Could Help.* Retrieved from https://www.healthaffairs.org/do/10.1377/hpb20180817.901935/full/

Kraut, A., Walld, R., Tate, R., & Mustard, C. (2001). Impact of diabetes on employment and income in Manitoba, Canada. *Diabetes Care*, *24*(1), 64-68.

Latif, E. (2009). The impact of diabetes on employment in Canada. *Health economics*, *18*(5), 577-589.

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer a pioneer of modern labor economics* (pp. 127-145). Springer, Boston, MA.

Lemieux, T. (2014). Occupations, fields of study and returns to education. *Canadian Journal of Economics/Revue canadienne d'économique*, *47*(4), 1047-1077.

Lin, S. J. (2011). Estimating the impact of diabetes on employment in Taiwan. *Econ Bull*, *31*(4), 3089-102.

Lundborg, P. (2008). The Health Returns to Education-What Can We Learn from Twins?. *Available at SSRN 1113685*.

Martikainen, P., Adda, J., Ferrie, J. E., Smith, G. D., & Marmot, M. (2003). Effects of income and wealth on GHQ depression and poor self rated health in white collar women and men in the Whitehall II study. *Journal of Epidemiology & Community Health*, *57*(9), 718-723.

Mazza, A., Battisti, M., Ingrassia, S., & Punzo, A. (2018, April). Modeling Return to Education in Heterogeneous Populations. In *PAA 2018 Annual Meeting*. PAA.

McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology*, *90*(1), 53.

Mincer, J. (1974). Schooling, Experience and Earnings. New York: Columbia University Press.

Minor, T. (2011). The effect of diabetes on female labor force decisions: new evidence from the National Health Interview Survey. *Health economics*, *20*(12), 1468-1486.

Murphy, K. M., Riddell, W. C., & Romer, P. M. (1998). *Wages, skills, and technology in the united states and canada* (No. w6638). National Bureau of Economic Research.

Ng, Y. C., Jacobs, P., & Johnson, J. A. (2001). Productivity losses associated with diabetes in the US. *Diabetes care*, *24*(2), 257-261.

Oreopoulos, P. (2006). The compelling effects of compulsory schooling: Evidence from Canada. *Canadian Journal of Economics/Revue canadienne d'économique*, *39*(1), 22-52.

Parliamentary Budget Officer (PBO) (2016). Federal Spending on Postsecondary Education. Retrieved from http://www.pbo dpb.gc.ca/web/default/files/Documents/Reports/2016/PSE/PSE_EN.pdf.

Ren, F., & Shannon, M. (2017). Cohort Effects and the Returns to Canadian Undergraduate Education. *Canadian Public Policy*, *43*(3), 221-245.

Seuring, T., Goryakin, Y., & Suhrcke, M. (2015). The impact of diabetes on employment in Mexico. *Economics & Human Biology*, *18*, 85-100.

Shelton Brown III, H., Pagán, J. A., & Bastida, E. (2005). The impact of diabetes on employment: genetic IVs in a bivariate probit. *Health economics*, *14*(5), 537-544.

Statistics Canada. (2016). Education and occupation of high-income Canadians. Statistics Canada Retrieved from http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-014-x/99-014-x2011003_2-eng.cfm.

Statistics Canada. (November 14,2018). Diabetes, 2017.Retrieved from https://www150.statcan.gc.ca/n1/pub/82-625-x/2018001/article/54982-eng.htm.

Statistics Canada. (2016). Education and occupation of high-income Canadians. Statistics Canada Catalogue no. 99-014-X2011003. Last modified: 2016-09-15. http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-014-x/99-014-x2011003_2-eng.cfm.

Statistics Canada. (2016). Tuition fees for degree programs, 2016/17. Retrieved from http://www.statcan.gc.ca/daily-quotidien/160907/dq160907a-eng.htm.

Stern, S. (1989). Measuring the effect of disability on labor force participation. *Journal of human Resources*, 361-395.

Thomas, K. C., & Ellis, A. R. (2013). Patterns of healthcare use and employment among people with disabilities. *Disability and Health Journal*, *6*(2), 133-140.

Trostel, P. A. (2005). Nonlinearity in the return to education. *Journal of Applied Economics*, *8*(1), 191-202.

Tunceli, K., Bradley, C. J., Nerenz, D., Williams, L. K., Pladevall, M., & Lafata, J. E. (2005). The impact of diabetes on employment and work productivity. *Diabetes care*, *28*(11), 2662-2667.

Van Pelt, D. N., & Emes, J. (2015). *Education Spending in Canada: What's Actually Happening?*. Fraser Institute.

Yelln, E. H., & Trupin, L. (2003). Disability and the characteristics of employment. *Monthly Lab. Rev.*, *126*, 20.

Chapter 2. A Dynamic Analysis of Diabetes and Employment in Canada

Abstract

This study investigates the dynamic effects of diabetes on employment in Canada using data from the 9 cycles of NPHS (National Population Health Survey), taking into account the possible endogeneity of diabetes through a recursive bivarite probit model. Previous studies have mostly used static or cross-sectional panel data models, but since employment status in the previous period has an important role on employment status in the current period, this study will consider a dynamic process suggested by Wooldridge (2005) to identify the impacts of diabetes on employment. Considering the dynamic process of employment, results show that diabetes has insignificant positive effect on employment status of men while women face a negative significant effect of diabetes on their employment status. Furthermore, the diabetes effect from the dynamic model is much smaller than those estimated by the static models. In fact, identifying the true relationship between diabetes and employment can be more helpful for policy makers to know the economic burden and design the appropriate policies to reduce related costs.

Keywords: diabetes, employment, dynamic panel data model

2.1. Introduction

The number of people living with diabetes is increasing rapidly due to increasing aging population, rising obesity rates and increasing sedentary lifestyles (Canadian Diabetes Association, 2009). Diabetes is a chronic disease that occurs either when the pancreas does not produce insulin or when the body cannot properly use the insulin it produces (World Health Organization). While we eat, our body turns food into sugars (glucose). Insulin is a hormone produced by the pancreas, which allows our body to use the glucose for energy. There are three main type of diabetes: type 1, type 2 and gestational diabetes. About 5 to 10 per cent of people with diabetes have type1 diabetes (Canadian Diabetes Association). Type 1, or insulin- dependent, diabetes occurs mostly among children and young adults, but it can develop in adulthood. In type 1 diabetes, the pancreas does not produce any insulin or insufficient insulin, which subsequently leads to an increase in blood sugar, rather than the sugar being used for energy. In other words, type 1 diabetes is a disorder of the immune system, which destroys the beta cells of the pancreas that release insulin (Diabetes Canada). Type 2, or insulin-independent, -diabetes is the most common form of diabetes. Type 2 diabetes accounts for about 90% of all individuals with diagnosed diabetes (International Diabetes Federation). In type 2 diabetes, the pancreas does not produce enough insulin, or the body cannot effectively use the insulin it makes (called insulin resistance), which leads to high blood sugar in our body. A third type, gestational diabetes, can develop during pregnancy. This type of diabetes is temporary and goes away after the baby is born. Between 3 and 20 percent of pregnant women have gestational diabetes (Canadian Diabetes Associations). In women with gestational diabetes, the body makes hormones that lead to insulin residence.

The prevalence of diabetes is increasing rapidly. According to the International Diabetes Federation, in 2015, 415 million adults had diabetes and this will increase to 642 million by 2040. Reports from the World Health Organization, estimated that in 2014, 9% of adults 18 years and older had diabetes. Wild et al. (2004) estimated the prevalence of diabetes in all age groups in the world to be 2.8% in 2000 and 4.4% in 2030. The Canadian chronic disease surveillance system reported that in 2008/09, approximately 2.4 million Canadian (6.8%) age one year and older were living with diagnosed diabetes (type 1 and type 2) and for adults age 20 and older the prevalence rate reaches 8.7% (1 in 11 Canadians) (Public Health Agency of Canada, 2011). The Canadian Diabetes Association estimated the prevalence of diabetes in Canada in 2015 and made predictions for 2025. They estimated that diabetes prevalence will experience an increasing trend of 44 percent during the 10 years time horizon (3.4 million to 5 million).

Diabetes and its complications can lead to direct and indirect costs. Logminiene et al. (2003) estimated direct and indirect cost of diabetes in the world. They found that the direct cost of diabetes increased from 1.7 billion US dollars in 1969 to 44.4 billion dollars in 1997; and the indirect costs increased 33 times, from 1.6 billion US dollars in 1969 to 54.1 billion US dollars in 1997. Apart from the direct costs of diabetes, such as hospitalization costs, specialist costs, and medication costs (Canadian Diabetes Association, 2009), there are indirect costs related to diabetes as well. People who have diabetes are at higher risk of serious health problems. Over time, complications develop and if the disease is not properly controlled, high blood sugar can cause life threatening and disabling complications such as heart disease, stroke, kidney failure, eye disease (diabetic retinopathy), and nerve damage (known as neuropathy). Most people with

diabetes cannot work as effectively as they did before they had diabetes-related complications (Vijan et al. 2004; DiBonaventura et al. 2011). Diabetes can affect employment through different ways. For instance, people with diabetes can have reduced productivity at work. This lost productivity may manifest through absenteeism (time lost from work due to sickness), presenteeism (less productive time at work due to illness) or early retirement. Individuals with diabetes may also face discrimination at workplace.

Vijan et al. (2004) investigated the economic burden of related diabetes complications on workforce participation. Results showed that diabetes is a significant predictor of lost productivity. In the US population of diabetic adults aged 51-61 in 1992, the diabetes related complications increase in risk of disability, mortality, sick days and retirement caused an incremental $58.6 billion in lost productivity. Particularly, $4.4 billion in lost income due to early retirement, $0.5 billion due to increased sick days, $31.7 billion due to disability, and $22.0 billion in lost income due to premature mortality. Using data from the GAZ and ELectricité (GAZEL)[1] cohort study, Herquelot et al. (2011) found a strong negative impact of diabetes on workforce participation in France. Results suggested that the risk of work disability and early retirement among diabetes of GAZEL cohort increased 60-70%. Moreover, diabetes had a sevenfold increase in the risk of death while in labor force. The risk of work cessation for people with diabetes increased too. For instance, people with diabetes aged 35-60 lost an estimated mean time of 1.1 year in the workforce.

[1] The GAZ and ELectricité (GAZEL) cohort was established in 1989 among employees of the French national electricity and gas company "EDF-GDF." In January 1989 (Herquelot et al. 2011).

DiBonaventura et al. (2011) investigated the impact of a special kind of diabetic complication named painful diabetic peripheral neuropathy (pDPN) on productivity. Results demonstrated that individuals with pDPN experienced significantly higher levels of absenteeism, presenteeism, overall work impairment and activity impairment compared to those with diabetes but without pDPN and the control group (people without diabetes).As the complications of diabetes appear gradually during their lifetime, patients with diabetes for longer duration may face early retirement from workforce. Using Danish Pension schemes, Cleal and Poulsen (2015) investigated the relationship between diabetes and retirement from workforce. They examined the contributions of diabetes cohorts before and after entering a normal retirement pension, voluntary early retirement or disability retirement. Results indicated that the highest proportion of diabetes was identified before the normal retirement. Particularly for people with diabetes, voluntary early retirement pension was the most common way of exit from workforce and individuals with prevalent diabetes had a higher relative risk of disability compared to the ones without this disease (2.98 vs. 2.40). People living with diabetes may face discrimination in the workplace and they worry that their conditions preventing them from getting the job they want.

Matsushima et al. (1993) examined the social and economic backgrounds of the youth-onset insulin-treated diabetes mellitus in Japan. The results suggested that individuals with diabetes were more likely to face job refusal in their lives compared to the sibling control groups (20 vs. 0%). In fact, most patients (6/7) who experienced a job refusal, told job interviewers about their disease. Robinson et al. (1990) examined a survey of employment problems in a random sample of diabetic patients aged 17-65 in

eight centers in the UK. Results demonstrated that 13% of patients with diabetes had difficulties in getting a job. Nine percent of diabetic patients had to change their job due to their illness and 7% of people with diabetes lost a job due to their illness. Furthermore, diabetic shift workers were twice as likely as their control groups (shift workers without diabetes) to face problems with their job (18 vs. 8%, p= 0.045).

2.2. Literature Review

The previous studies have examined the effect of diabetes on employment, and they mostly assumed that diabetes is an exogenous variable in their regression analysis. Kahn (1998) used the 1976 and 1989 waves of the National Health Interview Survey (NHIS) and the first wave of the Health and Retirement Survey (HRS) to analyze diabetic labor market performance. Results showed that between 1976 and 1992, diabetic women experienced significant improvements in their labor force participation relative to non-diabetic women while diabetic men have decreased their participation over time. Ng et al. (2001), using data from the 1989 US National Health Interview Survey, concluded that diabetes had a considerable impact on labor market outcomes. By controlling for demographic variables such as age, sex, and health status, they found that the presence of diabetes itself reduced employment by 3.5%, and the presence of complications reduced employment by 12% (compared to the absence of complications).

Kraut et al. (2001) also considered the complications of diabetes to examine employment and incomes of diabetic and non-diabetic people of working age. They applied longitudinal Manitoba data and found that individuals with diabetes and its complications were twice as likely to be out of the labor force as people without diabetes. However, for those without complications no difference was observed. Similarly, people

with diabetes and its complications had significantly lower incomes than people without diabetes; nevertheless, those without complications had the same incomes as non-diabetics. Bastida and Pagan (2002) analyzed the impact of diabetes on the employment and earning outcomes of Mexican Americans adults 45 years and older. The results demonstrated that diabetes resulted in lower productivity and earnings for women, however no significant effect on employment probability was observed. In contrast, the results for men showed that diabetes led to lower employment probability but no effect on earnings was observed. Tunceli et al., (2005) used data from the Health and Retirement Study (1992 and 1994) and found that diabetes had significant negative effects on the employment probability of both men and women. Among individual with diabetes, the absolute probability of working was 4.4 percentage points less for women and 7.1 percentage points less for men relative to that of their counterparts without this disease.

As mentioned above, these studies assumed diabetes as exogenous variable. However, diabetes may be correlated with unobserved factors that affect employment status as well. For instance, unobserved factors such as personal motivation that affect employment positively, may influence lifestyle choices, and consequently may lead to a decrease in the probability of diabetes (Brown et. al. 2005; Latif, 2009).

Endogeneity is a common issue in many areas of applied economics including health economics and health services research (Terza et al. 2008). According to the Gauss- Markov theorem, all explanatory variables must be uncorrelated with the error term in a statistical model. Violation of this assumption leads to an endogeneity problem with biased and inconsistent estimates. Typical cases of endogeneity might arise from

measurement error, simultaneity and omitted variable or unobserved heterogeneity (Wooldridge 2015). Measurement error occurs when some variables (dependent or explanatory variables) are measured with error. Simultaneity is the case where causal relationship between the dependent and explanatory variable occurs in both directions. In other words, variables are jointly determined. Omitted variable bias is where an explanatory variable is omitted from the model and included into the error term and unobserved heterogeneity that is the source of endogeneity of diabetes in our study, occurs when there is unobserved characteristic that influence both dependent and explanatory variables and this unobserved characteristic is included into the error term. So, in order to have unbiased and consistent effect of diabetes on employment, we must take into account the endogeneity of diabetes.

In this regard, there are studies that treated diabetes as endogenous to correctly examine the extent to which diabetes may affect an individual's labor outcome. For instance, Brown et al. (2005) used data from the Border Epidemiologic Study on Aging (BESA) to analyze the impact of diabetes on employment of Mexican Americans aged 45 and older residing in the US/Mexico border area of South Texas. They first took diabetes as exogenous and found a negative and highly significant effect of diabetes on employment for both women and men. They also measured marginal effects of diabetes for women and men -0.075 and -0.074 respectively and found them qualitatively consistent with previous studies that took diabetes as exogenous (e.g. Bastida and Pagan 2002; Kahn, 1998; Ng et al. 2001; Lavigne et al. 2003). Then, they considered diabetes as endogenous and modeled that as a binary regressor in a recursive simultaneous equations probit model of the determinants of employment and used family history of diabetes

(respondents were asked whether their parents, grand-parents, brothers and sisters have diabetes) as instruments for the diabetes variable. This time they could reach a substantial negative effect on employment for men but not for women. They believed that taking diabetes as exogenous led to overestimation of the negative impact of diabetes on women. However, in larger sample sizes, ignoring endogeneity could result in underestimation of negative effects of diabetes on male employment. Moreover, the authors mentioned that their results are consistent with previous studies (e.g. Bastida and Pagan, 2002) that found more severe effects of diabetes on labor market outcomes on men than on women.

Using data from the National Population Health Survey (1998), Latif (2009) estimated the impact of diabetes on employment probabilities of Canadians women and men between the age of 15 to 64. He could find a significant and negative impacts on employment probabilities for both women and men when he took diabetes as exogenous. Then he followed the approach used by Brown et al. (2005) and used a recursive bivariate probit model to consider the endogeneity of diabetes. Results showed that diabetes had a positive and insignificant impact on men employment while women experienced a significant and negative impact on employment probability. In other words, the author concluded that taking diabetes as exogenous led to an over-estimation of its impact on male employment. Furthermore, he found that his results were in contrast with those in Brown et al. (2005) who also considered diabetes as endogenous and he brought some reasons as sources of the differences such as the different populations studied, and Canadian women are more seriously affected than men.

Harris (2009) used a multivariate probit model with a recursive simultaneous structure to estimate the effect of cardiovascular disease (coronary heart disease and stroke) and diabetes on labour force participation for individuals aged 25 or more years. He used data from the AusDiab survey, a population-based cross-sectional survey of national diabetes mellitus prevalence and associated risk factors of the Australian population. He considered the potential endogeneity of the prevalence of both chronic diseases and the correlation across the diseases from co-morbidity and unobservable personal heterogeneity. Results showed that correlations between the error terms of the labour participation equations were not strong but statistically significant which confirmed that using a system of equations resulted in a better model compared to single probit equations. Furthermore, the estimates of the average treatment effect of cardiovascular disease and diabetes prevalence on labour force participation from the multivariate model were generally lower than a single equation model. He also compared the results from his multivariate probit model with other studies that took into account endogeneity of diabetes such as Brown et al. (2005). Particularly, He showed that the marginal effects of diabetes on labour force participation for men is similar to Brown et al. (2005). However, the marginal effect of diabetes for women was negative and significant, compared to Brown et al. (2005) where no significant effect was found.

Lin (2011) used 2005 National Health Interview Survey data to estimate the impact of diabetes on employment for individuals between the age of 45 and 64 in Taiwan. By taking diabetes as exogenous, he found that men had a statistically significant negative effects of diabetes on employment while women experienced an insignificant negative effect on employment. Particularly, with a significant marginal effect of -0.09

for men and insignificant marginal effect -0.11 for women. Then, he considered the endogeneity of diabetes and took a recursive bivariate probit model for the estimation. Results demonstrated that diabetes had a negative and significant effect on employment for men. However, women had insignificant negative effects of diabetes on employment. Furthermore, marginal effects were -0.19 and -0.15 for men and women respectively. The author concluded that ignoring endogeneity of diabetes resulted in underestimation of negative effects of diabetes on employment and he found that his results are consistent with those in Brown et al. (2005) and Minor (2011).

Minor (2011) used data from the 2006 National Health Interview Survey (NHIS) to examine the impact of diabetes (type 1 and type 2) on labour force participation, days out of work, average hours worked, and earnings for women between the age of 20 and 65. He considered type 1 as exogenous variable since it is prevalent early in life due to genetic disorder and he took type 2 diabetes as endogenous since it can occur over individual's lifetime due to individual's family history and health factors. Results showed that diabetes reduced the probability of employment, average hours worked, and earnings for females. The author demonstrated that type 2 diabetes was more detrimental to most labor market outcomes compared to type 1 diabetes and that much of negative effect was due to type 2 diabetes. Moreover, he showed that type 2 diabetes was endogenous with respect to a female's probability of employment. However, the results were in contrast with those in Brown et al. (2005) where they took diabetes as endogenous among Mexican-American females. Particularly, Brown et al. (2005) found that endogeneity caused diabetes effect to be overestimated while Minor (2011) found that endogeneity caused the effect to be underestimated regarding both employment and earnings. The

author brought some reasons for the differing results such as the different samples used and different control variables that might cause potential bias to the estimates.

Using Mexican Family life Survey (MxFLS), Seuring et al. (2015) examined the impact of diabetes on employment in Mexico for individuals between the age group of 15 and 64. In contrast to other studies such as Brown et al. (2005) and Latif (2009) that found diabetes endogenous either for men or women, Seuring et al. showed that diabetes was exogenous for both groups and it significantly reduced employment probability of men and women by 10 and 4.5 percentage points respectively. They also concluded that diabetes mainly influenced the employment probability of men and women over 44. Moreover, they found stronger effects of diabetes on the poor than on the rich particularly for men and finally, they discovered that diabetes had more adverse effect on those in informal labour market than those in formal employment.

The authors then compared the results with those in Brown et al. (2005); Latif (2009); Lin (2011) and Minor (2011) where all found diabetes endogenous variable either for men or women. In compare with Brown et al. (2005) who used the sample of Mexican American 45 and over, Seuring et al. found a stronger negative effect of diabetes on men and particularly women living in Mexico. Seuring et al. also compared the results with those in Lin (2011) who used the sample of Individuals between the age of 45 and 64 in Taiwan. In compare with Lin (2011), they found a stronger effect for women while the effect for men was lower. However, when the authors compared their results with those studies in more developed countries with more advanced health systems and very different populations such as Latif (2009) and Minor (2011), they discovered that these studies did not found very strong effect for women. Seuring et al. could not mention

precisely why their results were different from the above-mentioned studies. For instance, they indicated that there are larger informal and more physically demanding labour market where rely mostly on men in Mexico compared to Canada or the US, so men might experience more reduction in their employment opportunities due to diabetes. However, the author mentioned that women and men in high income countries have more similar jobs with less physical demanding and diabetes might not limit men to a higher extent than women.

In summary, to estimate the impact of diabetes on employment we have to consider the endogeneity problem (Brown et al. 2005; Harris, 2009; Latif, 2009; Lin, 2011; Minor, 2011; Seuring et al. 2015). In other words, if we take diabetes as exogenous, it leads to overestimation or underestimation of the real effect of diabetes on employment. Furthermore, all the mentioned studies used static or cross-sectional panel data models and they did not measure diabetes effect in a dynamic model. Since employment status at previous period has an important role on employment status at current time, it is important to consider diabetes as endogenous and modeled that as a binary regressor in a dynamic model.

Employment plays a crucial role in our life. Individuals can benefit from having a job, such as income, well-being and education. Therefore, identifying the factors which affect employment is important as well. Health status is an important predictor of employment (e.g. Bound et al., 1999; Currie and Madrian, 1999; Au *et al.*, 2005; Kalwij and Vermeulen, 2008). In other words, health has a great impact on individual's ability to work, being hired or retained in the labor force. In this regard, chronic diseases such as diabetes have a high contribution to morbidity related of employment outcomes. The goal

here is to find out whether diabetic individuals are likely to have lower employment status. For this purpose, this study uses the 9 cycles of NPHS (National Population of Health Survey) to examine the impacts of diabetes on employment. Since employment status at previous period has an important role on employment status at current time, unlike the previous studies which all used static or cross-sectional panel data models (e.g. Kahn, 1998; Ng et al. 2001; Kraut et al. 2001; Bastida and Pagan, 2002; Tunceli et al. 2005; Brown et al. 2005; Latif, 2009; Harris, 2009; Lin, 2011; Minor, 2011; Seuring et al. 2015), this study uses dynamic panel probit model to identify the impacts of diabetes on employment in Canada. The contribution of this study to the literature on diabetes and labour market outcomes is the the way the dynamics are estimated and in particular the approach used to deal with initial conditions. We followed Wooldridge (2005) approach to control for unobserved effects that may be correlated with diabetes and we discuss this further in the methodology section.

The rest of the study is organized as follows: section 2.3 presents an overview of the data and offers descriptive statistics; section 2.4 explains the methodology used in the study; section 2.5 describes the empirical results; section 2.6 concludes and provides some discussion for further research.

2.3. Data and Characteristics of the Sample

This study uses the 9-cycle longitudinal data of National Population Health Survey (NPHS). The NPHS collects information related to the health of the Canadian population and socio-demographic information. This study is restricted to the individuals age 25 and 64. The sample contains 4,469 and 4,573 year-person observation for men and

women, respectively. Table 2.1 summarizes the demographic characteristics of the sample.

In the present study, the exogenous variables are defined as age, sex, marital status, immigration status, education level, dwelling owned by household, capital income (dividends and interest), household size, number of children and province of residence. The educational level has four categories: less than secondary, secondary school graduation, some post-secondary, and post-secondary graduation. The dummy variable own indicates whether dwelling owned by a household or not. The dummy variable capital income shows if the individual received any income in the past 12 months (dividends and interest) or not. Household size is categorized into five groups (hsize1-hsize5). Hsize5 refers to household size 5 and up. The dummy variable East indicates whether the individual lives in any provinces including Newfoundland (NF), Prince Edward Island (PEI), New Brunswick (NB), Nova Scotia (NS) or not. The dummy variable Prairies shows if the individual lives in any provinces including Alberta (AB), Manitoba (MB), Saskatchewan (SK) or not.

Generally, NPHS specifies individuals as diabetic if he/she has diabetes and it has not separated the types of diabetes. In fact, the following question is asked individuals whether or not they have diabetes.

"Do you have any of the following long-term conditions that have been diagnosed by a health professional? – Diabetes"

The dummy variable Employed indicates whether an individual is employed or not (not employed or not in labor force).

There would be many variables that can affect diabetes and employment status in our model. However, the study uses a restricted number of variables in the estimations. Therefore, due to the nature of our dynamic model (a recursive bivariate probit), adding more control variables to the estimation make the convergence difficult. A table including list of variables used in the analyses is defined in the Appendix 2. A.

Table 2.1: Summary Statistics[2]

Description	Mean for Men	Mean for Women
Employed	0.7373	0.5142
Diabetes	0.0785	0.0663
Has capital income	0.2724	0.2546
Lives in BC	0.1015	0.1249
Lives in Ontario	0.3359	0.3589
Lives in Quebec	0.3140	0.2774
Lives in East	0.0891	0.0874
Lives in Prairies	0.1595	0.1513
Married	0.7606	0.6799
Owns a house	0.8591	0.7773
Less than secondary school	0.2588	0.2900
Secondary school graduated	0.1450	0.1696
Has some post-secondary	0.2599	0.2689
Post-secondary graduated	0.3363	0.2715
Immigrant	0.1983	0.1802
Household size is equal to 1	0.1155	0.1487
Household size is equal to 2	0.4174	0.4783
Household size is equal to 3	0.1909	0.1858
Household size is equal to 4	0.1784	0.1168
Household size is equal or greater than 5	0.0978	0.0704
Number of Children	0.2032	0.1652
Age (Age/10)	5.3840	5.4067
Age^2 (Age^2/100)	29.5555	29.9243
Birth mother ever had diabetes	0.1930	0.2108
Birth father ever had diabetes	0.0897	0.1111
Any siblings ever had diabetes	0.1473	0.1608
Mother died of diabetes	0.2528	0.2743
Father died of diabetes	0.3025	0.3131

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

Table 2.2 provides the demographic characteristics of the sample by diabetes and sex. It is obvious that individuals with diabetes are less likely to be employed compared to the ones who have not experienced this disease. The employment rate for men without

[2] In this chapter, the sample weight provided for the NPHS is used for the summary statistics and all the estimations.

diabetes is 74 percent while 70 percent of men with diabetes are employed. Women with diabetes also have lower rate of employment compared to the ones who do not have diabetes (32 and 53 percent respectively). Furthermore, individuals with diabetes are likely to be immigrant, older and less educated. Regional differences are not obvious among diabetic and non-diabetic individuals.

Table 2.2: Demographic Characteristics by Diabetes and Sex

	Men		Women	
	Diabetic	**Not-Diabetic**	**Diabetic**	**Not-Diabetic**
Employed	0.6992	0.7406	0.3173	0.5282
Has Capital Income	0.3440	0.2663	0.2453	0.2552
Lives in BC	0.0873	0.1027	0.1190	0.1253
Lives in Ontario	0.4479	0.3264	0.3632	0.3586
Lives in Quebec	0.2552	0.3190	0.2593	0.2787
Lives in East	0.0624	0.0913	0.0896	0.0873
Lives in Prairies	0.1472	0.1605	0.1689	0.1500
Married	0.8091	0.7565	0.7197	0.6771
Owns a house	0.8101	0.8633	0.7323	0.7805
Less than secondary school	0.3049	0.2549	0.3392	0.2865
Secondary school graduated	0.0941	0.1493	0.2333	0.1650
Has some post-secondary	0.2876	0.2576	0.2312	0.2716
Post-secondary graduated	0.3134	0.3383	0.1964	0.2769
Immigrant	0.2449	0.1943	0.2269	0.1769
Household size is equal to 1	0.1585	0.1118	0.1674	0.1474
Household size is equal to 2	0.3606	0.4222	0.4406	0.4810
Household size is equal to 3	0.1648	0.1931	0.2763	0.1794
Household size is equal to 4	0.1035	0.1848	0.0395	0.1222
Household size is equal or greater than 5	0.2127	0.0881	0.0762	0.0700
Number of Children	0.1269	0.2097	0.0715	0.1719
Age (Age/10)	5.5742	5.3678	5.7327	5.3836
Age2 ($Age^2/100$)	31.4408	29.3950	33.2428	29.6888
Birth mother ever had diabetes	0.3794	0.1771	0.4083	0.1968
Birth father ever had diabetes	0.1326	0.0861	0.1073	0.1113
Any siblings ever had diabetes	0.2928	0.1349	0.3464	0.1477
Mother died of diabetes	0.3120	0.2477	0.2866	0.2735
Father died of diabetes	0.2933	0.3033	0.3293	0.3119

Note: Statistics are computed by a pooled sample of all 9 cycle of NPHS

2.4. Methodology

The present study uses a dynamic panel probit model with random effects to measure the impact of diabetes on employment. Dynamic panel data models have advantages over static ones. Dynamic panel data models capture the lags of the dependent variable as explanatory variables. The presence of these lags is essential to control for the dynamics of the process and can lead to better identification. For instance, using the Irish component of the European Community Household Panel Survey 1995-2000, Gannon (2005) analyzed the dynamic effects of disability on participation in the labour force. Using a dynamic model and controlling for unobserved heterogeneity and past participation, she demonstrated that static models overestimate the effect of disability on participation between 40-60% for men and 5-10% for women. Oguzoglu (2012) used the first 8 cycles of the NPHS and examined the dynamic impact of disability on labour force participation in Canada. He concluded that high persistence in labour force participation made disability shock influence participation for multiple periods while static models by ignoring this persistency, underestimated the overall effect of disability. Thus, we controlled for state dependence and unobserved heterogeneity by estimating a dynamic model with random effects to provide more accurate estimates of diabetes on employment.

Table 2.3 demonstrates that employment is highly persistent. 93 percent of individuals who are employed at time t-1 are still employed at time t and 76 percent of individuals who are unemployed or not in labor force at t-1 are unemployed or not in labor force at t as well.

Table 2.3: State Dependence in Employment

Variable Name	Employed at t	Unemployed or not In labor force at t
Employed at t-1	93.25	23.68
Unemployed or not in labor force at t-1	6.75	76.32

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

Table 2.4 also emphasized the state dependency of employment. Regardless of being diabetic or not people who are employed at last period are more likely to be employed at current time.

Table 2.4: Employment Rates by Diabetes and Lag of Employment

Variable Name	Employment rate for Men (if $EMP_{t-1}=1$)	Employment rate for Women (if $EMP_{t-1}=1$)
Diabetic	0.873	0.803
Not Diabetic	0.898	0.833

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

Table 2.5 presents the employment rates without considering the status of employment for the last period. Compared to the employment rates by lag of employment, the numbers are much smaller for diabetic and not diabetic individuals. This highlights the importance of state dependency of employment.

Table 2.5: Employment Rates by Diabetes

Variable Name	Employment rate for Men	Employment rate for Women
Diabetic	0.70	0.32
Not Diabetic	0.74	0.53

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

In summary, cross- sectional or static panel data models may not completely identify the dynamic relationship between diabetes and employment, and it is necessary to consider diabetes effect in a dynamic approach.

Following the previous studies; (Brown et al., 2005; Latif, 2009; Harris, 2009; Minor, 2011; Seuring et al. 2015) that considered diabetes as an endogenous variable, a recursive bivariate panel probit model can be suggested as follows:

$$EMP_{it} = \gamma_1 EMP_{i\,t-1} + \beta_1 X_{it} + \delta_1 D_{it} + \alpha_i + \varepsilon_{it} \quad (1)$$

$$D_{it} = \Upsilon_1 X_{it} + \gamma_2 Z_{it} + \delta_i + \phi_t \quad (2)$$

The parameters α_i and δ_i indicate individual specific unobserved heterogeneity and ε_{it} and ϕ_t are the random disturbance terms. In equation (2), Z_{it} refers to instrumental variables that can affect diabetes such as mother has diabetes, father has diabetes, siblings have diabetes, mother died of diabetes, and father died of diabetes (Latif, 2009). In other words, when diabetes is endogenous, unobserved factors related to health conditions may be correlated with unobserved factors for the employment equation. So a univariate panel probit model fails to show the consistent estimates of the effects of diabetes on employment. In addition, employment coefficient for males and females can be overestimated or underestimated.

The study uses random effects for the panel data model. In fact, due to the non-linear nature of the model, α_i cannot be eliminated through differencing or within transformation. Furthermore, adding N-1 individual dummy variables to the probit model to estimate α_i , will cause severely biased estimates of β unless T is large. With T small estimates of α_i and β are inconsistent. This is known as incidental parameters problem. The main issue to estimate equation (1) is that the unobserved effect α_i is correlated with $EMP_{i\,t-1}$ by definition. This is usually called the initial condition problem that it will become less serious when T gets large. However, with T small the problem can result in severe bias. This study tries to use the approach suggested by Wooldridge (2005) to solve the initial conditions problem. In this approach, α_i can be written as a function of EMP_{i0}

(the employment status from the first cycle) and X_i. This approach takes into account the employment status from the first cycle as well as time averages of X_{it} as additional repressors.

In other words, Wooldridge suggests to model the distribution of $\{EMP_{i1}, \ldots, EMP_{iT}\}$ given EMP_{i0} by using conditional Maximum likelihood approach. For this purpose, the distribution for α_i, given EMP_{i0} and other exogenous variables can be specified as follows:

$$f(EMP_{i1}, \ldots, EMP_{iT} \mid EMP_{i0}) = \int f(EMP_{i1}, \ldots, EMP_{iT} \mid EMP_{i0}, X_{it}, \alpha_i)\, g(\alpha_i \mid EMP_{i0}, X_{it})\, d\alpha_i$$

$$g(\alpha_i \mid EMP_{i0}, Xit) \sim N(\alpha_i = \psi + \xi_0 EMP_{i0} + \xi \bar{X}_i, \sigma^2_a)$$

As in the panel probit model the assumption is that there is no correlation between unobserved heterogeneity effect (α_i) and the explanatory variables, the variable $\bar{X}_i$ indicates time averages of all time- varying variables in X_{it} to relax this limitation. Particularly, we can write:

$$\alpha_i = \psi + \xi_0 EMP_{i0} + \xi \bar{X}_i + a_i \quad (3)$$

$a_i \sim$ Normal $(0, \sigma^2_a)$ and independent of EMP_{i0} and X_i.

Substituting (3) in (1) will result in the following equation:

$$EMP^*_{it} = \beta_1 X_{it} + \delta_1 D_{it} + \gamma_1 EMP_{i\,t-1} + \xi_0 EMP_{i0} + \xi \bar{X}_i + a_i + \varepsilon it \quad (4)$$

Now we can estimate equations (4) and (2) as a recursive dynamic panel probit by random effects:

$$EMP^*_{it} = \beta_1 X_{it} + \delta_1 D_{it} + \gamma_1 EMP_{i\,t-1} + \xi_0 EMP_{i0} + \xi \bar{X}_i + a_i + \varepsilon it \quad (4)$$

$$D^*_{it} = \beta_2 X_{it} + \Theta_2 Z_{it} + \delta_i + \phi_t \quad (2)$$

Since employment status at previous period has an important role on employment status at current time, unlike the previous studies which all used static or cross-sectional models

(e.g. Kahn, 1998; Ng et al. 2001; Kraut et al. 2001; Bastida and Pagan, 2002; Tunceli et al. 2005; Brown et al. 2005; Latif, 2009; Harris, 2009; Lin, 2011; Minor, 2011; Seuring et al. 2015), this study uses the above-mentioned Wooldridge approach (2005) to estimate a dynamic panel probit model to provide more accurate estimates of the effect of diabetes on employment.

2.5. Empirical Findings

As an initial estimation, a static panel probit model can be estimated. Table 2.6 shows the results of a static panel probit model in which endogeneity of diabetes and the dynamic nature of the individual's employment status has not been considered. Results shows that diabetes has significant negative effect for both men and women. Other control variables have the expected signs. The study also estimated the marginal effect of diabetes to show the effects of diabetes variable on employment probability status. The marginal effects are -0.003 and -0.111 for men and women respectively; that is, the employment propensity of diabetic men are 0.3 percentage points lower than that of non-diabetic men while, the employment propensity of diabetic women are 11.1 percentage points lower than that of non- diabetic counterparts. These results are consistent with those studies that considered diabetes as exogenous and found a significant negative diabetes effect on employment status of both men and women (Latif, 2009; Brown et al , 2005; Tunceli et al. 2005; Lavigne et al, 2003; Bastida and Pagan, 2002; Ng et al. 2001; Kahn, 1998).However, since there would be unobserved factors related to employment status that might be correlated with diabetes, the estimated results are inconsistent and biased.

Table 2.6: Static Panel Probit Model

	Coefficients for Men	Coefficients for Women
Diabetes	-0.0193***	-0.4573***
	(0.00497)	(0.00479)
Has Capital Income	- 0.3898***	-0.0897***
	(0.00233)	(0.00220)
Lives in BC	0.0684***	0.4613***
	(0.00676)	(0.00635)
Lives in Ontario	0.5430***	0.4772***
	(0.00495)	(0.00489)
Lives in East	-0.2642***	0.1744***
	(0.00673)	(0.00726)
Lives in Prairies	0.6498***	0.6207***
	(0.00592)	(0.00595)
Married	0.6173***	-0.7153***
	(0.00680)	(0.00455)
Owns a house	-0.1037***	0.2177***
	(0.00457)	(0.00385)
Secondary school graduated	0.1493***	0.3636***
	(0.00603)	(0.00514)
Has some post-secondary	-0.3991***	0.8135***
	(0.00463)	(.00446)
Post-secondary graduated	0.3333***	0.7474***
	(0.00478)	(0.00445)
Immigrant	0.7916***	0.2389***
	(0.00513)	(0.00488)
Household size is equal to 2	-0.6761***	0.0290***
	(0.00480)	(0.00373)
Household size is equal to 3	-0.6866***	0.3010***
	(0.00542)	(0.00441)
Household size is equal to 4	-0.5444***	-0.1093***
	(0.00623)	(0.00508)
Household size is equal or greater than 5	-1.1070***	0.6431***
	(0.00747)	(0.00668)
Number of Children	-0.1877***	-0.3621***
	(0.00256)	(0.00247)
Age	5.9161***	3.6680***
	(0.01763)	(0.01439)
Age2	-0.6723***	-0.4466***
	(0.00168)	(0.00138)
Constants	-20.1487***	-16.9794***
	(0.11119)	(0.07762)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Model include time average of all time- varying control variables. Reference groups are not diabetic, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

By controlling the lag of employment status, we can reach to a dynamic panel probit model. Table 2.7 shows the results of a dynamic panel probit model in which the diabetes binary variable is still considered as an exogenous variable. Results show a high state dependency in employment variable and the coefficient for Employed t-1 is positive and significant. Thus, individuals who were employed in the previous period are more likely to be currently employed. Specifically, even by controlling the lag of employment, diabetes has still significant negative effect on employment's status of both groups. This time, the estimated marginal effect of diabetes are -0.035 and -0.017 for men and women respectively. Compared to the cross-sectional and static previous studies (Latif, 2009; Brown 2005), the diabetes effects are much smaller which indicate that those studies overestimate the diabetes effect for both groups. However, in this study, the estimated dynamic model has controlled for dynamic nature of individuals' employment status, it is still necessary to correct for endogeneity of diabetes in order to have an unbiased and consistent diabetes effect.

Table 2.7: Dynamic Panel Probit Regression for Employment Status

	Coefficients for Men	Coefficients for Women
Employed t-1	1.7145***	1.2876***
	(0.00289)	(0.00257)
Initial Employed	0.7378***	1.0992***
	(0.00411)	(0.00403)
Diabetes	-0.2219***	-0.0899***
	(0.00332)	(0.00426)
Has Capital Income	-0.3907***	-0.1376***
	(0.00239)	(0.00252)
Lives in BC	-0.1385***	0.1534***
	(0.00349)	(0.00405)
Lives in Ontario	0.1060***	0.1354***
	(0.00263)	(0.00316)
Lives in East	-0.0895***	0.0413***
	(0.00339)	(0.00440)
Lives in Prairies	0.1630***	0.1904***
	(0.00305)	(0.00383)
Married	0.1337***	-0.6034***
	(0.00671)	(0.00538)
Owns a house	-0.0029***	0.0814***
	(0.00508)	(0.00498)
Secondary school graduated	0.0189***	0.1541***
	(0.00320)	(0.00354)
Has some post-secondary	-0.0981***	0.4704***
	(0.00266)	(0.00319)
Post-secondary graduated	0.0742***	0.3809***
	(0.00266)	(0.00319)
Immigrant	0.3885***	0.1680***
	(0.00293)	(0.00323)
Household size is equal to 2	-0.3630***	0.0841***
	(0.00387)	(0.00356)
Household size is equal to 3	-0.3223***	0.3890***
	(0.00428)	(0.00426)
Household size is equal to 4	-0.0686***	0.3805***
	(0.00512)	(0.00503)
Household size is equal or greater than 5	-0.5398***	0.7704***
	(0.00622)	(0.00670)
Number of Children	-0.2132***	-0.3696***
	(0.00271)	(0.00293)
Age	2.5348***	1.5122***
	(0.02165)	(0.01976)
Age2	-0.3130***	-0.2174***
	(0.00206)	(0.00190)

Constants	-6.570***	-9.6295***
	(0.06147)	(0.05145)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. EMPt-1 refers to the employment status from the previous period. Initial EMP is the employment status at first cycle. Model includes time average of all time -varying control variables.

In order to estimate the dynamic probit model by controlling for the endogeneity of diabetes, a Two-Stage Least Squares (2SLS) regression analysis can be suggested. Table 2.8 presents the results of a probit panel data regression for diabetes as an endogenous variable for the first stage. This regression includes exogenous variables and instrumental variables as well. Following Brown et al. (2005) and Latif (2009), this study used family history of diabetes as instruments for diabetes. According to the figures, all variables are significant at 99 confidence intervals.

Table 2.8: Probit Regression for Diabetes

	Coefficients for Men	Coefficients for Women
Has Capital Income	0.1435***	0.2248***
	(0.00559)	(.00531)
Lives in BC	-0.1526***	0.1431***
	(0.01872)	(.01397)
Lives in Ontario	0.5579***	-0.1842***
	(0.01171)	(.01093)
Lives in East	-0.1499***	0.2178***
	(0.01845)	(.01511)
Lives in Prairies	0.0322**	0.4036***
	(0.01491)	(.01231)
Married	0.9460***	0.5174***
	(0.01573)	(.01153)
Owns a house	-0.3414***	-0.1087***
	(0.01137)	(.00906)
Secondary school graduated	-0.5249***	0.9421***
	(0.01560)	(.01102)
Has some post-secondary	0.1575***	-0.5995***
	(0.01180)	(.01143)
Post-secondary graduated	-0.5430***	-0.1463***
	(0.01250)	(.01016)
Immigrant	0.3820***	0.3467***
	(0.01108)	(.01088)
Household size is equal to 2	-1.9301***	-0.7062***
	(0.01598)	(.01051)
Household size is equal to 3	-1.4612***	-0.3782***
	(0.01630)	(.01188)
Household size is equal to 4	-1.6138***	-1.0131***
	(0.01820)	(.01683)
Household size is equal or greater than 5	-0.3307***	-0.4550***
	(0.02158)	(.01939)
Number of Children	-0.8174***	-0.5326***
	(0.01061)	(.01017)
Age	4.7442***	1.3640***
	(0.07403)	(.06314)
Age2	-0.1471***	0.1032***
	(0.00681)	(.00581)
Birth mother ever had diabetes	1.2667***	0.9398***
	(0.01083)	(.00890)
Birth father ever had diabetes	1.3051***	-0.1350***
	(0.01568)	(.01363)
Any siblings ever had diabetes	1.3078***	1.2039***
	(0.01237)	(.00945)
Mother died of diabetes	0.4979***	-0.1535***

	(0.01008)	(.00907)
Father died of diabetes	-0.1055***	0.05894***
	(0.01011)	(.00849)
Constants	-30.6301***	-18.0247***
	(0.20103)	(.17186)

Note: Significance level: *10%, **5%, ***1%. Standard Errors are in parenthesis.

In 2SLS analysis, when there is endogenous dummy (binary) variable, one common mistake is to use the fitted value of the endogenous variable from the first stage into the second stage which is called forbidden regression (Angrist and Pischke, 2009). To avoid this, one alternative to the forbidden second step is to use nonlinear fitted values as instruments (Angrist and Pischke, 2009). For this study, I used the fitted value of the endogenous diabetes dummy variable as well as family history of diabetes as instruments to use in the second step. [3]Table 2.9 shows the results of a dynamic panel data model. The figures indicate that diabetes have insignificant positive effects on men's employment. However, the effects of diabetes on women's employment is negative and significant. In fact, by controlling the endogeneity of diabetes in our dynamic model, the results are consistent with the cross-sectional and static study by Latif (2009). However, the numbers are much smaller.

[3] . The same approach has been taken by Windmeijer and Santos Silva, 1997.

Table 2.9: Dynamic Panel Probit Regression for Employment Status

	Coefficients for Men	Coefficients for Women
Employed t-1	0.5811***	0.5541***
	(.01700)	(.01545)
Initial Employed	0.1201***	0.0866***
	(.01823)	(.01667)
Diabetes	0.1327	-0.2103*
	(.09887)	(.12379)
Has Capital Income	-0.0387**	-0.0397**
	(.01764)	(.01828)
Lives in BC	-0.0238	0.0229
	(.02429)	(.02289)
Lives in Ontario	0.0101	0.0189
	(.01863)	(.01852)
Lives in East	-0.0005	-0.0090
	(.01656)	(.01752)
Lives in Prairies	0.0352*	0.0394**
	(.01796)	(.01862)
Married	-0.0243	-0.0077
	(.04316)	(.03772)
Owns a house	-0.0272	-0.0620*
	(.03708)	(.03696)
Secondary school graduated	0.0003	0.0223
	(.02030)	(.01850)
Has some post-secondary	-0.0115	0.0578***
	(.01627)	(.01612)
Post-secondary graduated	0.0235	0.0486***
	(.01715)	(.01638)
Immigrant	0.0427**	0.0293
	(.01801)	(.01897)
Household size is equal to 2	-0.0034	0.0001
	(.02149)	(.01970)
Household size is equal to 3	0.0091	0.0406*
	(.02500)	(.02465)
Household size is equal to 4	0.0285	0.0575*
	(.02899)	(.03064)
Household size is equal or greater than 5	-0.0134	0.0948**
	(.03784)	(.04160)
Number of Children	-0.0142	-0.0331*
	(.01802)	(.02000)
Age	0.5172***	0.2187
	(.14065)	(.13939)
Age2	-0.0600***	-0.0290**
	(.01335)	(.01332)

Constants	-0.4666	-0.5776**
	(.32159)	(.26971)

Note: Significance level: *10%, **5%, ***1%. Standard Errors are in parenthesis. EMPt-1 refers to the employment status from the previous period. Initial EMP is the employment status at first cycle. Model includes time average of all time- varying control variables.

In order to check the robustness of the results, the study uses a bivariate probit model using cluster id. Specifically, a bivariate probit model is more efficient than a two- step analysis since a 2sls estimation analysis does not take into account the correlation of error terms of employment and diabetes equations (Green, 1998). Table 2.10 and Table 2.11 show the results of the dynamic bivariate probit model for men and women respectively. According to the results, men experience a significant positive effect of diabetes on their employment status; however, the effect of diabetes on women's employment is negative and significant. The marginal effects are 0.16 and -0.22 for men and women respectively; that is, the employment propensity of diabetic men are 16 percentage points bigger than that of non-diabetic men while, the employment propensity of diabetic women are 22 percentage points lower than that of non- diabetic counterparts. Furthermore, by controlling the lag of employment in our dynamic bivariate probit model, the sign of diabetes effect for both men and women is still the same as our first suggested estimation method (2SLS analysis estimation). Contrary to the first estimation approach, this time the diabetes effect on men's employment status is significant. However, the lag of employment has still positive and significant effect on employment status of both groups.

Table 2.10: A Bivariate Dynamic Panel Probit Model for Men

	Employment Status Coefficients	Diabetes Coefficients
Employed t-1	1.7872***	
	(0.10453)	
Initial Employed	0.4569***	
	(0.10315)	
Diabetes	0.8955***	
	(0.27121)	
Has Capital Income	-0.3499***	-0.0590
	(0.12303)	(0.09688)
Lives in BC	-0.1240	0.0816
	(0.15151)	(0.30392)
Lives in Ontario	0.0179	0.3736
	(0.12184)	(0.22987)
Lives in East	0.0931	-0.0215
	(0.11236)	(0.23311)
Lives in Prairies	0.1069	0.2392
	(0.12269)	(0.23082)
Married	-0.0093	0.5776**
	(0.27821)	(0.28641)
Owns a house	0.0007	0.0746
	(0.19522)	(0.21683)
Secondary school graduated	0.1084	-0.4371
	(0.12486)	(0.28460)
Has some post-secondary	-0.0257	-0.1029
	(0.10866)	(0.20001)
Post-secondary graduated	0.1349	-0.3863*
	(0.10874)	(0.21292)
Immigrant	0.2776**	0.1202
	(0.11687)	(0.25475)
Household size is equal to 2	-0.1436	-0.6149**
	(0.19899)	(0.24895)
Household size is equal to 3	-0.1280	-0.5040*
	(0.21322)	(0.29625)
Household size is equal to 4	0.0834	-0.4580
	(0.21648)	(0.33428)
Household size is equal or greater than 5	-0.5137*	0.4425
	(0.29056)	(0.43702)
Number of Children	-0.1016	-0.4591***
	(0.08991)	(0.15489)
Age	1.7875**	-0.7660
	(0.78412)	(0.9693)
Age2	-0.2388***	0.1289
	(0.07274)	(0.89771)

Birth mother ever had diabetes		0.5398***
		(0.19790)
Birth father ever had diabetes		0.1604
		(0.27657)
Any siblings ever had diabetes		0.4298**
		(0.20051)
Mother died of diabetes		-0.0332
		(0.18481)
Father died of diabetes		0.0068
		(0.17127)
Constants	-4.3936**	-16.6085***
	(2.08393)	(6.22635)

Note: Significance level: *10%, **5%, ***1%. Standard Errors are in parenthesis. EMPt-1 refers to the employment status from the previous period. Initial EMP is the employment status at first cycle. Model includes time average of all time -varying control variables.

Table 2.11: A Bivariate Dynamic Panel Probit Model for Women

	Employment Status Coefficients	Diabetes Coefficients
Employed t-1	1.6797***	
	(0.09715)	
Initial Employed	0.4199***	
	(0.09976)	
Diabetes	-1.0571**	
	(0.47581)	
Has Capital Income	-0.1017	0.1435
	(0.09643)	(0.10315)
Lives in BC	0.1014	0.0290
	(0.11752)	(0.24302)
Lives in Ontario	0.0747	0.0151
	(0.11176)	(0.18679)
Lives in East	-0.0139	-0.0458
	(0.11260)	(0.24300)
Lives in Prairies	0.1191	0.1742
	(0.11688)	(0.19908)
Married	-0.3993**	0.7683***
	(0.20012)	(0.22845)
Owns a house	-0.0044	-0.2101
	(0.20543)	(0.17549)
Secondary school graduated	0.1699	0.1969
	(0.11169)	(0.19530)
Has some post-secondary	0.2792**	0.0011
	(0.10821)	(0.19073)
Post-secondary graduated	0.2446**	-0.2358
	(0.11301)	(0.20035)
Immigrant	0.1083	0.2187
	(0.10737)	(0.19922)
Household size is equal to 2	0.0820	-0.4501**
	(0.11625)	(0.22643)
Household size is equal to 3	0.3266**	-0.1235
	(0.14745)	(0.28842)
Household size is equal to 4	0.3754**	-0.5962*
	(0.17551)	(0.33057)
Household size is equal or greater than 5	0.5194**	-0.0678
	(0.23838)	(0.39509)
Number of Children	-0.2916**	-0.5583***
	(0.11553)	(0.16496)
Age	0.0454	-0.8407
	(0.68177)	(0.72711)
Age2	-0.0450	0.1450**
	(0.06601)	(0.06862)

Birth mother ever had diabetes		0.4952***
		(0.16007)
Birth father ever had diabetes		0.0196
		(0.21485)
Any Sibling ever had diabetes		0.4501**
		(0.19502)
Mother died of diabetes		0.1258
		(0.14908)
Father died of diabetes		0.0640
		(0.14866)
Constant	-7.1415***	-6.2806
	(1.66528)	(3.87011)

Note: Significance level: *10%, **5%, ***1%. Standard Errors are in parenthesis. EMPt-1 refers to the employment status from the previous period. Initial EMP is the employment status at first cycle. Model includes time average of all time -varying control variables.

2.6. Discussion and Conclusion

This study uses longitudinal data of National Population Health Survey (NPHS) to examine the dynamic effect of diabetes on employment in Canada. The study controls for state dependence and unobserved heterogeneity by estimating a dynamic panel model and takes into account men and women separately for individuals age 25-64.

Results indicate that diabetes have insignificant positive effects on men' employment. However, the effects of diabetes on women's employment is negative and significant. In fact, by controlling the lag of employment in our dynamic model, the results confirm the signs and significance of diabetes coefficients estimated by Latif (2009). However, the numbers are much smaller (0.13 vs. 0.96 for men and -0.21 vs. 0-.44 for women). That is, static and cross-sectional models overestimate the diabetes effect on individual's employment status. The negative and significant diabetes effect for women may indicate that women are more affected than men by this disease that make them unemployed or out of labour force market. However, the source of this difference is not clear, and it needed further investigation (Latif, 2009). Since the prevalence and incidence of diabetes is increasing, it is important to have accurate estimates of the labour market costs of diabetes.

From a health policy perspective, the results derived from this study suggest that in estimating labour market related costs associated with diabetes, we should also control for state dependency and unobserved heterogeneity for more accurate estimates of the effect of diabetes on employment. Particularly, a precise estimates of diabetes effect on employment would be helpful for policy makers, government or employers to know the economic burden and design the appropriate policies to reduce related costs.

For instance, employers should accommodate employees with diabetes to be more effective at work. In this regard, employers should provide updated guidelines and protocols about diabetes and its management at work (American Diabetes Association, 2011). They also can specify ways to reduce the risk of severe hypoglycemia such as additional breaks to check blood glucose levels or permission to access food or beverages at their desk (American Diabetes Association, 2011). Furthermore, diabetic employees may need leave or a flexible work schedule for medical appointments or other diabetes care needs (American Diabetes Association, 2011). Providing all accommodations for diabetic individuals at work is not possible; however, employers can ensure whether the provided accommodations help diabetes to perform their jobs more effectively and that may reduce the negative effects of diabetes on employment.

In this study, our goal was to provide more accurate estimates of the effect of diabetes on employment. However, we acknowledge some limitations.

Our study has not differentiated between the main types of diabetes (type 1 and type 2), since NPHS has taken diabetes as a whole and it has not specified the types of diabetes. In fact, this issue is so important because the impact of diabetes on individual's employment may vary depending on which types of diabetes are considered. As type 2 diabetes accounts for about 90% of all individuals with diagnosed diabetes, this study restricted the age group to 25-64 to moderate this problem.

The static and cross-sectional studies such as Brown et al. (2005) and Latif (2009) also estimated probit regressions between instrumental variables and conditions related to diabetes (such as heart disease, high blood pressure, cataracts, glaucoma , stroke and visual impairment) and to conditions unrelated to diabetes(such as cancer, arthritis,

bon/joint problems and hearing impairment) and they did not find any strong statistical evidence that these health conditions are correlated with instruments. As a robustness check, Latif (2009) also included health conditions related to diabetes in his bivariate probit model, however he mentioned that this inclusion changed the magnitude of the diabetes coefficients, qualitative interpretation remained unchanged. Our study has not taken into account health conditions related to diabetes and health conditions unrelated to diabetes. As our data has a panel nature and complication of diabetes can influence diabetic individuals gradually over their life, there is a high possibility that diabetic individuals with complications might increase the effect on individuals' employment status. Inclusion of diabetes complications may be applied as an interaction variable with diabetes or separate control variables in the regression or diabetes may restricted to diabetes with complications as a control variable in employment equation for future research.

Another limitation of this study is that the measure for diabetes is self-reported which can bring bias to the model/inference.

In order to check the robustness of our main results (obtained from the main model-2SLS Analysis), the study used a bivariate probit model. However, results confirmed the sign of diabetes effect on both groups' employment status from the 2SLS analysis, this time the positive diabetes effect on men's employment is significant. One possibility could result from the model specification. In other words, the main estimation (2SLS Analysis estimation) was measured by using panel data with random effects while the alternative bivariate probit model is a pooled probit model. In fact, the syntax that is used for estimating a bivariate probit model in Stata is designed for cross- sectional

models not panel data. Thus, in order to make a good alternative model, the study used the cluster option to show that the observations are clustered into individuals and that the observations may be correlated within individuals but would be independent between individuals. However, this is not equivalent to random effects model but it non-parametrically considers the correlation of residuals[4].

[4] Another possibility is to follow conditional mixed-process (CMP) framework suggested by Roodman's cmp command (2011). Specifically, this user-written command can be used to jointly estimate two or more equations and allow for the correlation of their error terms. However, this technique cannot be implemented with standard software such as STATA.

Appendix 2.A

Table 2.A: Variables Definitions

Variable Name	Definition of variables
Employed	= 1 if employed
Diabetes	=1 if has diabetes
Has capital income	=1 has capital income
Lives in BC	= if lives in BC
Lives in Ontario	=1 if lives in Ontario
Lives in Quebec	=1 if lives in Quebec
Lives in East	=1 if lives in NF, PEI, NB, NS
Lives in Prairies	=1 if lives in ALB, MB or SASK
Married	= 1 if married
Owns a house	= 1 if owns a house
Less than secondary school	=1 has less than secondary education
Secondary school graduated	= 1 if is secondary school graduated
Has some post-secondary	=1 if has some post-secondary education
Post-secondary graduated	= if is post-secondary school graduated
Immigrant	=1if immigrant
Household size is equal to 1	=1if household size =1
Household size is equal to 2	=1if household size =2
Household size is equal to 3	=1if household size =3
Household size is equal to 4	=1if household size =4
Household size is equal or greater than 5	=1if household size =>5
Number of Children	Number of children
Age	Age/10
Age^2	$Age^2/100$
Birth mother ever had diabetes	=1 if birth mother ever had diabetes
Birth father ever had diabetes	=1 if birth father ever had diabetes
Any siblings ever had diabetes	=1 if any siblings ever had diabetes
Mother died of diabetes	=1 if mother died of diabetes
Father died of diabetes	=1 if father died of diabetes

References:

American Diabetes Association. Diabetes and employment.*Diabetes care* 34, no. Supplement 1, S82-S86.

Au, D. W. H., Crossley, T. F., & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health economics*, *14*(10), 999-1018.

Bastida, E., & Pagán, J. A. (2002). The impact of diabetes on adult employment and earnings of Mexican Americans: findings from a community based study. *Health economics*, *11*(5), 403-413.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics*, *6*(2), 179-202.

Canadian Diabetes Association. (2016). Living with Gestational Diabetes. Retrieved from http://www.diabetes.ca/diabetes-and-you/living-with-gestational-diabetes.

Canadian Diabetes Association. (2016). Types of Diabetes. Retrieved from http://www.diabetes.ca/about-diabetes/types-of-diabetes .

Cleal, B., & Poulsen, K. (2015). The Relationship between Diabetes and Retirement from Work: A Ten-Year Follow-Up in National Registers. *Int J Diabetes Clin Res*, *2*, 035.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, *3*, 3309-3416.

daCosta DiBonaventura, M., Cappelleri, J. C., & Joshi, A. V. (2011). A longitudinal assessment of painful diabetic peripheral neuropathy on health status, productivity, and health care utilization and cost. *Pain Medicine*, *12*(1), 118-126

Diabetes Canada. (2018). Types of Diabetes. Retrieved from https://www.diabetes.ca/about-diabetes/types-of-diabetes

Gannon, B. (2005). A dynamic analysis of disability and labour force participation in Ireland 1995–2000. *Health economics*, *14*(9), 925-938.

World Health Organization. (2014). *Global status report on noncommunicable diseases 2014* (No. WHO/NMH/NVI/15.1). World Health Organization.

Greene, W. H. (1998). Gender economics courses in liberal arts colleges: Further results. *The Journal of Economic Education*, *29*(4), 291-300.

Harris, A. (2009). Diabetes, cardiovascular disease and labour force participation in Australia: An endogenous multivariate probit analysis of clinical prevalence data. *Economic Record*, *85*(271), 472-484.

Herquelot, E., Guéguen, A., Bonenfant, S., & Dray-Spira, R. (2011). Impact of diabetes on work cessation: data from the GAZEL cohort study. *Diabetes care*, *34*(6), 1344-1349.

International Diabetes Federation. (2016). About Diabetes. Retrieved from http://www.idf.org/about-diabetes.

International Diabetes Federation. (2015) .Facts & Figures. Retrieved from http://www.idf.org/WDD15-guide/facts-and-figures.html.

Kahn, M. E. (1998). Health and labor market performance: the case of diabetes. *Journal of Labor Economics*, *16*(4), 878-899.

Kalwij, A., & Vermeulen, F. (2008). Health and labour force participation of older people in Europe: what do objective health indicators add to the analysis?. *Health economics*, *17*(5), 619-638.

Kraut, A., Walld, R., Tate, R., & Mustard, C. (2001). Impact of diabetes on employment and income in Manitoba, Canada. *Diabetes Care*, *24*(1), 64-68.

Latif, E. (2009). The impact of diabetes on employment in Canada. *Health economics*, *18*(5), 577-589.

Lavigne, J. E., Phelps, C. E., Mushlin, A., & Lednar, W. M. (2003). Reductions in individual work productivity associated with type 2 diabetes mellitus. *Pharmacoeconomics*, *21*(15), 1123-1134.

Lin, S. J. (2011). Estimating the impact of diabetes on employment in Taiwan. *Econ Bull*, *31*(4), 3089-102.

Logminiene, Z., Norkus, A., & Valius, L. (2004). Direct and indirect diabetes costs in the world. *Medicina*, *40*(1), 16-26.

Matsushima, M., Tajima, N., Agata, T., Yokoyama, J., Ikeda, Y., & Isogai, Y. (1993). Social and economic impact on youth-onset diabetes in Japan. *Diabetes Care*, *16*(5), 824-827.

Minor, T. (2011). The effect of diabetes on female labor force decisions: new evidence from the National Health Interview Survey. *Health economics*, *20*(12), 1468-1486.

Ng, Y. C., Jacobs, P., & Johnson, J. A. (2001). Productivity losses associated with diabetes in the US. *Diabetes care*, *24*(2), 257-261.

Oguzoglu, U. (2012). Dynamics of disability and work in Canada.

Public Health Agency of Canada. (2016). Facts and figures from a public health perspective. Retrieved from http://www.phac-aspc.gc.ca/cd-mc/publications/diabetes-diabete/facts-figures-faits-chiffres-2011/chap1-eng.php#endnote7.

Robinson, N., Yateman, N. A., Protopapa, L. E., & Bush, L. (1990). Employment problems and diabetes. *Diabetic medicine*, *7*(1), 16-22.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal*, *11*(2), 159-206.

Seuring, T., Goryakin, Y., & Suhrcke, M. (2015). The impact of diabetes on employment in Mexico. *Economics & Human Biology*, *18*, 85-100.

Shelton Brown III, H., Pagán, J. A., & Bastida, E. (2005). The impact of diabetes on employment: genetic IVs in a bivariate probit. *Health economics*, *14*(5), 537-544.

Somerville, R. (2009). An economic tsunami: the cost of diabetes in Canada. *Toronto, Canada: Canadian Diabetes Association*.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics*, *27*(3), 531-543.

Tunceli, K., Bradley, C. J., Nerenz, D., Williams, L. K., Pladevall, M., & Lafata, J. E. (2005). The impact of diabetes on employment and work productivity. *Diabetes care*, *28*(11), 2662-2667.

Vijan, S., Hayward, R. A., & Langa, K. M. (2004). The impact of diabetes on workforce participation: results from a national household sample. *Health services research*, *39*(6p1), 1653-1670.

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes care*, *27*(5), 1047-1053.

Windmeijer, F. A., & Santos Silva, J. M. (1997). Endogeneity in count data models: an application to demand for health care. *Journal of applied econometrics*, *12*(3), 281-294.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach*. Nelson Education.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, *20*(1), 39-54.

World Health Organization. (2016). Diabetes. Retrieved from http://www.who.int/mediacentre/factsheets/fs312/en/.

World Health Organization. (2016). 10 Facts about diabetes. Retrieved from http://www.who.int/features/factfiles/diabetes/facts/en/index9.html.

World Health Organization. (2014). Global Health Estimates: Deaths by Cause, Age, Sex and Country, 2000-2012. Geneva.

Chapter 3. Analyzing the Relationship Between Health and Employment in Canada

Abstract

This study examines the interaction between health and employment through a simultaneous equation model to take into account the endogeneity between the two variables. For this purpose, the study uses a sample on Canadian men and women between the age of 25 and 64 and analyzes them separately. Results show that health has a positive and significant effect on employment status for both men and women. For the reverse effect, employment also has a positive and significant effect on health status of both groups. The results also show a high state dependency of health and employment variables for both men and women which confirm the dynamic of the two key variables.

Keywords: health, employment, dynamic panel data model

3.1. Introduction

Health is a state of complete physical, mental and social well-being and not only the absence of disease or infirmity (World Health Organization, 1958). Health is fundamental to well being and can play a substantial role in economic progress. For instance, good health can have a positive, sizable and statistically significant effect on aggregate output (Bloom, Canning and Sevilla, 2004). Health status can contribute as an important determinant of an individual's happiness as well. In fact, health can be a more powerful determinants of an individual's happiness than other factors such as income, marital status or age (Guo and Hu, 2011). On the other hands, health conditions such as extreme pain, extreme anxiety, and the difficulties with usual activities (it refers to individuals' typical daily routine such as going to work and school, leisure time and etc) have negative effects on health satisfaction as well (Graham, Higuera and Lora, 2011). Health status can influence individuals' income. Particularly, poor health is one of the major problems associated with low income. Individuals can not apply for paid employment or lose their job once they become ill or retire early (Martikainen et al., 2003). A poor health status can weaken the incentive of rural laborers to participate in the migrant labor force and significantly reduces their earnings (Qin et al., 2015). On the other hand, income can affect health and longevity through different mechanism such as clinical, behavioral, social, and environmental factors (Khullar and Chokshi, 2018). As noted by Preston (1975), based on the Absolute Income hypothesis, the relationship between health and income is positive and concave which means that individuals with higher income have better health outcomes.

Health may affect individual's educational attainments or outcomes. For instance, individual's health status measured by the probability of sickness, can significantly affect

academic success (Gan and Gong, 2007). In addition, poor health in childhood may prevent children from attending school (Case et al., 2009). On the other hand, education may determine health as well. Particularly, individuals with more years of schooling are more likely to have better health and well being and healthier behaviours (Feinstein et al. 2006). Moreover, higher educational levels positively affect self-reported health and reduce the number of chronic conditions (Lundborg, 2008).

Although improvement in well-being can result in improved workplace performance such as profitability, labor productivity and the quality of outputs or services (Bryson, Forth and Stokes, 2014), poor health has its own direct and indirect costs as well. According to the Economic Burden of Illness in Canada (EBIC) that gives comparable costing information for all major illnesses, the estimated total EBIC in 2008, including direct costs (hospital care, physician care and drug expenditures) and indirect costs (the value of lost production due to premature mortality and morbidity) was $192.8 billion and it increased by 13.8% from 2005 to 2008 (Public Health Agency of Canada, 2005-2008). Morbidity and mortality costs accounted for 8.7% ($16.4 billion) and 0.2% ($0.5 billion) of total costs, in 2008, respectively. Moreover, morbidity costs made a contribution of over 97% of indirect costs for all five categories (Injuries, respiratory infections, musculoskeletal diseases, neuropsychiatric conditions and certain infectious and parasitic diseases) (Public Health Agency of Canada, 2005-2008).

Poor health conditions also can result in lost productivity as major source of indirect costs. Particularly, there are two measures of lost productivity: absenteeism (time lost at work due to illness) and presenteeism (time impaired at work due to illness). In 2011, the average absenteeism rate was 9.3 days per full time Canadian employee

(Dabboussy and Uppal, 2012). The direct cost of absenteeism including time off for minor illness to longer-term leaves of absence was estimated $16.6 billion for the Canadian economy in 2012 (Stewart, 2013).

A report from Morneau Shepell in 2015, a survey of employees, employers and physicians across Canada, released that eight out of ten employee respondents self-reported experience with presenteeism. According to the report, 53 percent of employees considered presenteeism as a serious issue in their workplace while 43 percent of employees indicated absenteeism as a serious issue.

Although health is fundamental to our well-being, work acts as an important determinant of health as well. Particularly, individuals who are employed are healthier than those who are not employed (Yelln and Trupin, 2003; Thomas and Ellis, 2013; McKee-Ryan et al., 2005). Employment can have social, psychological, and financial benefits that improve health (Goodman, 2015). For instance, employment can increase household income and decrease economic hardship which consequently improve physical and psychological well-being (Goodman, 2015).

Job satisfaction is another positive aspect of working which may determine worker health. In fact, there can be a positive relationship between job satisfaction and subjective health measures which means that employer with higher or improved job satisfaction levels are healthier and more satisfied with their health (Fischer and Sousa-Poza, 2009). In other words, improvements in job satisfaction over time can prevent workers from further health deterioration (Fischer and Sousa-Poza, 2009).

Apart from positive aspects of working, it can be harmful for overall health status as well. For instance, workplace stress is a harmful physical and emotional response that

occurs when there is a conflict between job demands on the employee and the amount of control an employee has, to meet these demands (Canadian Center for Occupational Health and Safety, 2018). Long working hours, heavy workload, job insecurity, work environment can be common examples of work-relates stress that may lead to deterioration of health and well-being (Better Health, 2012). Moreover, work-related stress could have a negative effect on the health of employees and work organizations (Ippoliti et al., 2017). Discrimination at workplace can also have negative effects on health outcomes. Employment discrimination occurs when an employee is treated unfavorably due to his/her race, skin color, national origin, gender, disability, religion, or age (Doyle, 2019). Particularly, self-reported experience of racial discrimination can lead to poor or fair self-reported health; lower physical functioning; lower mental health; smoking; and cardiovascular disease (Harris et al. 2006).

In sum, if the risks of work are properly controlled and the demands of work are adjusted to match individual capacity exclusive of very specific conditions, the positive effects of work on health and well-being are greater than the risks (Waddell and Burton, 2006).

Health status is considered an important factor in determining labor supply decisions (Stern, 1989; Bound et al., 1999; Currie and Madrian, 1999; Au et al., 2005; Kalwij and Vermeulen, 2008). Canadian studies that have investigated the impact of health on labour market outcomes are cross-sectional or static (Breslaw and Stelener, 1987; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have not examined the dynamic interaction between health and labour market outcomes. Moreover, they have solved the endogeneity of health through a two-stage method or instrumental variable

(IV) approach which does not give an efficient estimate for health effect. A two- stage method does not take into account the endogeneity of health through the correlation of error terms and it may give incorrect inferences on the endogeneity of health (Cai, 2010). In fact, in a two-stage method, the endogeneity of health is only arising from the coefficient on employment status. Furthermore, using a simultaneous equation model facilitate finding instruments for health variable. In this regard, this study uses a sequential method in which the health status in the previous period can affect employment status at current time and the employment status at previous period can affect health status at current time. [5]This study contributes to the Canadian literature by estimating a dynamic panel data model through a simultaneous equation model to obtain more precise model specification for the interaction of health and labour market outcome.

Understanding the relationship between health and labor outcomes is even more important as more individuals reach the life stage where their health status gradually depreciates to the point where it can influence their labor market behavior (Currie and Madrian, 1999). Furthermore, a good understanding of how health influences labor force participation is also important for policy makers to understand the relationship between health and labor outcomes and apply more efficient policies to improve health status of labor market participants.

The rest of the study is organized as follows: Section 3.2 reviews the Canadian and International literature; section 3.3 presents an overview of the data and offers descriptive statistics; section 3.4 explains the methodology used in the study; section 3.5

[5] Please refer to the further explanation in the methodology section.

describes the empirical results; section 3.6 does some sensitivity analysis as a robustness check; section 3.7 concludes and provides some discussion for further research.

3.2. Literature Review

3.2.1. International Literature

The number of studies that have investigated the relationship between health and labor force outcomes is extensive. These studies can be categorized in three groups; the first group consider the one way causality from health to employment, which means that health is a key factor in determining labor market decisions (Currie and Madrian, 1999; Bound et al., 1999; Pelkowski and Berger, 2004; Lindeboom et al., 2006). Currie and Madrian, (1999) provides an overview of the empirical literature linking health and labor market outcomes to confirm that health has a persistent effect on labor market outcomes including wages, earnings, labor force participation, hours worked, retirement, job turnover and benefits package. They particularly indicate that health can affect wages through different ways. First, poor health may reduce productivity and lead to lower wages. Second, the employer costs of supporting a worker in poor health may be appeared in the form of lower wages and finally those in poor health have higher probability to face discrimination. In general, they conclude that health has an important role in labor market outcomes due to its direct impact on productivity and its indirect effect by trading off between income and leisure. However, they cannot reach an agreement on the magnitude of the health effect due to both the choice of health measures and to identification assumptions. They mention that the way health is measured is really important. In fact, they believe that estimates of the health effects of labor supply are highly sensitive to the measure used such as multiple measure or more comprehensive

measures (e.g., an indicator whether health limits the ability to work, a specific limitation on an activity during life time).For instance, for a cost/benefit analysis of a specific treatment they recommend to focus on a particular disease. However, if the goal is to understand the effect of health on hours worked, they suggest a broader definition of health. In absence of sensible identification assumptions, they recommend taking the" production function" approach to health and look into the medical determinants of various conditions. In this regard, they indicate to use some risk factors such as family history of a particular disease to explain health while being excluded from equations for labor market outcomes.

Using the first three waves of the longitudinal Health and Retirement Survey (HRS), Bound et al., (1999) investigate the dynamic relationship between health and labor market behavior for the sample of men and women between the age of 50 and 62 at the time of first wave. The results suggest that poor health makes many older workers quit their job and more than half of those who exit the labor force apply for disability insurance. Among those who continue working, many change their jobs within several years of the beginning of their poor health status. In addition, the authors find that when a health shock occurs in early ages, it is less likely to lead to labor force exits.

Pelkowski and Berger (2004) investigate the effect of health problems (temporary and permanent health conditions) on employment, annual hours worked and hourly wages. They use The Health and Retirement Survey (HRS) to examine the impact of temporary and permanent illness on labor market outcomes. The sample is the non-institutionalized men and women born in US between 1931 and 1941. The results indicate that current permanent health problems have significant negative effects on

average hourly wages of workers. In this regard, women reduce their wages slightly larger than men. Furthermore, permanent health problems result in reduction in the annual hours worked. Although women have a greater reduction in wages, men experience a larger reduction in hours worked. For men, the permanent health problems that occurred in the 40s have the largest negative effect, while women have the negative effects peak in the 30s. The authors mention that individuals experience the negative health problems in those age groups because the severity of the health shocks occurs near the peak of their life-cycle earnings. The results also show that temporary health conditions have little impact on either hourly wages or hours worked for both men and women. The authors emphasize that ignoring the health conditions as temporary and permanent underestimate the long-term impact of health problems on labor market outcomes.

Lindeboom et al. (2006) uses data from the British National Child Development Study (NCDS) for individuals from birth up to 42 years old in order to examine the relationship between the onset of disability and employment outcomes. The results suggest that using health shocks as an instrumental variable show a causal effect of disability on employment. In other words, when disability occurs at the age of 25, the employment rate is reduced at the age of 40 by about 21%. Furthermore, the authors find that those individuals who experience bad conditions during early childhood have a greater rate of health deterioration in their adulthood and they have higher tendency to be non-employed compared to the ones who have not experienced bad conditions in their early childhood.

The second group of studies focus on the causality from employment to health, which indicate that lack of employment can lead to physical or mental health problems (Clark and Oswald, 1994; Lucas et al., 2004; Bambra and Eikemo, 2009; Pharr, Moonie, and Bungum, 2011). Clark and Oswald (1994) use data from the British Household Panel Study from the 1990's to analyze a random sample of about 6000 British workers to investigate whether unemployed individuals are relatively happy or unhappy. The results demonstrate that unemployed individuals in Great Britain in 1991 have much lower levels of mental well-being compared to those who work. However, the data indicates that high unemployment rates across regions and age groups are correlated with relatively low disutility from joblessness. The authors also find that young workers and high-unemployment areas, like the North, have less distress from unemployment. Furthermore, people with long term unemployment status have less distress and are happier than those with short term unemployment status.

Lucas et al., (2004) use the set-point theories of subjective wellbeing (individuals show reaction to life events; however, they return to their baseline levels of satisfaction over time) to examine the effect of unemployment on people's life satisfaction. They use 15 years of the longitudinal German Socio-Economic Panel Study (SOEP) for more than 24000 individuals living in Germany. The authors find that people react strongly to unemployment and then they return to their baseline levels of life satisfaction. However, the individuals do not completely reach their former levels of satisfaction even after being reemployed implying that life events such as unemployment have a strong influence on long-term levels of personal well-being.

Using data from the first two waves (2002-2004) of the cross-sectional European Social Survey of individuals aged 25-65 for 23 countries (classified in five state regimes) Bambra and Eikemo (2009) examine whether the relationship between unemployment and increased risk of morbidity and mortality varies between welfare states which have different levels of social protection for the unemployed. The results demonstrate that in all countries, unemployed individuals report higher rates of poor health than those who are employed. In other words, the negative relationship between unemployment and health is consistent across Europe but varies moderately by welfare state regimes. Moreover, this negative relationship is stronger for women than men as unemployed women are more likely to receive lower than average wage replacement rates than unemployed men.

Pharr, Moonie, and Bungum (2011) take data from Nevada's 2009 Behavioral Risk Factor Surveillance System (BRFSS) to estimate the impact of employment status and unemployment duration on perceived health, access to health care and health risk behaviors. In this regard, they compare individuals who are unemployed (greater than and less than one year) to the ones who are employed and those who are voluntarily out of labor force (OLF). The results suggest that unemployed individuals have significantly lower level of perceived mental health compared to the employed and OLF participants. Moreover, unemployed individuals have a higher probability to delay health care services and they have a lower chance of access to health care than employed and OLF participants. In other words, after adjusting for age, income, gender, race and education, unemployed participants are about four times more likely to lack health care services and about two times more likely to delaying health care services compared to employed

participants. The authors also mention that OLF participants have better mental health scores and they are better off financially compared to those participants who are involuntarily unemployed. However, in terms of health risk behaviors, unemployed participants are not more likely to drink, smoke or be physically inactive.

In contrast to the previous studies that rely on the one-way causality from health or labour market outcomes, the recent studies mostly consider the two-way causality between health and labour market outcomes and they also control for the persistency of the health or labour market outcomes (Haan and Myck, 2009; Lindeboom and Kerkhofs, 2009; Cai, 2010; Cai, 2018; Delattre et al. 2019). For instance, Haan and Myck (2009) use 12 years (1996–2007) of data from the German Socio-Economic Panel (SOEP) on a sample of German men aged 30-59 to study the relationship between health and labour market risks (poor health and non- employment (all individuals who report working zero hours in their 'actual working week')).They use a bivariate dynamic random effects model to allow for the correlation in the unobserved heterogeneity of health and labor market risks. They find a significant interaction between poor health and non-employment. Particularly, they show that poor health at time t-1 is a significant determinant of current labour market risk and lagged non-employment has a positive effect on the probability of being in poor health at time t. They suggest that this interaction can be partly related to unobserved individual specific effects and to the effect of the lagged labour market risks and lagged poor health. In addition, they show a very strong level of persistency in health and labour market risks and conclude that accounting for unobserved effects reduces the magnitude of the estimated coefficients on the lagged

endogenous variables and significantly reduces the persistence of both health and labor market risks.

Using data from the first two waves (1993 and 1995) of the Leiden University Center for Research on Retirement and Aging (CERRA) panel survey, Lindeboom and Kerkhofs (2009) explore the intertemporal relationship between health and work decisions of older men workers in The Netherlands. They take into account the endogeneity of the health variable and use simulated maximum likelihood techniques to control for the unobserved heterogeneity that affect the two outcomes. The results confirm that financial incentives are as important factors in explaining retirement decisions. In other words, pension and social security reforms which are planned to increase the labor force participation of older workers have adverse effects on health of elderly people as the increased working efforts at older ages lead to a deterioration of health.

Using data from the first four waves of the Household, Income and Labor Dynamics in Australia (HILDA) Survey, Cai (2010) investigates the relationship between health and labor force status among a sample of working aged Australian men and women (men aged 25-64 and women 25-59). The author uses a simultaneous equation method to examine this relationship, allowing for the endogeneity of health to avoid overestimation of health effects on labor force participation. The results indicate that health has a positive and significant effect on labor force participation for both men and women. However, the effect of labor force participation on health is different for men and women. Specifically, labor force participation has a negative effect on health for men while this effect is positive for women.

Using the Household, Income and Labour Dynamics in Australia (HILDA) Survey, Cai (2018) estimates a dynamic probit model for married Australian women. He shows that both observed individual heterogeneity such as education and long health conditions and unobserved individual heterogeneity such as preferences between work and leisure, motivation and ability may explain the persistence of labour force participation of married Australian women. The author also finds that ignoring serially correlated unobserved transitory shocks such as deterioration in an individual's health may lead to underestimation of state dependence of labour force participation. Moreover, he mentions that the state dependence labour force participation varies with age, education, health, immigration status and number of children under the age school. The author concludes that due to a positive state dependence in married Australian women's labour force participation, policy intervention that increases married women's labour supply would have a long- lasting effect.

Delattre et al. (2019) examine the causal relationship between health and employment status for individuals between the age of 20 and 74. Using French longitudinal survey on health and work, they estimate a bivariate dynamic probit model to take into account the persistence effects, initial conditions and unobserved heterogeneity. The authors find persistence in both health and employment status. They also demonstrate that being health at t-1 is a significant determinant of current labour market status, and lagged employment has a significant effect on the probability of being ill at time t. The authors conclude that since there is a strong dual causality between health and employment status, the joint design of health and employment policies can improve health and employment simultaneously.

3.2.2. Canadian Literature

Studies that investigate the relation between health and employment on Canadian population are rare and they are mostly examine the interaction between disability and employment of Canadians (Hum and Simpson, 1996; Campolieti, 2002; Oguzoglu, 2012). A few studies have investigated the relationship between health and employment. For instance, Breslaw and Stelener (1986) use data from the Canada Health Survey (CHS) in 1979 to examine the influence of health on labour force behavior of elderly men (age 50 and over). They conclude that health is a strongly significant variable in influencing the labour market state of the elderly men. However, the authors have not considered the endogeneity of health. Moreover, since the data they have used is cross-sectional, they have not controlled for unobserved heterogeneity as well. Some studies have taken into account the endogeneity of health and use static panel data model to control for unobserved heterogeneity. For instance, Au et al. (2005) uses four cycles of longitudinal National Population Health Survey (NPHS) from 1994-95 to 2000-2001 for older Canadians (age 50 and over) to study the relationship between health and employment. For this purpose, they use different measure of health such as self-assessed health, Health Utilities Index Mark 3(HUI3) and a purged health measure (estimated health shock) and compare the effect of health on employment. They find statistically significant effects of health on employment for both men and women in all model specifications (OLS and fixed effects (IV & reduced form)). In other words, they report that a one standard deviation increase in health raises the probability of employment between three to eighteen percentage points based on the sample and specification. In addition, they show that health effects using more objective health measures (HUI3 or estimated health shock) are very similar to those estimates that using self-assessed health

measure. The authors find this surprising and suggest that justification bias arising from self-assessed health measure may be outweighed by a counteracting attenuation bias caused by measurement error. They also find smaller and significant health effects when the models are estimated with fixed effects compared to those estimates in levels for both men and women. Thus, they conclude that health shocks are important in the work decision.

Hum, Simpson and Fissuh (2008) use different Canadian data set to estimate the impact of health on total annual hours worked. They use longitudinal Survey of Labour and Income Dynamics (SLID) from 1996 to 2001 and compare the health effects among different age groups. They estimate pooled, RE and FE models to examine the impact of unobserved effects for exogenous and endogenous health for different men age groups. They conclude that health impacts are smaller when health is considered as endogenous and still smaller when the FE model is estimated. For instance, they report that for young men 21-34, the impact of good health on total hours worked decreases from 589 hours (31.4 percentage points) with the RE model when health is exogenous to 395 hours or 21.1 percentage points when an instrument is used for health. The health impact also decreases more to 80 hours or 4.2 percentage points when the FE model is used with the same health instrument. For middle-aged men 35-50, they report that the estimated impact of good health for the RE model with exogenous health decreases from 495 hours or 24.9 percentage points to 312 hours or 15.7 percentage points when an instrument is used for health. The estimated impact is even decreased further to 86 hours or 4.3 percentage points when the FE model is used with the same instrument for health. For older men 51-65, they report that with the RE model when health is exogenous, the

impact of good health decrease from 839 hours or 61.3 percentage points to 494 hours or 36 percentage points when an instrument is used for health. With the FE model, the estimated impact is decreased more to 83 hours or 6 percentage points with the same health instrument. Thus, they suggest that however the effect of good health on men labour supply is small (with FE model and when health is treated as endogenous), it is widespread among the population and increases only modestly in older men.

According to the literature we can conclude that in order to investigate the relationship between health and employment, endogeneity of health have to be considered. In other words, taking into account health as exogenous can lead to biased and inconsistent estimates of the impact of health on employment.

Endogeneity is a common issue in many areas of applied economics including health economics and health services research (Terza, Basu and Rathouz, 2008). Endogeneity problem might occur when an explanatory variable is correlated with error term (Wooldridge, 2010). Typical sources of endogeneity might arise from measurement error, simultaneity, omitted variable bias or unobserved heterogeneity (Wooldridge, 2010). Measurement error occurs when the variable of interest is measured with error and we can only observe an imperfect measure of it in a statistical model (Wooldridge, 2010). Omitted variable bias is where an explanatory variable is omitted from the model and included into the error term. In other words, when we would like to control for additional variables, but data is not available, the variables cannot be estimated in a regression model and they become part of the error term (Wooldridge, 2010). Simultaneity is the case where causal relationship between the dependent and explanatory variable occurs in both directions. Particularly, at least one of the explanatory variables is determined

jointly along with dependent variable and unobserved heterogeneity is a source of endogeneity where there is unobserved characteristic that influence both dependent and explanatory variables and this unobserved characteristic is included into the error term. (Wooldridge, 2010).

In this study, endogeneity of health is coming from simultaneity and unobserved heterogeneity. According to the literature analyzing the two-way causality between health and labor market outcomes (e.g. Haan and Myck, 2009; Cai, 2010; Lindeboom and Kerkhofs, 2009) there are two ways to solve the endogeneity of health to employment; A two-stage method or instrumental variable (IV) approach and a simultaneous equation model. Although the IV approach can give a consistent estimate for the health effect, the source of endogeneity of health cannot be examined by this approach (Cai, 2010). This approach is also not efficient since it does not take into account the correlation between health and employment equations (Cai, 2010). Moreover, by IV approach, exogeneity of health cannot be examined (Cai, 2010). In contrast, simultaneous equation model does consider the correlation of health and employment equations through their error terms and can give an efficient estimate for health effect. In addition, exogeneity of health can be examined. Considering the advantage of a simultaneous equation model, this study takes this approach to solve the endogeneity of health to employment to obtain an efficient estimate for health effect.

Unobserved heterogeneity is another issue arising from cross-sectional studies that leads to biased estimates of the impact of health on employment. Thus, using longitudinal data can solve the biased estimates of health effect. This study uses

longitudinal National Population Health Survey (NPHS) in order to remove the biased estimated of health effects arises from cross-sectional studies.

Another important issue that have to be considered in regard to interaction between health and employment is state dependency of health and employment status. Particularly, individuals who are in good health status at previous period are more likely to be in good health status at current time. Moreover, those individuals who are employed at previous period are more likely to be employed at current time. Thus, in order to have more precise estimation, dynamics of health and employment status should be taken into account as well (the importance of this fact is shown in the methodology section).

As already mentioned, Canadian studies that have investigated the impact of health on labour market outcomes are cross-sectional or static (e.g. Breslaw and Stelener, 1986; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have not examined the dynamic interaction between health and labour market outcomes. Moreover, they have solved the endogeneity of health through a two-stage method or instrumental variable (IV) approach which does not give an efficient estimate for health effect. This study contributes to the Canadian literature by estimating a dynamic panel data model through a simultaneous equation model to obtain more precise model specification for the interaction of health and labour market outcome.

3.3. Data and Variables

The data is taken from 9 cycles (1994/95- 2010/11) of longitudinal data of National Population Health Survey (NPHS). The NPHS collects information related to the health of the Canadian population and socio-demographic information every two years. This study is restricted to the individuals between the age of 25 and 64. The sample

contains 33,186 and 33,863 year-person observations for men and women, respectively. Table 3.1 and Table 3.2 summarize the demographic characteristics of the sample. The variables include age, sex, marital status, immigration status, education level, dwelling owned by household, capital income (dividends and interest), household size, number of children and province of residence. The educational level has four categories: less than secondary, secondary school graduation, some post-secondary, and post- secondary graduation. The dummy variable own indicates whether dwelling owned by a household or not. The dummy variable capital income shows if the individual received any income in the past 12 months (dividends and interest) or not. Household size is categorized into five groups (hsize1- hsize5). Hsize5 refers to household size 5 and up.

Particularly, there is a question in NPHS' questionnaire where the individual is interviewed to rate its own health status; "In general, would you say your health is: Excellent, very good, good, fair, and poor". The study classifies individuals into two categories: "good health" when they announce their health status as excellent, very good and good and "poor health" when they are belonging to the fair or poor group.

The dummy variable Employed indicates whether an individual is employed or non-employed (not employed or not in labor force).

There would be many variables that affect health and employment. However, a restricted number of variables is used in the estimations to facilitate the convergence of our non-recursive bivariate probit model.[6]

[6] A table including list of variables used in the estimation is defined on Appendix 3.B.

Table 3.1: Summary Statistics for Men[7]

Description	All	Mean (If E=1)	Mean (If E=0)	Mean (If H=1)	Mean (If H=0)	Mean (If E &H=1)	Mean (If E&H=0)
Employed	0.8496	1	0	0.8769	0.5351	1	0
Healthy	0.9200	0.9497	0.7530	1	0	1	0
Has capital income	0.2229	0.2242	0.2155	0.2291	0.1512	0.2260	0.1070
Lives in BC	0.1177	0.1181	0.1157	0.1179	0.1157	0.1175	0.0999
Lives in Ontario	0.3680	0.3772	0.3166	0.3679	0.3700	0.3760	0.3370
Lives in Quebec	0.2639	0.2514	0.3346	0.2646	0.2557	0.2533	0.3028
Lives in NF, PEI, NB, NS	0.0810	0.0738	0.1214	0.0781	0.1145	0.0733	0.1511
Lives in ALB, MB or SASK	0.1694	0.1796	0.1118	0.1716	0.1441	0.1799	0.1090
Married	0.6243	0.6385	0.5441	0.6251	0.6149	0.6373	0.5623
Owns a house	0.7673	0.7818	0.6851	0.7729	0.7018	0.7835	0.6453
Less than secondary school	0.1538	0.1293	0.2921	0.1416	0.2946	0.1245	0.3790
Secondary school graduated	0.1312	0.1357	0.1057	0.1327	0.1146	0.1360	0.0964
Has some post-secondary	0.2648	0.2637	0.2710	0.2650	0.2625	0.2633	0.2529
Post-secondary graduated	0.4502	0.4713	0.3311	0.4608	0.3283	0.4762	0.2715
Immigrant	0.1752	0.1782	0.1579	0.1735	0.1942	0.1749	0.1400
Household size is equal to 1	0.1390	0.1271	0.2062	0.1344	0.1917	0.1263	0.2481
Household size is equal to 2	0.2882	0.2667	0.4096	0.2837	0.3393	0.2660	0.4080
Household size is equal to 3	0.2010	0.2042	0.1829	0.2016	0.1945	0.2033	0.1633
Household size is equal to 4	0.2320	0.2523	0.1169	0.2377	0.1659	0.2546	0.1160
Household size is equal or greater than 5	0.1002	0.1083	0.0545	0.1025	0.0735	0.1087	0.0427
Number of Children	0.4881	0.5312	0.2447	0.5081	0.2589	0.5412	0.1624
Age (Age/10)	4.3845	4.2804	4.9728	4.3415	4.8801	4.2617	5.1635
Age^2 (Age^2/100)	20.3852	19.3596	26.1789	19.9903	24.9329	19.1886	27.6332
Observations	31,044	25,667	5,377	28,258	2,786	24,323	1,442

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

[7] In this chapter, the sample weight provided for the NPHS is used for the summary statistics and all the estimations.

Table 3.2: Summary Statistics for Women

Description	All	Mean (If E=1)	Mean (If E=0)	Mean (If H=1)	Mean (If H=0)	Mean (If E &H=1)	Mean (If E&H=0)
Employed	0.7209	1	0	0.7499	0.4360	1	0
Healthy	0.9076	0.9441	0.8133	1	0	1	0
Has capital income	0.2048	0.2017	0.2127	0.2120	0.1342	0.2051	0.1264
Lives in BC	0.1261	0.1294	0.1177	0.1255	0.1320	0.1279	0.1142
Lives in Ontario	0.3662	0.3778	0.3364	0.3654	0.3748	0.3760	0.3495
Lives in Quebec	0.2538	0.2408	0.2876	0.2572	0.2207	0.2453	0.2649
Lives in EAST	0.0874	0.0791	0.1091	0.0856	0.1055	0.0787	0.1217
Lives in ALB, MB or SASK	0.1663	0.1730	0.1492	0.1663	0.1670	0.1721	0.1498
Married	0.6118	0.5926	0.6615	0.6203	0.5284	0.5965	0.5300
Owns a house	0.7520	0.7655	0.7173	0.7625	0.6489	0.7697	0.6140
Less than secondary school	0.1373	0.0887	0.2626	0.1215	0.2920	0.0832	0.3765
Secondary school graduated	0.1582	0.1507	0.1776	0.1575	0.1651	0.1498	0.1645
Has some post-secondary	0.2654	0.2677	0.2595	0.2660	0.2597	0.2675	0.2501
Post-secondary graduated	0.4391	0.4929	0.3003	0.4550	0.2832	0.4996	0.2089
Immigrant	0.1812	0.1734	0.2014	0.1767	0.2255	0.1688	0.2057
Household size is equal to 1	0.1113	0.1106	0.1131	0.1057	0.1670	0.1090	0.1887
Household size is equal to 2	0.3133	0.2920	0.3683	0.3079	0.3663	0.2916	0.4182
Household size is equal to 3	0.2236	0.2342	0.1963	0.2257	0.2034	0.2343	0.1803
Household size is equal to 4	0.2179	0.2303	0.1857	0.2237	0.1602	0.2319	0.1270
Household size is equal or greater than 5	0.0926	0.0942	0.0885	0.0953	0.0665	0.0945	0.0485
Number of Children	0.5111	0.4982	0.5445	0.5299	0.3269	0.5034	0.2613
Age (Age/10)	4.3639	4.2149	4.7486	4.3228	4.7670	4.2017	5.0203
Age^2 (Age^2/100)	20.2148	18.7424	24.0185	19.8473	23.8256	18.6281	26.263
Observations	36,005	25,379	10,626	32,366	3,639	23,900	2,160

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

Table 3.1 provides the demographic characteristics of the sample for men. The first column shows the descriptive statistics for the whole sample men. Column 2 and column 3 are classified by employment status (Employed/non-employed (not employed or not in labour force), column 4 and 5 shows the descriptive statistics of men classified by health status (healthy/unhealthy (fair or poor)) and the last two columns are for the men who are employed and healthy and non-employed and unhealthy.

The average age of men in the sample is 43.85 years. In this sample, 92 percent of the men reported at least a good health status and 85 percent reported to be employed. Based on the column 2 and column 3, 95 percent of men who are employed are healthy while this will decrease to 75 percent for those men who are non-employed. It is obvious that individuals who are unhealthy are less likely to be employed compared to the ones who are healthy. The employment rate for healthy men is 88 percent while 54 percent of unhealthy men are employed. Furthermore, unhealthy and non-employed men are likely to be immigrant, older and less educated compared to those men who are healthy and employed. Table 3.2 represents the descriptive statistics for women. The average age of women in the sample is 43.64 years. In this sample, 91 percent of the women reported at least a good health status and 72 percent reported to be employed. Similar to men, 95 percent of women who are employed are healthy; however, 81 percent of non-employed women are healthy. Unhealthy women also have lower rate of employment compared to the ones who are healthy (75 and 44 percent respectively). Moreover, women who are unhealthy and non-employed are likely to be immigrant, older and less educated to those who are healthy and employed.

3.4. Methodology

The present study uses a simultaneous equation model to measure a dynamic interaction between health and employment. Particularly, a simultaneous equation model is a solution for endogeneity of health and can take into account unobservable individual characteristics that might jointly affect health and employment. Thus, it can give an efficient estimate for the impact of health on employment. In other words, controlling for endogeneity of health can lead to unbiased and consistent estimate for health effects of employment.

In addition to controlling for endogeneity of health, state dependency of health and employment has been considered. In contrast to static or cross-sectional models, dynamic models control the lags of the dependent variable as explanatory variables and this can help to obtain a more precise model specification. For instance, using the Irish component of the European Community Household Panel Survey 1995-2000, Gannon (2005) analyzes the dynamic effects of disability on participation in the labour force. Using a dynamic model and controlling for unobserved heterogeneity and past participation, she demonstrates that static models overestimate the effect of disability on participation between 40-60% for men and 5-10% for women. Oguzoglu (2012) uses the first 8 cycles of the NPHS and examines the dynamic impact of disability on labour force participation in Canada. He concludes that high persistence in labour force participation makes disability shock influences participation for multiple periods while static models by ignoring this persistency, underestimate the overall effect of disability. Using the German Socio-Economic Panel (SOEP) Survey 1996-2007, Haan and Myck (2009) estimate the dynamic relationship between health and labour market risks. They show that poor health on previous period is a significant determinant of current labour market

risk. In addition, the lagged non-employment has a positive effect on the probability of being in poor health at current time. They also find a very strong level of persistency in health and labour market risks.

As already mentioned, Canadian studies that have investigated the impact of health on labour market outcomes are cross-sectional or static (e.g. Breslaw and Stelener, 1986; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have not examined the dynamic interaction between health and labour market outcomes. Moreover, they have solved the endogeneity of health through a two-stage method or instrumental variable (IV) approach which does not give an efficient estimate for health effect. This study contributes to the Canadian literature by estimating a dynamic panel data model through a simultaneous equation model to obtain more precise model specification for the interaction of health and labour market outcome.

In line with Haan and Myck (2009), this study uses the simultaneous equations to estimate the relationship between heath and employment:

$$P(E_{it}=1) = f(\alpha_2 + \gamma_2 H_{i\,t-1} + \gamma_3 E_{i\,t-1} + \Upsilon_2 X_{it} + u_i + \phi_t) \qquad (1)$$

$$P(H_{it}=1) = f(\alpha_1 + \beta_2 H_{i\,t-1} + \beta_3 E_{i\,t-1} + \Upsilon_1 X_{it} + v_i + \epsilon_{it}) \qquad (2)$$

In the equations above, H refers to self-perceived health (SPH) for measuring health status. Self-perceived health which is also known as self-assessed health (SAH) or self-rated health (SRH) refers as an indicator to measure health status. In other words, individuals rate their current health status based on a scale ranging from 1-4 or 5 (excellent-poor). Although there are studies which indicate that SAH may lead to bias compared to more objective measure of health (Bound, 1989; Bound et al., 1999), there is an extensive literature which confirm that self- reported health is a good predictor of

mortality and morbidity (Idler and Benyamini, 1997; Damian et al., 1999; Goldberg et al., 2001; Larsson et al., 2002). Particularly, there is a question in NPHS' questionnaire where the individual is interviewed to rate its own health status ("In general, would you say your health is: Excellent, very good, good, fair, and poor"), this study classifies individuals into two categories: "good health" when they announce their health status as excellent, very good and good and "poor health" when they are belonging to the fair or poor group. E indicates employment status. In NPHS, current labor force status (LFS) is categorized into three groups: employed, unemployed and not in labor force. In this study the dummy variable E_{it} is equal to 1 if the individual is employed and 0 if he/she is not employed or not in labor force, H_{t-1} and E_{t-1} denote the SPH and the employment status in the previous period, respectively. Similar to Haan and Myck (2009), it is supposed that the previous period of health can affect the current period of employment in equation (1) and the previous period of employment can also affect the current period of health in equation (2). In other words, when health is considered as endogenous, unobserved factors related to health conditions may be correlated with unobserved factors for the employment equation. Thus, a univariate panel probit model fails to show the consistent estimates of the effects of health on employment. In addition, employment coefficient for men and women can be overestimated or underestimated. X_{it} is a vector of exogenous variables such as age, sex, marital status, immigration status, education level, dwelling owned by household, capital income (dividends and interest), household size, number of children and province of residence. The v_i and u_i indicate individual specific unobserved heterogeneity and ϵ_{it} and ϕ_t are the random disturbance terms.

This study uses random effects for the panel data model. In fact, due to the non-linear nature of the model, v_i and u_i cannot be eliminated through differencing or within transformation. Furthermore, adding N-1 individual dummy variables to the probit model to estimate v_i and u_i will cause severely biased estimates. This is known as incidental parameters problem.

The main issue to estimate equation (1) or (2) is that the unobserved effects v_i and u_i are correlated with $H_{i\,t-1}$ and $E_{i\,t-1}$ respectively by definition. This is usually called the initial condition problem which can result in severe bias. This study uses the approach suggested by Wooldridge (2005) to solve the initial conditions problem. In this approach, u_i can be written as a function of E_{i0} (the employment status from the first cycle) and X_i while the unobserved effect v_i is as function of H_{i0} (the health status from the first cycle) and Xi. This approach takes into account the employment or health status from the first cycle as well as time averages of X_{it} as additional regressors.

In other words, Wooldridge suggests modelling the distribution of $\{E_{i1}, \ldots, E_{iT}\}$ given E_{i0} by using conditional Maximum likelihood approach. For this purpose, the distribution for u_i given E_{i0} and other exogenous variables can be specified as follows:

$$f(E_{i1}, \ldots, E_{iT} \mid E_{i0}) = \int f(Ei1, \ldots, EiT \mid E_{i0}, X_{it,} u_i)\, g(u_i \mid E_{i0}, X_{it})\, du_i$$

$$g(u_i \mid E_{i0}, X_{it}) \sim N(u_i = \lambda_2 \bar{X}_i + \gamma_1 E_{i0}, \sigma^2_\delta)$$

As in the panel probit with random effects, the assumption is that there is no correlation between unobserved heterogeneity effect (u_i) and the explanatory variables, the variable $\bar{X}_i$ indicates time averages of all time-varying variables in X_{it} to relax this limitation and allow for the correlation between unobserved effect and explanatory variables. Particularly, we can write

$u_i = \lambda_2 \bar{X}_i + \gamma_1 E_{i0} + \delta_i$; $\delta_i \sim N(0, \sigma^2_\delta)$ and it is independent of E_{i0} and $X_{i.}$ (3)

The same process can be written for the unobserved effect (v_i) that influence the health status (H_{it}), so the final equation suggested by Wooldridge is as follows:

$v_i = \lambda_1 \bar{X}_i + \beta_1 H_{i0} + \vartheta_i$; $\vartheta_i \sim N(0, \sigma^2_\vartheta)$ and it is independent of H_{i0} and $X_{i.}$ (4)

If the equations (3) and (4) substituted into equations (1) and (2) respectively, the following equations can be reached:

$P(E_{it}=1) = f(\alpha_2 + \gamma_2 H_{i\,t-1} + \gamma_3 E_{i\,t-1} + \Upsilon_2 X_{it} + \lambda_2 \bar{X}_i + \gamma_1 E_0 + \delta_i + \phi_t)$ (5)

$P(H_{it}=1) = f(\alpha_1 + \beta_2 H_{i\,t-1} + \beta_3 E_{i\,t-1} + \Upsilon_1 X_{it} + \lambda_1 \bar{X}_i + \beta_1 H_0 + \vartheta_i + \epsilon_{it})$ (6)

Consequently, equations (5) and (6) can be estimated as a bivariate dynamic panel probit model.

3.5. Empirical Findings

As an initial estimation, a static panel probit model can be estimated. Table 3.3 shows the results of a static panel probit model in which the dynamic nature of the individual's employment status and individual's health status has not been considered. Results show that there is a significant and positive interaction between health and employment variables for both men and women. In other words, in the health equation, the coefficient for the latent employment (probability of employment) is positive and significant for both men and women (0.1350 vs. 0.0537). This is also in line with the literature that lack of employment can lead to physical and mental health problems (Clark and Oswald, 1994; Lucas et al., 2004; Bambra and Eikemo, 2009; Pharr, Moonie, and Bungum, 2011). On the other hand, in the employment equation, the coefficient for the latent health (probability of health) is significant and positive for men and women (0.3975 vs. 0.3831). This confirms the literature that better health increases the

probability of employment (Currie and Madrian, 1999; Bound et al., 1999; Pelkowski and Berger, 2004; Lindeboom et al., 2006).

Other explanatory variables have the expected signs for both men and women. For instance, education levels are positively correlated with the health and employment variables. In the health equation, the probability of being healthy decreases when age increases and the negative sign for age square show that health variable is convex in age. On the other hand, the probability of being employed increases when age increases, and employment variable is concave in age. Number of children has a positive and significant effect on the probability of health variable while the effect on the probability of employment is negative and significant for both men and women.

As it was mentioned earlier, exogeneity of health can be tested through a simultaneous equation model. On the end of table 3.3, a parameter called atanhrho has been shown. This parameter represents the correlation between error terms in the bivariate panel probit model. The value for this parameter is negative and highly significant for both men and women respectively (-0.3005 vs. -0.2299) which confirms the endogeneity between health and employment. Furthermore, it indicates that the equations for health and employment variables should be estimated simultaneously in order to have unbiased and consistent coefficients.

Table 3.3: Static Bivariate Panel Probit Estimation

	Coefficients for Men	Coefficients for Women
Health Equation		
$Employed_t$	0.1350***	0.0537***
	(0.00004)	(0.00008)
Has Capital Income	0.1750***	0.2178***
	(0.00002)	(0.00002)
Lives in BC	-0.1495***	-0.3477***
	(0.00004)	(0.00005)
Lives in Ontario	-0.2199***	-0.3001***
	(0.00002)	(0.00005)
Lives in East	-0.3613***	-0.3297***
	(0.00007)	(0.00007)
Lives in Prairies	-0.1330***	-0.3675***
	(0.00004)	(0.00005)
Married	-0.0371***	0.2919***
	(0.00004)	(0.00004)
Owns a house	0.0262***	0.2055***
	(0.0000)	(0.00002)
Secondary school graduated	0.4497***	0.5154***
	(0.00004)	(0.00006)
Has some post-secondary	0.4511***	0.5570***
	(0.00003)	(0.00005)
Post-secondary graduated	0.6207***	0.8277***
	(0.0000)	(0.00009)
Immigrant	-0.0976***	-0.2947***
	(0.00002)	(0.00003)
Household size is equal to 2	0.2352***	0.0247***
	(0.00002)	(0.00004)
Household size is equal to 3	0.2070***	0.0305***
	(0.00002)	(0.00003)
Household size is equal to 4	0.1751***	0.0240***
	(0.00002)	(0.00003)
Household size is equal or greater than 5	0.1199***	0.1419***
	(0.00003)	(0.00004)
Number of Children	0.1117***	0.0311***
	(0.00001)	(0.00004)
Age	-0.4926***	-0.6600***
	(0.0000)	(0.0000)
Age2	0.0232***	0.0456***
	(0.000009)	(0.00001)
Constants	3.2988	3.4528
Employment Equation		
$Healthy_t$	0.3975***	0.3831***

	(0.0000)	(0.00008)
Has Capital Income	-0.1609***	-0.1575***
	(0.00001)	(0.00002)
Lives in BC	0.2327***	0.3436***
	(0.00003)	(0.00004)
Lives in Ontario	0.3293***	0.4394***
	(0.00002)	(0.0000)
Lives in East	-0.1430***	0.0139***
	(0.00006)	(0.0000)
Lives in Prairies	0.6443***	0.4621***
	(0.00004)	(0.00005)
Married	0.4785***	-0.4595***
	(0.0000)	(0.00003)
Owns a house	0.2455***	0.1062***
	(0.00001)	(0.00002)
Secondary school graduated	0.3841***	0.5008***
	(0.00004)	(0.00006)
Has some post-secondary	0.2233***	0.6345***
	(0.00003)	(0.00006)
Post-secondary graduated	0.4569***	0.8757***
	(0.00004)	(0.0009)
Immigrant	0.0446***	-0.1162***
	(0.00002)	(0.00004)
Household size is equal to 2	-0.0861***	-0.0231***
	(0.00002)	(0.00002)
Household size is equal to 3	-0.0442***	-0.0078***
	(0.00002)	(0.00002)
Household size is equal to 4	0.0870***	-0.0374***
	(0.00002)	(0.00002)
Household size is equal or greater than 5	0.0174***	-0.0826***
	(0.00003)	(0.0002)
Number of Children	-0.1481***	-0.3811***
	(0.0001)	(0.00001)
Age	2.9567***	3.2171***
	(0.00006)	(0.00007)
Age2	-0.3780***	-0.4005***
	(0.000006)	(0.000006)
Constants	-5.2946***	-6.2673***
	(0.00019)	(0.00035)
atanrho	-0.3005***	-0.2299***
	(0.00002)	(0.00001)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

By controlling the lag of health and the lag of employment on the health and employment equations respectively, a dynamic bivariate panel probit model can be reached. Table 3.4 shows the results of a dynamic bivariate panel probit model in which both health and employment variables are considered endogenous. Results demonstrate that there is a high state dependency in employment variable for both men and women. The coefficient for $Employed_{t-1}$ is positive and significant in employment equation for both men and women respectively (1.2109 vs. 1.2325). Thus, individuals who were employed in the previous period are more likely to be currently employed. In the health equation, the coefficient for $healthy_{t-1}$, is also positive and significant for both men and women (0.6271 vs. 0.5883); however, it is much lower compared to the coefficient of $Employed_{t-1}$ in employment equation. This may indicate that there is also state dependency for health variable for both men and women, but it is not as strong as that of for the employment variable.

Specifically, even by controlling the lag of employment in the employment equation, there is still a significant and positive interaction between health and employment for both men and women. However, the coefficient of the latent health is smaller than that of in the static model for both men (0.3246 vs. 0.3975) and women (0.1274 vs. 0.3831). This may indicate that the static model overestimates the health effect for both groups. On the other hand, by controlling the lag of health in the health equation, the latent employment has still positive and significant effect on the latent health for both men and women respectively (0.2257 vs. 0.2207). However, this time the coefficient for the latent employment is greater compared to its counterpart in the static model for men (0.2257 vs. 0.1350) and women (0.2207 vs. 0.0537).

As it was shown in the dynamic model, the lag of employment and the lag of health variable in the health and employment equations have been controlled respectively. In the employment equation, being healthy at previous period increases the probability of employment for both men and women; however, this effect is much higher for women than men (0.1800 vs. 0.0113). In the health equation, being employed at previous period has a negative and significant effect on the probability of health for both men and women respectively (-0.1549 vs. -0.1650).

Table 3.4: Dynamic Bivariate Panel Probit Estimation

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	1.1596***	1.3016***
	(0.00004)	(0.00003)
$Healthy_{t-1}$	0.6271***	0.5883***
	(0.00003)	(0.00002)
$Employed_t$	0.2257***	0.2207***
	(0.00004)	(0.00003)
$Employed_{t-1}$	-0.1549***	-0.1650***
	(0 .0000)	(0.0000)
Has Capital Income	0.0393***	0.1584***
	(0.00002)	(0.00002)
Lives in BC	-0.2557***	-0.2895***
	(0.00003)	(0.00003)
Lives in Ontario	-0.2310***	-0.2888***
	(0.00002)	(0.00002)
Lives in East	-0.1810***	-0.1614***
	(0.00006)	(0.00006)
Lives in Prairies	-0.2305***	-0.3318***
	(0.00003)	(0.00003)
Married	-0.0752***	0.0976***
	(0.00004)	(0.00004)
Owns a house	-0.1771***	0.0428***
	(0.00003)	(0.00003)
Secondary school graduated	0.2106***	0.1930***
	(0.0000)	(0.00003)
Has some post-secondary	0.1921***	0.1984***
	(0.00003)	(0.00003)
Post-secondary graduated	0.2279***	0.2987***
	(0.00003)	(0.00003)
Immigrant	-0.0791***	-0.2080***
	(0.0000)	(0.00002)
Household size is equal to 2	0.1791***	0.0566***
	(0.00002)	(0.00002)
Household size is equal to 3	0.1028***	0.0397***
	(0.00002)	(0.00002)
Household size is equal to 4	0.0702***	-0.0011***
	(0.00003)	(0.00003)
Household size is equal or greater than 5	-0.0648***	0.0988***
	(0.00003)	(0.00003)
Number of Children	0.0307***	0.0445***
	(0.00001)	(0.00001)
Age	-0.7888***	-0.6947***
	(0.00013)	(0.00014)
Age2	0.0556***	0.0583***
	(0.00001)	(0.00001)
Constants	0.2824***	3.3035***
	(0.00029)	(0.00033)

Employment Equation		
Initial Employed	0.5855***	0.6679***
	(0.00002)	(0.00002)
Employed$_{t-1}$	1.2109***	1.2325***
	(0.00003)	(0.00001)
Healthy$_t$	0.3246***	0.1274***
	(0.00004)	(0.00002)
Healthy$_{t-1}$	0.0113***	0.1800***
	(0.00006)	(0.00003)
Has Capital Income	-0.1635***	-0.0989***
	(0.00002)	(0.00001)
Lives in BC	0.1076***	0.1137***
	(0.00002)	(0.00002)
Lives in Ontario	0.1505***	0.1684***
	(0.00002)	(0.00002)
Lives in East	-0.0825***	-0.0804***
	(0.00005)	(0.00004)
Lives in Prairies	0.3412***	0.1576***
	(0.00003)	(0.00002)
Married	0.2080***	-0.2762***
	(0.00003)	(0.00003)
Owns a house	0.0435***	-0.0237***
	(0.00003)	(0.00002)
Secondary school graduated	0.2311***	0.2123***
	(0.00003)	(0.00002)
Has some post-secondary	0.1499***	0.2485***
	(0.00003)	(0.00002)
Post-secondary graduated	0.2927***	0.4035***
	(0.00003)	(0.0000)
Immigrant	0.0845***	-0.0098***
	(0.00002)	(0.00001)
Household size is equal to 2	-0.0392***	-0.0270***
	(0.00002)	(0.00002)
Household size is equal to 3	0.0577***	0.0091***
	(0.00002)	(0.00002)
Household size is equal to 4	0.1561***	0.0339***
	(0.00002)	(0.00002)
Household size is equal or greater than 5	0.0692***	0.0975***
	(0.00003)	(0.0000)
Number of Children	-0.0761***	-0.2169***
	(0.0000)	(0.00001)
Age	1.7989***	2.1860***
	(0.00011)	(0.00009)
Age2	-0.2337***	-0.2681***
	(0.00001)	(0.000009)
Constants	-4.5199***	-6.0613***
	(0.00020)	(0.00018)
atanhrho	-0.3797***	-0.1187***
	(0.00004)	(0.00003)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Employed$_{t-1}$ refers to the employment status from the previous period. Initial Employed is the employment status at first cycle. Healthy$_{t-1}$ refers to the health status from the previous

period. Initial Healthy is the health status at first cycle. Model includes time average of all time -varying control variables. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Since the model is nonlinear, the estimated coefficients in Table 3.4 can be useful just for the sign and significance. In order to interpret the results, marginal effects (ME) should be reported. In other words, ME measures how much predicted probabilities change as the binary independent variable changes from 0 to 1 after controlling all other variables at their means.

Table 3.5 shows the marginal effect of probability of employment arising from unit change in covariate, by holding all other covariates at their mean values.

The marginal effects of the employment equation can be briefly summarized as follows: The probability of being employed for men, increases 2.32 percentage points when the individual health status in the previous period changes from a poor health status to a good health status. For women, the probability of employment increases 5.89 percentage points when the health status in the previous period goes from poor to a good health status. This confirms that the health status at previous period has a significant and positive effect on probability of employment for both men and women.

Individuals who are employed at previous period are also more likely to be employed at current time. For instance, for men, the probability of employment increases 21.70 percentage points when the individuals are employed at the previous period. For women the probability of employment increases 32.78 percentage points when the individual is employed at previous period. This may indicate that the state dependency of employment for women is higher compared to men.

Capital Income has significant and negative effect on probability of employment for both men and women. The probability of employment for men decreases 1.49

percentage points when they have capital income. Women also experience 1.66 percentage point decrease in the probability of employment when they have capital income.

The effect of marital status on probability of employment is different among men and women. For instance, being married has significant and positive effect on probability of employment for men. In other words, for men, the probability of employment increases 1.75 percentage points when the individual is married. While women experience 5.26 percentage points decrease in probability of employment when they are married.

It seems that higher level of education has positive larger effects on probability of being employed for both men and women. For instance, individuals who are post-secondary graduated are more likely to be employed compared to those who are secondary school graduated or has some post-secondary. In fact, the probability of employment for men increases 3.23 percentage points when they are post- secondary graduated. Women also experience 8.80 percentage point increase in probability of being employed.

The effect of being immigrant on probability of employment is not the same among men and women. For instance, for men, being immigrant increase the probability of employment 0.51 percentage points while women experience 0.76 percentage points decrease on probability of being employed.

Number of children (less than 12 years old) has significant and negative effect on probability of employment for both men and women. However, women experience a much larger negative effect compared to men (-4.33 vs. -0.60).

The effect of age on probability of employment is negative and significant for both men and women. In other words, for men, the probability of being employed decreases 3.72 percentage points and women also experience 4.88 decrease in probability of employment due to one-year increase in age. This may indicate that as age increase to a certain point the effect becomes negative and increasingly as age continues to increase. Thus, the probability of employment is concave in age effect.

Table 3.5: Marginal Effect of Employment Equation

	Men	Women
$Healthy_{t-1}$	0.0232***	0.0589***
	(0.00000)	(0.00000)
$Employed_{t-1}$	0.2170***	0.3278***
	(0.00001)	(0.00000)
Has Capital Income	-0.0149***	-0.0166***
	(0.00000)	(0.00000)
Lives in BC	0.0022***	0.0151***
	(0.00000)	(0.00000)
Lives in Ontario	0.0066***	0.0258***
	(0.00000)	(0.00000)
Lives in East	-0.0137***	-0.0213***
	(0.00000)	(0.0000)
Lives in Prairies	0.0212***	0.0227***
	(0.00000)	(0.00000)
Married	0.0175***	-0.0526***
	(0.00000)	(0.00000)
Owns a house	-0.0013***	-0.0037***
	(0.00000)	(0.00000)
Secondary school graduated	0.0221***	0.0437***
	(0.00000)	(0.00000)
Has some post-secondary	0.0176***	0.0520***
	(0.00000)	(0.00000)
Post-secondary graduated	0.0323***	0.0880***
	(0.00000)	(0.00001)
Immigrant	0.0051***	-0.0076***
	(0.00000)	(0.00000)
Household size is equal to 2	0.0017***	-0.0041***
	(0.00000)	(0.00000)
Household size is equal to 3	0.0078***	0.0029***
	(0.00000)	(0.00000)
Household size is equal to 4	0.0148***	0.0068***
	(0.00000)	(0.00000)
Household size is equal or greater than5	0.0042***	0.0213***
	(0.00000)	(0.00001)
Number of Children	-0.0060***	-0.0433***
	(0.00000)	(0.00000)
Age	-0.0372***	-0.0488***
	(0.00000)	(0.00000)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

3.6. Sensitivity Analysis

For robustness check, the dynamic model is estimated without taking into account endogeneity of health and employment variables (Appendix 3.A, Table 3.A1 and 3.A2). In other words, by estimating each equation separately, the coefficients for health and employment variables are biased and inconsistent. Particularly, the magnitude of the estimated coefficients on lagged endogenous variables is exaggerated for men and women. For men, in the health equation, the coefficient of the lagged employment variable is 0.3520 which is positive and higher than the one in the simultaneous dynamic estimation (-0.1549). In the employment equation, the coefficient of the lagged health variable is 0.5233 which is much higher than its counterpart in the simultaneous dynamic estimation (0.0113). For women, the coefficient of the lagged employment variable is 0.2768 in the health equation, which is positive and higher than the one in the simultaneous dynamic estimation (-0.1650). In the employment equation, the coefficient of the lagged health variable is 0.4213 which is higher than its counterpart in the simultaneous dynamic estimation (0.1800). The persistency of health and employment variables is also overestimated. For instance, in the employment equation, the lagged employment variable is higher than its counterpart in the simultaneous dynamic estimation for men (1.5520 vs. 1.2109) and women (1.5294 vs. 1.2325). In the health equation, the lagged health variable is also higher than the one in the simultaneous dynamic estimation for men (1.2646 vs. 0.6271) and women (1.3147 vs. 0.5883). These confirm Haan and Myck (2009) finding that a separate specification leads to upward biased effects of labour market and health variables and suggest a joint modelling through a simultaneous equation model to account for unobserved characteristics which might jointly affect both health and employment.

Another robustness check is to show the importance of controlling for unobserved heterogeneity which is ignored by cross-sectional studies. Thus, a pooled dynamic bivariate model is estimated (the estimation is provided in Appendix 3.A, Table 3.A3). This also leads to biased and inconsistent coefficients. Similar to the case where endogeneity of health has not been considered, the estimated coefficients for the endogenous variables are overestimated. For instance, for men, in the health equation, the estimated coefficient for the lagged employment variable is positive and higher than its counterpart in the dynamic simultaneous equation model where individual unobserved heterogeneity has been taken into account (0.3417 vs. -0.1549). In the employment equation, the estimated coefficient for the lagged health variable is also higher than the one in the dynamic simultaneous equation model for men (0.0162 vs. 0.0113). For women, in the health equation, the estimated coefficient for the lagged employment variable is higher compared to the one in the dynamic simultaneous equation model (-0.5963 vs. -0.1650). However, in the employment equation, the estimated coefficient of the lagged health variable is lower than its counterpart in the dynamic simultaneous equation model, it is not significant (0.0028 vs. 0.1800).

Persistency of the health and employment variables are also overestimated by the pooled bivariate dynamic model. For instance, for men, in the health equation, the estimated coefficient for the lagged health variable is higher than the one in the dynamic simultaneous equation model (1.2541 vs. 0.6271). In the employment equation, men also have higher estimated coefficient for the lagged employment variable (1.5312 vs. 1.2109). For women, in the health equation, the estimated coefficient for the lagged health variable is higher than the one in the dynamic simultaneous equation model

(1.0527 vs. 0.5883). In the employment equation, women also have higher estimated coefficient for the lagged employment variable (1.5620 vs. 1.2325).

These results confirm the finding by Canadian studies such as Au et al (2005) and Hum, Simpson and Fissuh (2008) that suggest ignoring unobserved heterogeneity can overestimate the effect of health on labour market outcomes.

The study also uses excellent health instead of healthy (excellent, very good and good) as a sensitivity analysis, results show the persistency of health and employment variables. Moreover, health is endogenous for both men and women[8]. However, it seems that using excellent health status instead of healthy status (excellent, very good or good) as a health measure in the estimations, gives a different relationship between health and employment. For instance, for men, results show that the latent employment has a negative and significant effect on the latent excellent health (-0.1053). However, the lag of employment has a positive and significant effect on the probability of excellent health (0.2642). This may indicate that employment has a positive effect on men with excellent health with lag not simultaneously. On the other hand, in the employment equation, results show that the latent excellent health has a positive and significant effect on the latent employment (0.2101). However, the lag of excellent health has a negative and significant effect on the probability of employment (-0.0263). This may show that, for men, having excellent health may have a negative effect on employment with lag. For women, results demonstrate that the latent employment has a positive and significant effect on the latent excellent health (0.1863). While, the lag of employment has a

[8] Please refer to the estimation on the appendix 3.C.

negative and significant effect on the probability of excellent health (-0.3086). This may indicate that in contrast to men, employment has a negative effect on women with excellent health with lag not simultaneously. On the other hand, in the employment equation, results indicate that the latent excellent health has a positive effect on the latent employment (0.0883). However, the lag of excellent health has a negative and significant effect on the probability of employment (-0.1173). This may show that similar to men, having excellent health may have a negative effect on employment with lag for women.

Age is an important factor influencing both health and employment variables. In addition, as noted by Lindeboom and van Doorslaer (2004), self-assessed health is not homogenous among different age groups. Thus, another robustness check is to use different sub-sample groups such as 25-34, 35-49, 50-64 and investigate whether health and employment effects are different among different age groups over time.[9]

Table 3.6 shows the results of the dynamic model for the individuals between the age of 25 and 34 for men and women. Results show a positive effect of probability of employment (latent employment) on the probability of health (latent health) for men and women (0.6831 vs. 0.1943). However, this effect is not significant for women. Compared to the dynamic estimation for the full sample (25-64), the effect is higher and less significant for men (0.6831 vs. 0.2257). Moreover, the lag of employment has a negative and significant effect on the probability of health for men and it is much higher than its counterpart in the estimation for the full sample. For women, the lag of employment has a negative effect on the probability of health, but it is not significant. On the other hand, the

[9] In the estimated results for the sub-samples, [iweight] has not been used due to error problems.

effect of latent health on the latent employment is positive and significant and they are higher than those in the full sample estimation for men (0.5240 vs. 0.3246) and women (0.2070 vs. 0.1274). The lag of health has a negative and insignificant effect on the probability of employment for both men and women.

As already mentioned, the parameter atanhrho shows the correlation between error terms in the bivariate panel probit model. Since the value for this parameter is not significant for both men and women, thus, the health variable is not endogenous for the individuals between the age of 25 and 34. As a result, univariate probit models have to be used for each equation (health and employment equations). Table 3.A4 and 3.A5 on Appendix 3.A show the estimations for health and employment equations respectively. In the health equation, results show that the lag of employment has significant and positive effect on the probability of health for women. However, the effect of lag of employment on the probability of health is positive and insignificant for men. In the employment equation, the lag of health has significant and positive effect on the probability of employment for men and women.

Therefore, in contrast to the dynamic model for the full sample age group (25-64), the health variable is not endogenous for the individuals between the age of 25 and 34. In other words, health and employment equations have not to be estimated simultaneously and they are not correlated through their error terms. Compared to the individuals between the age of 25 and 64 that have negative and significant effect of the lag of employment on the probability of health for both men and women, this effect is positive and significant for women but not significant for men between the age of 25 and 34. On the other hand, the positive and significant effect of lag of health on the probability of

employment for the individuals between the age of 25 and 34 is much higher than its counterparts for men (0.5185 vs. 0.0113) and women (0.3693 vs. 0.1800) between the age 25 and 64.

Table 3.6: Dynamic Bivariate Panel Probit Estimation- Age (25-34)

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	0.5512	1.3181***
	(0.44437)	(0.30340)
$Healthy_{t-1}$	0.5942**	0.5305***
	(0.28278)	(0.19478)
$Employed_t$	0.6831**	0.1943
	(0.33162)	(0.26676)
$Employed_{t-1}$	-0.6974*	-0.1236
	(0.37690)	(0.33463)
Has Capital Income	-0.0170	0.0592
	(0.17792)	(0.17774)
Lives in BC	-0.4339*	-0.2579
	(0.2521)	(0.17466)
Lives in Ontario	-0.5048***	-0.2406*
	(0.1688)	(0.14262)
Lives in East	-0.0058	-0.0508
	(0.21231)	(0.14183)
Lives in Prairies	-0.3683**	-0.1752
	(0.15629)	(0.13958)
Married	-0.1317	-0.3062
	(0.25638)	(0.19249)
Owns a house	-0.1472	0.2243
	(0.16144)	(0.15059)
Secondary school graduated	-0.1270	0.5109**
	(0.22951)	(0.23002)
Has some post-secondary	-0.0967	0.3315
	(0.19656)	(0.20287)
Post-secondary graduated	0.0006	0.4921*
	(0.28061)	(0.27177)
Immigrant	0.0678	-0.1944
	(0.17587)	(0.14581)
Household size is equal to 2	0.2818*	0.1738
	(0.15977)	(0.1323)
Household size is equal to 3	0.1743	0.0267
	(0.13310)	(0.14136)
Household size is equal to 4	0.0496	-0.1153
	(0.1527)	(0.15481)
Household size is equal or greater than 5	-0.0619	0.1900
	(0.25340)	(0.21713)
Number of Children	0.0556	0.3573***
	(0.1200)	(0.10871)
Age	2.2338	6.9984
	(6.0472)	(5.44793)

Age2	-0.3936	-1.2201
	(0.99135)	(0.88875)
Constants	-30.0479*	27.9945*
	(15.43524)	(16.9157)
Employment Equation		
Initial Employed	0.2192	0.5038***
	(0.16511)	(0.10297)
$Employed_{t-1}$	0.8920***	0.7260***
	(0.14738)	(0.08932)
$Healthy_t$	0.5240**	0.2070*
	(0.25637)	(0.11746)
$Healthy_{t-1}$	-0.2549	-0.0014
	(0.3153)	(0.23277)
Has Capital Income	0.0006	-0.0009
	(0.15084)	(0.10756)
Lives in BC	0.1291	0.0710
	(0.17872)	(0.10989)
Lives in Ontario	0.3501**	0.1351
	(0.13851)	(0.08995)
Lives in East	-0.1944	-0.1140
	(0.1281)	(0.08013)
Lives in Prairies	0.3382***	0.1109
	(0.12523)	(0.08471)
Married	0.3021	-0.0994
	(0.19781)	(0.12142)
Owns a house	0.1080	-0.0912
	(0.1343)	(0.09821)
Secondary school graduated	0.3139**	0.2803**
	(0.14558)	(0.12846)
Has some post-secondary	0.2580**	0.3218***
	(0.12122)	(0.11104)
Post-secondary graduated	0.3190**	0.5705***
	(0.14907)	(0.12995)
Immigrant	-0.1819	-0.0657
	(0.13394)	(0.09638)
Household size is equal to 2	-0.1228	-0.0480
	(0.11931)	(0.09479)
Household size is equal to 3	-0.1434	-0.1994**
	(0.11200)	(0.09066)
Household size is equal to 4	-0.0857	-0.0246
	(0.12993)	(0.09736)
Household size is equal or greater than 5	-0.1133	-0.1381
	(0.20091)	(0.12142)
Number of Children	-0.0113	-0.3496***
	(0.1011)	(0.06447)

Age	-1.6889	-0.7511
	(5.1331)	(3.4677)
Age2	0.3079	0.1932
	(0.84217)	(0.56726)
Constants	24.1886*	-12.1533
	(13.47145)	(10.80984)
atanhrho	-1.2585	-0.3487
	(0.92404)	(0.40259)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. $Employed_{t-1}$ refers to the employment status from the previous period. Initial Employed is the employment status at first cycle. $Healthy_{t-1}$ refers to the health status from the previous period. Initial Healthy is the health status at first cycle. Model includes time average of all time -varying control variables. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1

Table 3.7 presents the results of the dynamic estimations for men and women between the age of 35 and 49. Results indicate that the latent employment has significant and positive effect on the latent health for men and women. Compared to the full sample model estimation, this effect is not significantly different for men (0.2256 vs. 0.2257) and women (0.2538 vs. 0.2207). Furthermore, the lag of employment has not any significant effect on the probability of health for both men and women between the age of 35 and 49. On the other hand, the latent health has positive and significant effect on the latent employment for both men and women. However, women have higher value for the coefficient of the latent health than their counterparts for the full sample size estimation (0.2213 vs. 0.1274). The lag of health has not any significant effect on the probability of employment for both men and women between the age of 35 and 49.

The value for the parameter atanhrho is not significant for men which indicates that the health variable is not endogenous and univariate probit models have to be estimated for each equation (health and employment equations) for men.

Appendix 3.A, table 3.A6 and 3.A7 show the estimation for health and employment equations separately for men. In the health equation, results indicate that the lag of

employment has significant and positive effect on the probability of health for men. On the other hand, in the employment equation, the lag of health has significant and positive effect on the probability of employment for men.

In summary, it can be concluded that health is endogenous for women but exogenous for men between the age of 35 and 49. However, health is endogenous in the estimations for the individuals between the age of 25 and 64. In contrast to the estimation for the individuals between the age of 25 and 64 that have negative and significant effect of the lag of employment on the probability of health for both men and women, this effect is positive and significant for men but negative and insignificant for women between the age of 35 and 49. On the other hand, however, the effect of lag of health on the probability of employment was positive and significant for both men and women between the age of 25 and 64, this effect is only significant for men between the age of 35 and 49 and it is much higher than its counterpart in the full sample estimation (0.7015 vs. 0.0113).

Table 3.7: Dynamic Bivariate Panel Probit Estimation- Age (35-49)

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	1.1949***	1.4830***
	(0.16089)	(0.13236)
$Healthy_{t-1}$	0.5415***	0.4666***
	(0.12005)	(0.08979)
$Employed_t$	0.2256**	0.2538***
	(0.09807)	(0.0729)
$Employed_{t-1}$	0.1160	-0.1936
	(0.19128)	(0.14497)
Has Capital Income	0.0344	0.4335***
	(0.11230)	(0.10456)
Lives in BC	-0.3774**	-0.4339***
	(0.15009)	(0.13107)
Lives in Ontario	-0.2832**	-0.3956***
	(0.11895)	(0.10726)
Lives in East	-0.0711	-0.2253**
	(0.11574)	(0.10372)
Lives in Prairies	-0.2571**	-0.2997***
	(0.12443)	(0.10929)
Married	-0.1789	0.03449
	(0.18252)	(0.14077)
Owns a house	-0.1095	0.0774
	(0.14981)	(0.11832)
Secondary school graduated	0.2276*	0.2155*
	(0.13539)	(0.11827)
Has some post-secondary	0.1425	0.1545
	(0.10952)	(0.10488)
Post-secondary graduated	0.2852**	0.2916***
	(0.11529)	(0.11164)
Immigrant	-0.0341	-0.0665
	(0.11663)	(0.10054)
Household size is equal to 2	-0.1007	-0.0363
	(0.10585)	(0.09401)
Household size is equal to 3	-0.0143	-0.1810*
	(0.11389)	(0.09451)
Household size is equal to 4	0.0427	-0.0590
	(0.11795)	(0.10312)
Household size is equal or greater than 5	-0.1631	-0.0340
	(0.13877)	(0.12718)
Number of Children	0.0142	0.1069*
	(0.06122)	(0.06122)
Age	-1.2152	2.9477
	(2.2823)	(1.88099)
Age2	0.0983	-0.3889*
	(0.26413)	(0.21727)
Constants	-3.6079	2.535031
	(10.14775)	(9.4709)

Employment Equation		
Initial Employed	1.1504***	1.0520***
	(0.14987)	(0.09101)
$Employed_{t-1}$	0.7295***	0.9604***
	(0.11203)	(0.06519)
$Healthy_t$	0.3169***	0.2213***
	(0.09200)	(0.05604)
$Healthy_{t-1}$	0.2712	0.1162
	(0.17233)	(0.11273)
Has Capital Income	-0.1751	-0.0531
	(0.10814)	(0.0782)
Lives in BC	0.2404	0.0319
	(0.15031)	(0.10606)
Lives in Ontario	0.2044*	0.1786**
	(0.11093)	(0.08584)
Lives in East	-0.2081**	-0.0954
	(0.10227)	(0.07975)
Lives in Prairies	0.3448***	0.1328
	(0.11583)	(0.08471)
Married	0.3354**	-0.0226
	(0.17055)	(0.12232)
Owns a house	-0.0996	-0.1579
	(0.138279)	(0.10110)
Secondary school graduated	0.3135**	0.3409***
	(0.12670)	(0.09391)
Has some post-secondary	0.2592**	0.3515***
	(0.10153)	(0.08370)
Post-secondary graduated	0.4094***	0.5325***
	(0.10555)	(0.08671)
Immigrant	0.0242	-0.0276
	(0.11724)	(0.08151)
Household size is equal to 2	0.1115	-0.0055
	(0.09578)	(0.08030)
Household size is equal to 3	0.1610	0.0904
	(0.10340)	(0.07958)
Household size is equal to 4	0.2308**	0.0926
	(0.10816)	(0.08098)
Household size is equal or greater than 5	0.3062**	0.1118
	(0.13791)	(0.09562)
Number of Children	0.0069	-0.1450***
	(0.05965)	(0.04693)
Age	2.2765	0.5393
	(2.18921)	(1.52474)
Age2	-0.2644	-0.0398
	(0.25368)	(0.17699)
Constants	-20.6824**	-22.8403***
	(9.42124)	(7.494024)
atanhrho	-0.2503	-0.2907**
	(0.17210)	(0.12053)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. $Employed_{t-1}$ refers to the employment status from the previous period. Initial Employed is the employment status at first cycle.

$Healthy_{t-1}$ refers to the health status from the previous period. Initial Healthy is the health status at first cycle. Model includes time average of all time -varying control variables. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3.8 provides the results of the dynamic estimations for men and women between the age of 50 and 64. Results demonstrate that the latent employment has a significant and positive effect on the latent health for both men and women. In comparison to the model estimation for full sample size (25-64), the coefficient for the latent employment is higher than those in the full sample estimation especially for men (0.2587 vs. 0.2257) and women (0.3484 vs. 0.2207). The effect of the lag of employment on the probability of heath is significant and negative for women and it is higher than its counterpart for the full sample size estimation (-0.3865 vs. -0.1650). However, the lag of employment does not have any effect on the probability of health for men. On the other hand, the effect of the latent health on the latent employment is significant and positive for men but it is not significant for women. Compared to the full sample model estimation, this effect is smaller for men (0.1921 vs. 0.3246). The lag of health does not have any significant effect on the probability of employment for both men and women between the age of 50 and 64.

The value for the parameter atanhrho is not significant for women. That means that health is not endogenous for women between the age of 50 and 64. Thus, univariate probit models have to be estimated for health and employment equations.

Appendix 3.A, table 3.A8 and 3.A9 provide the estimations for health and employment equations separately for women between the age 50 and 64. In the health equation, results demonstrate that the lag of employment has significant and positive effect on the

probability of health. On the other hand, the lag of health has significant and positive effect on the probability of employment.

In brief, for the individuals between the age of 50 and 64, health is endogenous for men but exogenous for women. While health is endogenous for both men and women between the age of 25 and 64. Moreover, although the lag of employment has significant and negative effect on the probability of health for both men and women between the age of 25 and 64, this effect is positive and only significant for women between the age 50 and 64. On the other hand, the lag of health has positive and significant effect on the probability of employment only for women and this effect is higher than its counterpart for women between the age of 25 and 64.

Table 3.8: Dynamic Bivariate Panel Probit Estimation- Age (50-64)

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	1.2670***	1.1911***
	(0.13831)	(0.13005)
$Healthy_{t-1}$	0.5949***	0.7402***
	(0.09733)	(0.09908)
$Employed_t$	0.2587**	0.3484***
	(0.10432)	(0.09322)
$Employed_{t-1}$	-0.2505	-0.3865*
	(0.23477)	(0.21644)
Has Capital Income	0.0294	0.0626
	(0.08366)	(0.08308)
Lives in BC	-0.1867	-0.2336*
	(0.13449)	(0.12498)
Lives in Ontario	-0.2492**	-0.1212
	(0.10695)	(0.09903)
Lives in East	-0.3844***	-0.1309
	(0.10247)	(0.09487)
Lives in Prairies	-0.2902***	-0.3846***
	(0.10970)	(0.10103)
Married	0.3463	0.0828
	(0.22099)	(0.19639)
Owns a house	-0.0073	-0.2936*
	(0.16957)	(0.17258)
Secondary school graduated	0.1984*	0.1733
	(0.11440)	(0.10619)
Has some post-secondary	0.1916**	0.0852
	(0.09075)	(0.08608)
Post-secondary graduated	0.1855**	0.2182**
	(0.08702)	(0.09015)
Immigrant	-0.1209	-0.2074**
	(0.09869)	(0.09034)
Household size is equal to 2	0.0214	0.0573
	(0.10348)	(0.08639)
Household size is equal to 3	-0.0321	0.0642
	(0.12414)	(0.10892)
Household size is equal to 4	-0.0641	-0.1384
	(0.14739)	(0.14794)
Household size is equal or greater than5	-0.3742*	0.0398
	(0.206846)	(0.23267)
Number of Children	0.1102	0.0830
	(0.14471)	(0.15360)
Age	-1.0471	2.4321
	(2.64950)	(2.53190)
Age2	0.0782	-0.1861
	(0.22751)	(0.21799)
Constants	28.79829	26.6680
	(17.57069)	(16.29275)

Employment Equation		
Initial Employed	0.9422***	0.8867***
	(0.13022)	(0.11158)
$Employed_{t-1}$	1.3898***	1.4516***
	(0.08462)	(0.07219)
$Healthy_t$	0.1921***	0.0872
	(0.07358)	(0.07490)
$Healthy_{t-1}$	0.1087	0.1972
	(0.14266)	(0.14934)
Has Capital Income	-0.1370**	-0.1376**
	(0.06650)	(0.06187)
Lives in BC	0.0884	0.2533**
	(0.10816)	(0.10219)
Lives in Ontario	0.0731	0.1776**
	(0.08847)	(0.08040)
Lives in East	0.0497	0.0522
	(0.08571)	(0.07747)
Lives in Prairies	0.2957***	0.3058***
	(0.08963)	(0.08472)
Married	-0.3259*	-0.0956
	(0.19032)	(0.1587)
Owns a house	0.2045	-0.1773
	(0.15305)	(0.14547)
Secondary school graduated	-0.0261	0.1665*
	(0.09753)	(0.08737)
Has some post-secondary	-0.0766	0.1546**
	(0.07883)	(0.07342)
Post-secondary graduated	0.0332	0.2179***
	(0.07493)	(0.07423)
Immigrant	0.2991***	0.0485
	(0.08243)	(0.07594)
Household size is equal to 2	0.0863	-0.0574
	(0.08969)	(0.07585)
Household size is equal to 3	0.2614**	0.0007
	(0.10434)	(0.09191)
Household size is equal to 4	0.3286**	0.1743
	(0.12803)	(0.12701)
Household size is equal or greater than 5	0.1826	0.0111
	(0.19357)	(0.19303)
Number of Children	-0.2322*	-0.1743
	(0.13417)	(0.14802)
Age	-1.0298	-2.6553
	(2.34988)	(2.0932)
Age2	-0.0008	0.1425
	(0.20255)	(0.18092)
Constants	3.7874	-10.7371
	(15.55798)	(14.75304)
atanhrho	-0.3012*	-0.2481922
	(0.17194)	(0.16162)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. $Employed_{t-1}$ refers to the employment status from the previous period. Initial Employed is the employment status at first cycle. $Healthy_{t-1}$ refers to the health status from the previous

period. Initial Healthy is the health status at first cycle. Model includes time average of all time -varying control variables. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

3.7. Discussion and Conclusion

Canadian studies that have investigated the impact of health on labour market outcomes are cross-sectional or static (Breslaw and Stelener, 1986; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have not examined the dynamic interaction between health and labour market outcomes. Moreover, they have solved the endogeneity of health through a two-stage method or instrumental variable (IV) approach which does not give an efficient estimate for health effect. This study contributes to the Canadian literature by estimating a dynamic panel data model through a simultaneous equation model to obtain more precise model specification for the interaction of health and labour market outcome.

For this purpose, the study uses the 9-cycle of longitudinal data of National Population Health Survey (NPHS), on a sample of men and women between the age of 25 and 64.

The results confirm that there is a high state dependency in employment variable for both men and women. In the health equation, the coefficient for $healthy_{t-1}$, is positive and significant while it is much lower compared to the coefficient of $Employed_{t-1}$ in employment equation. This may indicate that there is state dependency for health variable for both men and women while it is not as strong as that of employment variable. This confirms the earlier work by Hann and Myck (2009) that use German Data for men aged 30-59 and find strong persistence in the dynamics of health and labour market risks. Specifically, even by controlling the lag of employment, the latent health has still significant positive effect on the latent employment for both groups in employment equation. In addition, compared to the static model the coefficient for the latent health are

smaller especially for women which indicate that the static model overestimates the health effect for both groups.

On the other hand, by controlling the lag of health, the latent employment also has a significant positive effect on the latent health for both men and women in health equation. However, this time the coefficient for the latent employment is greater especially for women compared to the static model.

In the dynamic bivariate probit model, the lag of employment and the lag of health variable in the health and employment equations are also controlled respectively. Results show that in the employment equation, being healthy at previous period increases the probability of employment for both men and women; however, this effect is much higher for women than men. On the other hand, in the health equation, being employed at previous period has a negative and significant effect on the probability of health for both men and women.

For robustness check, the dynamic model is estimated without taking into account endogeneity of health and employment variables. In other words, by estimating each equation separately, the coefficients for health and employment variables are biased and inconsistent. Particularly, the magnitude of the estimated coefficients on the lagged endogenous variables is exaggerated and the persistence of health and employment is also overestimated. This confirms Haan and Myck (2009) finding that a separate specification leads to upward biased effects of labour market and health variables and suggest a joint modelling through a simultaneous equation model to account for unobserved characteristics which might jointly affect both health and employment.

Another robustness check is to show the importance of controlling for unobserved heterogeneity which is ignored by cross-sectional studies. Thus, a pooled dynamic bivariate model is estimated. This also leads to biased and inconsistent coefficients. Similar to the case where endogeneity of health has not been considered, the estimated coefficient for the endogenous variables and persistency of health and employment variables are overestimated. This confirms the finding by Canadian studies such as Au et al (2005) and Hum, Simpson and Fissuh (2008) that suggest ignoring unobserved heterogeneity can overestimate the effect of health on labour market outcomes.

Age is an important factor influencing both health and employment variables. In addition, as noted by Lindeboom and van Doorslaer (2004), self-assessed health is not homogenous among different age groups. Thus, another robustness check is to use different sub-sample groups such as 25-34, 35-49, 50-64 and investigate whether health and employment effects are different among different age groups over time. Results demonstrate that the interactions between health and employment is different for different age groups. Particularly, health is only endogenous for women 35 to 49 and men 50 to 64 while it is endogenous for both men and women between the age 25 and 64. Moreover, the effect of lag of employment on the probability of health is positive and significant for most age groups (this effect is negative for women 35 to 49 and men 50 to 64 but they are not significant). While, this effect is negative and significant for both men and women between the age of 25 and 64. On the other hand, the effect of lag of health on the probability of employment is positive and significant for most age groups (this effect is positive for women 35 to 49 and men 50 to 64 but they are not significant). In addition,

this effect is higher for all age groups compared to the individuals between the age of 25 and 64.

The purpose of this study is to provide more precise estimation for the interaction between health and employment; however, we acknowledge some limitations and suggest more avenues for future research.

In the current study, self-perceived health is used as a health measure in the estimation. However, there maybe be justification bias arising from this health measure. Particularly, individuals who are not employed may report poor health to justify their employment status. Thus, using this health measure may overestimate the effect of health on employment. The evidence on the justification hypothesis is mixed. Some studies find that self-reported health suffers from justification bias (Bazzoli, 1985; Kreider, 1999, Kerkhofs et al., 1999; Lindeboom and Kerkhofs, 2002; Au et al, 2005). On the other hands, some others do not find any strong evidence for this bias (Boaz and Muller, 1990; Dwyer and Mitchell, 1999; Benitez-Silva et al, 2004). Studies that find the justification endogeneity for self-reported health, have tried to use more objective health measures such as Health Utility Index[10] or specific health conditions as instruments (Bound et al., 1999; Disney et al., 2006; Au et al. 2005). Thus, one possibility is to use more objective measure of health to solve the justification endogeneity of self-reported health. However, the smaller effect of health on labour supply using more objective measure of health can be due to measurement error (Bound, 1989). Baker, Stabile and Deri (2004) also confirm

[10] Please refer to the appendix 3.D that shows that subjective and objective measure of health coincides.

that even more objective self-reported measures of health may suffer from measurement error and justification bias.

We could also use alternative measures of labour force outcome such as labour force participation, work for pay or profit, total hours worked, part-time/full-time status, job search or volunteer work and compare the results with the original estimate (in which employment status is used as a measure for labour supply outcome). However, if we want to consider more objective measure of health along with detailed information on income and earnings, there is a need for a new longitudinal survey to investigate the relationship between health and labour supply. Since longitudinal National Population Health Survey (NPHS) and longitudinal Survey of Labour and Income Dynamics (SLID) only have detailed information on health or income and earnings respectively.

As it was mentioned earlier in the study, one alternative approach is to use a two-stage method that also gives consistent estimate for the equation parameters, but it is not efficient since it does not consider the correlation of error terms between the two equations (health and employment). There are two instrumental variables approaches (IV) to correct for endogeneity bias in non-linear models; two stage predictor substitution (2SPS) and two-stage residual inclusion (2SRI) (Terza, Basu and Rathouz, 2008). In the 2SPS approach, the endogenous variable is regressed on the related control variables in the first stage then the non-linear predicted values of the endogenous variable is obtained and in the second stage the endogenous variables is replaced with its predicted value (Terza, Basu and Rathouz, 2008). In this regard it is important to use non-linear fitted value of the endogenous (binary) variable as an instrument from the first stage into the second stage. Since, using linear fitted value of the endogenous variable results in

forbidden regression (Angrist and Pischke, 2009). The SRI approach is similar to 2SPS in the first stage; however, the endogenous variable is not replaced, and the first stage residuals is used as an additional regressor in the second-stage equation (Terza, Basu and Rathouz, 2008). Cai (2010) also uses a two-stage method and compares the results with a simultaneous equation model. He mentions that both methods give consistent parameter estimates but the two-stage method is inefficient. He indicates that in the two-stage method, the correlation between the two equations (employment and health equations) is not taken into account. Moreover, a true test for the exogeneity of the health variable cannot be conducted by the two-stage method. Consequently, he suggests that relying only on the two-stage method, may give incorrect inferences on the endogeneity of health variable. Thus, using the two-stage approaches and comparing the results with the simultaneous equation model in which correlation of error terms is taken into account can be a good extension as a future research.

Controlling for state dependency along with endogeneity of the health and employment variables help reach a more precise estimation which would be more useful for policy makers to conduct efficient policies. Since the results confirm that there is a positive and significant interaction between health and employment, thus, health policies that have positive and direct effect on health, can have positive and indirect effect on employment status.

In this regard, policy makers can invest on workplace health programs. For instance, investing on wellness culture can be a good way of boosting employees' health status. In other words, a wellness culture is a consistent determinant of personal health and job satisfaction (The Sanofi Canada Healthcare Survey, 2020). Among the ways in

which the work environment encourages health and wellness; safety, good relationships with co workers and a good relationship with their immediate supervisor can be more effective ways that workplace environment can encourage health and wellness (The Sanofi Canada Healthcare Survey, 2020).

In addition, workplace polices through business plans and policy manuals can support individual's health. In other words, they can encourage more physical activity (e.g. transit passes and fitness centre memberships), healthy eating (e.g. healthy snacks at meetings) and reduced alcohol consumption (e.g. education programs) (Workplace Wellness, n.d.).Workplaces can also have wellness committees to address more issues or polices that have a direct impact on health and wellness (Workplace Wellness, n.d.). Thus, workplace health programs can influence individual's health, such as health behaviors; health risks for disease; and current health status (Centers for Disease Control and Prevention, 2016). In addition, they can influence areas such as health care costs, absenteeism, productivity, recruitment/retention, culture and employee morale (Centers for Disease Control and Prevention, 2016).

A**ppendix 3.A**

Table 3. A1: Dynamic Panel Probit Model for Health Equation

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	0.5534***	0.6286***
	(0.04349)	(0.03539)
$Healthy_{t-1}$	1.2646***	1.3147***
	(0.03792)	(0.03272)
$Employed_{t-1}$	0.3520***	0.2768***
	(0.03429)	(0.02750)
Has Capital Income	0.0182	0.1193***
	(0.04559)	(0.04429)
Lives in BC	-0.2450***	-0.2453***
	(0.05460)	(0.04871)
Lives in Ontario	-0.1840***	-0.1693***
	(0.04408)	(0.03972)
Lives in East	-0.1431***	-0.1004***
	(0.04213)	(0.03841)
Lives in Prairies	-0.1504***	-0.1585***
	(0.0443)	(0.04023)
Married	0.0690	0.0368
	(0.07552)	(0.0616)
Owns a house	-0.0308	-0.0453
	(0.06381)	(0.05552)
Secondary school graduated	0.1852***	0.1461***
	(0.04810)	(0.04263)
Has some post-secondary	0.1095***	0.1213***
	(0.0384)	(0.0363)
Post-secondary graduated	0.1959***	0.2420***
	(0.03731)	(0.03613)
Immigrant	0.0053	-0.0637*
	(0.04236)	(0.0379)
Household size is equal to 2	0.0370	-0.0303
	(0.04297)	(0.03682)
Household size is equal to 3	0.0339	-0.0632
	(0.04844)	(0.04153)
Household size is equal to 4	0.0457	-0.0194
	(0.05265)	(0.04810)
Household size is equal or greater than 5	-0.1344**	0.0390
	(0.06805)	(0.06495)
Number of Children	0.0079	0.0500*
	(0.03047)	(0.0277)
Age	-0.3232	-0.3710**
	(0.21095)	(0.18950)
Age2	0.0068	0.0255
	(0.2194)	(0.01999)
Constants	-0.4198	1.0860***
	(0.35122)	(0.32248)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A2: Dynamic Panel Probit Model for Employment Equation

	Coefficients for Men	Coefficients for Women
Employment Equation		
Initial Employed	0.3452***	0.3397***
	(0.03453)	(0.02424)
$Employed_{t-1}$	1.5520***	1.5294***
	(0.03175)	(0.02342)
$Healthy_{t-1}$	0.5233**	0.4213***
	(0.03906)	(0.03278)
Has Capital Income	-0.1082***	-0.0352
	(0.03920)	(0.03297)
Lives in BC	0.0410	0.0491
	(0.04931)	(0.03994)
Lives in Ontario	0.0872**	0.0837***
	(0.03867)	(0.03193)
Lives in East	-0.0862**	-0.0652**
	(0.03562)	(0.02992)
Lives in Prairies	0.2283***	0.1061***
	(0.03972)	(0.03183)
Married	0.0939	-0.1385***
	(0.06910)	(0.05053)
Owns a house	-0.0116	-0.0843*
	(0.05617)	(0.04589)
Secondary school graduated	0.1625***	0.2047***
	(0.04386)	(0.03587)
Has some post-secondary	0.1052***	0.2113***
	(0.03558)	(0.03164)
Post-secondary graduated	0.2356***	0.3476***
	(0.03415)	(0.03077)
Immigrant	0.0658*	-0.0182
	(0.03984)	(0.0317)
Household size is equal to 2	0.0442	-0.0356
	(0.0376)	(0.03221)
Household size is equal to 3	0.1285***	-0.0329
	(0.04293)	(0.03491)
Household size is equal to 4	0.1851***	0.0306
	(0.04785)	(0.03799)
Household size is equal or greater than 5	0.0833	0.0588
	(0.06545)	(0.04781)
Number of Children	-0.0505*	-0.1281***
	(0.02840)	(0.02025)
Age	1.1088***	1.4066***
	(0.1830)	(0.14364)
Age2	-0.1544***	-0.1778***
	(0.01901)	(0.01541)
Constants	-3.5869***	-4.2832***
	(0.30851)	(0.2554)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A3: Pooled Dynamic Bivariate Probit Model

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	0.5952***	0.5650***
	(000116)	(0.00136)
$Healthy_{t-1}$	1.2541***	1.0527***
	(0.00107)	(0.0018)
$Employed_t$	0.0189***	0.4832***
	(0.00384)	(0.00222)
$Employed_{t-1}$	0.3417***	-0.5963***
	(0.00652)	(0.00417)
Has Capital Income	-0.0046***	0.1462***
	(0.0012)	(0.00117)
Lives in BC	-0.1939***	-0.2488***
	(0.00123)	(0.0011)
Lives in Ontario	-0.1562***	-0.2371***
	(0.00102)	(0.00092)
Lives in East	-0.1353***	-0.0961***
	(0.00135)	(0.00132)
Lives in Prairies	-0.1300***	-0.2570***
	(0.00150)	(0.00109)
Married	-0.0249***	0.1322***
	(0.00197)	(0.00176)
Owns a house	-0.1680***	0.0532***
	(0.00153)	(0.00155)
Secondary school graduated	0.1797***	0.0530***
	(0.00147)	(0.00126)
Has some post-secondary	0.1366***	0.0565***
	(0.00113)	(0.00120)
Post-secondary graduated	0.1817***	0.0655***
	(0.00138)	(0.00142)
Immigrant	-0.0476***	-0.1228***
	(0. 00090)	(0.00086)
Household size is equal to 2	0.1246***	0.0279***
	(0.00110)	(0.00105)
Household size is equal to 3	0.0616***	0.0147***
	(0.00122)	(0.00113)
Household size is equal to 4	0.0861***	-0.0223***
	(0.00143)	(0.00124)
Household size is equal or greater than 5	-0.0700***	0.0068***
	(0.00155)	(0.00160)
Number of Children	-0.0039***	0.0736***
	(0.00076)	(0.00079)
Age	-0.4831***	-0.9287***
	(0.00668)	(0.00586)
Age2	0.0260***	0.0964***
	(0.00079)	(0.00065)
Constants	-0.6518***	3.6537***
	(0. 01716)	(0.0130)

Employment Equation		
Initial Employed	0.2717***	0.2821***
	(0.00097)	(0.00073)
$Employed_{t-1}$	1.5312***	1.5620***
	(0.00147)	(0.00076)
$Healthy_t$	0.3729***	0.2646***
	(0.00224)	(0.00168)
$Healthy_{t-1}$	0.0162***	0.0028
	(0.00352)	(0.00264)
Has Capital Income	-0.1433***	-0.0930***
	(0.00103)	(0.00091)
Lives in BC	0.1002***	0.1276***
	(0.00123)	(0.00102)
Lives in Ontario	0.1367***	0.1450***
	(0.00092)	(0.00080)
Lives in East	-0.0655***	-0.0535***
	(0.00129)	(0.00106)
Lives in Prairies	0.2772***	0.1479***
	(0.00111)	(0.00095)
Married	0.1657***	-0.2250***
	(0.00188)	(0.00143)
Owns a house	0.0400***	-0.0257***
	(0.00152)	(0.00129)
Secondary school graduated	0.1790***	0.1482***
	(0.00132)	(0.00101)
Has some post-secondary	0.1134***	0.1604***
	(0.00105)	(0.00094)
Post-secondary graduated	0.2286***	0.2757***
	(0.00107)	(0.00098)
Immigrant	0.0604***	0.0207***
	(0.00091)	(0.00077)
Household size is equal to 2	-0.0150***	-0.0163***
	(0.00110)	(0.00092)
Household size is equal to 3	0.0754***	0.00113***
	(0.00115)	(0.00097)
Household size is equal to 4	0.1703***	0.0518***
	(0.00125)	(0.00102)
Household size is equal or greater than 5	0.0935***	0.1006***
	(0.00155)	(0.00127)
Number of Children	-0.0679***	-0.1531***
	(0.00071)	(0.00057)
Age	1.1842***	1.4553***
	(0.00511)	(0.00409)
Age2	-0.1603***	-0.1818***
	(0.00052)	(0.00043)
Constants	-3.8952***	-5.2222***
	(0.00846)	(0.00738)
atanhrho	-0.0966***	-0.5423***
	(0.00455)	(0.00357)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A4: Dynamic Panel Probit for Health Equation- Age (25-34)

	Coefficients for Men	Coefficients for Women
Health Equation		
Initial Healthy	0.5383***	0.6585***
	(0.15898)	(.11394)
$Healthy_{t-1}$	1.1143***	1.1199***
	(0.14757)	(0.10791)
$Employed_{t-1}$	0.1732	0.1859**
	(0.10886)	(.07352)
Has Capital Income	-0.0304	0.0349
	(0.16294)	(0.1484)
Lives in BC	-0.5202***	-0.2133*
	(0.15714)	(0.12753)
Lives in Ontario	-0.4129***	-0.1720
	(0.13627)	(0.10516)
Lives in East	-0.2011	-0.0704
	(0.13633)	(0.10082)
Lives in Prairies	-0.2357*	-0.1138
	(0.13889)	(0.10292)
Married	0.1075	-0.2840*
	(0.19309)	(0.14969)
Owns a house	-0.1323	0.1856
	(0.14844)	(0.12636)
Secondary school graduated	0.1129	0.4081***
	(0.14096)	(0.12758)
Has some post-secondary	0.1059	0.2985***
	(0.1202)	(0.10556)
Post-secondary graduated	0.3087**	0.4541***
	(0.12100)	(0.10559)
Immigrant	-0.0527	-0.1476
	(0.13641)	(0.10484)
Household size is equal to 2	0.3464***	0.1248
	(0.11618)	(0.10309)
Household size is equal to 3	0.1206	-0.0235
	(0.11961)	(0.10481)
Household size is equal to 4	0.0079	-0.0956
	(0.13335)	(0.12164)
Household size is equal or greater than 5	-0.2274	0.1282
	(0.19180)	(0.17279)
Number of Children	0.0999	0.2545***
	(0.10530)	(0.07790)
Age	1.4574	3.5064
	(5.5377)	(4.6163)
Age2	-0.2643	-0.6332
	(0.90759)	(0.752860)
Constants	-21.8308	21.8736*
	(13.94203)	(13.0369)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A5: Dynamic Panel Probit for Employment Equation- Age (25-34)

	Coefficients for Men	Coefficients for Women
Employment Equation		
Initial Employed	0.2103**	0.3051***
	(0.09217)	(0.05743)
$Employed_{t-1}$	0.9957***	0.9537***
	(0.09082)	(0.05689)
$Healthy_{t-1}$	0.5185***	0.3693***
	(0.12078)	(0.08846)
Has Capital Income	-0.0337	0.0175
	(0.12682)	(0.09513)
Lives in BC	-0.1386	0.0068
	(0.11750)	(0.08762)
Lives in Ontario	0.1337	0.0734
	(0.09899)	(0.0716)
Lives in East	-0.2953***	-0.1224*
	(0.08600)	(0.06430)
Lives in Prairies	0.2148**	0.0575
	(.09927)	(0.06794)
Married	0.3664**	-0.1737*
	(0.16793)	(0.10236)
Owns a house	0.0435	-0.0287
	(0.11232)	(0.08571)
Secondary school graduated	0.3542***	0.3369***
	(0.11300)	(0.08874)
Has some post-secondary	0.2845***	0.3536***
	(0.09250)	(0.07913)
Post-secondary graduated	0.4480***	0.6111***
	(0.09078)	(0.07797)
Immigrant	-0.2061*	-0.1045
	(0.11029)	(0.07728)
Household size is equal to 2	0.0495	-0.0126
	(0.08377)	(0.08064)
Household size is equal to 3	-0.0901	-0.1708**
	(0.09271)	(0.07623)
Household size is equal to 4	-0.0973	-0.0398
	(0.1112)	(0.08107)
Household size is equal or greater than 5	-0.2249	-0.0689
	(0.16763)	(0.10077)
Number of Children	0.0363	-0.2514***
	(0.08543)	(0.05054)
Age	-0.7741	-0.1981
	(4.36347)	(3.06817)
Age2	0.1398	0.0735
	(0.71568)	(0.50105)
Constants	14.4130	-5.9723
	(11.14594)	(8.97089)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A6: Dynamic Panel Probit For Health Equation- Men (35-49)

	Coefficients for Men
Health Equation	
Initial Healthy	0.4841***
	(0.08226)
$Healthy_{t-1}$	1.2423***
	(0.07187)
$Employed_{t-1}$	0.5209***
	(0.0639)
Has Capital Income	0.0146
	(0.0867)
Lives in BC	-0.2161**
	(0.09414)
Lives in Ontario	-0.1654**
	(0.07412)
Lives in East	-0.0711
	(0.07206)
Lives in Prairies	-0.1136
	(0.07614)
Married	-0.0856
	(0.14354)
Owns a house	-0.1185
	(0.12025)
Secondary school graduated	0.2585***
	(0.08409)
Has some post-secondary	0.1319**
	(0.0668)
Post-secondary graduated	0.2706***
	(0.06519)
Immigrant	-0.0313
	(0.07458)
Household size is equal to 2	-0.0911
	(0.07436)
Household size is equal to 3	0.0121
	(0.07939)
Household size is equal to 4	0.0401
	(0.07936)
Household size is equal or greater than 5	-0.0984
	(0.09494)
Number of Children	-0.0147
	(0.04878)
Age	-0.9144
	(1.82112)
Age2	0.0710
	(0.21061)
Constants	-3.8804
	(6.97793)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A7: Dynamic Panel Probit for Employment Equation- Men (35-49)

	Coefficients for Men
Employment Equation	
Initial Employed	0.4635***
	(0.06584)
$Employed_{t-1}$	1.3486***
	(0.06390)
$Healthy_{t-1}$	0.7015***
	(0.07074)
Has Capital Income	-0.1142
	(0.08361)
Lives in BC	0.1133
	(0.09522)
Lives in Ontario	0.1060
	(0.06877)
Lives in East	-0.1611**
	(0.06235)
Lives in Prairies	0.2306***
	(0.07338)
Married	0.2268*
	(0.13427)
Owns a house	-0.1384
	(0.1088)
Secondary school graduated	0.3097***
	(0.07819)
Has some post-secondary	0.2174***
	(0.06304)
Post-secondary graduated	0.3609***
	(0.06106)
Immigrant	-0.0008
	(0.07601)
Household size is equal to 2	0.0462
	(0.06671)
Household size is equal to 3	0.1180
	(0.07169)
Household size is equal to 4	0.2006***
	(0.07362)
Household size is equal or greater than 5	0.2020**
	(0.09478)
Number of Children	-0.0078
	(0.04713)
Age	1.1016
	(1.7329)
Age2	-0.1434
	(0.2008)
Constants	-14.4747**
	(6.52926)

Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A8: Dynamic Panel Probit For Health Equation- Women (50-64)

	Coefficients for Women
Health Equation	
Initial Healthy	0.6086***
	(0.05972)
$Healthy_{t-1}$	1.3109***
	(0.05652)
$Employed_{t-1}$	0.3322***
	(0.04524)
Has Capital Income	0.0399
	(0.06951)
Lives in BC	-0.1243
	(0.08371)
Lives in Ontario	-0.0326
	(0.06696)
Lives in East	-0.0865
	(0.06391)
Lives in Prairies	-0.2063***
	(0.06667)
Married	0.0148
	(0.16229)
Owns a house	-0.3279**
	(0.14239)
Secondary school graduated	0.1628**
	(0.06979)
Has some post-secondary	0.1090*
	(0.05636)
Post-secondary graduated	0.2214***
	(0.05679)
Immigrant	-0.1479**
	(0.06123)
Household size is equal to 2	0.0268
	(0.06423)
Household size is equal to 3	0.0576
	(0.08173)
Household size is equal to 4	-0.0832
	(0.11040)
Household size is equal or greater than 5	-0.0464
	(0.17763)
Number of Children	0.0563
	(0.13050)
Age	1.2014
	(2.13725)
Age2	-0.1026
	(0.18445)
Constants	18.0986
	(12.02888)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Table 3. A9: Dynamic Panel Probit for Employment Equation- Women (50-64)

	Coefficients for Women
Employment Equation	
Initial Employed	0.3677***
	(0.05042)
$Employed_{t-1}$	1.7790***
	(0.04610)
$Healthy_{t-1}$	0.3761***
	(0.05624)
Has Capital Income	-0.1038*
	(0.05533)
Lives in BC	0.1824**
	(0.07270)
Lives in Ontario	0.1155**
	(0.05731)
Lives in East	0.0254
	(0.05524)
Lives in Prairies	0.2031***
	(0.05801)
Married	-0.0857
	(0.14038)
Owns a house	-0.1774
	(0.12662)
Secondary school graduated	0.1397**
	(0.0624)
Has some post-secondary	0.1328**
	(0.05257)
Post-secondary graduated	0.1719***
	(0.05049)
Immigrant	0.0262
	(0.05394)
Household size is equal to 2	-0.0154
	(0.05959)
Household size is equal to 3	0.0194
	(0.07270)
Household size is equal to 4	0.1678*
	(0.10161)
Household size is equal or greater than 5	0.0262
	(0.15643)
Number of Children	-0.1423
	(0.13030)
Age	-3.0216
	(1.88401)
Age2	0.2047
	(0.16267)
Constants	-8.0315
	(11.56512)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not healthy, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Appendix 3.B

Table 3.B: Variables Definitions

Variable Name	Definition of variables
Employed	= 1 if employed
Healthy	=1 if the individual health status is excellent, very good and good
Has capital income	=1 has capital income
Lives in BC	= if lives in BC
Lives in Ontario	=1 if lives in Ontario
Lives in Quebec	=1 if lives in Quebec
Lives in East	=1 if lives in NF, PEI, NB, NS
Lives in Prairies	=1 if lives in ALB, MB or SASK
Married	= 1 if married
Owns a house	= 1 if owns a house
Less than secondary school	=1 has less than secondary education
Secondary school graduated	= 1 if is secondary school graduated
Has some post-secondary	=1 if has some post-secondary education
Post-secondary graduated	= if is post-secondary school graduated
Immigrant	=1if immigrant
Household size is equal to 1	=1if household size =1
Household size is equal to 2	=1if household size =2
Household size is equal to 3	=1if household size =3
Household size is equal to 4	=1if household size =4
Household size is equal or greater than 5	=1if household size =>5
Number of Children	Number of children
Age	Age/10
Age^2	$Age^2/100$

A**ppendix 3.C**

Table 3.C: Dynamic Bivariate Panel Probit Estimation

	Coefficients for Men	**Coefficients for Women**
Health Equation		
Initial Excel	0.9588***	1.1519***
	(0.00002)	(0.00002)
$Excel_{t-1}$	0.3508***	0.3424***
	(0.00001)	(0.00002)
$Employed_t$	-0.1053***	0.1863***
	(0.00007)	(0.00003)
$Employed_{t-1}$	0.2642***	-0.3086***
	(0.00010)	(0.00005)
Has Capital Income	-0.0498***	0.0089***
	(0.00002)	(0.00002)
Lives in BC	-0.1193***	-0.1760***
	(0.00003)	(0.00003)
Lives in Ontario	-0.15199***	-0.1117***
	(0.00002)	(0.00002)
Lives in East	-0.2071***	-0.2185***
	(0.00006)	(0.00006)
Lives in Prairies	-0.1318***	-0.2459***
	(0.00003)	(0.00003)
Married	-0.0602***	0.0085***
	(0.00004)	(0.00003)
Owns a house	0.0030***	-0.0071***
	(0.00002)	(0.00003)
Secondary school graduated	0.1798***	0.1399***
	(0.00004)	(0.00003)
Has some post-secondary	0.1085***	0.0485***
	(0.00004)	(0.00003)
Post-secondary graduated	0.2304***	0.1239***
	(0.00004)	(0.00003)
Immigrant	0.0325***	-0.1418***
	(0.00002)	(0.00002)
Household size is equal to 2	-0.0011***	-0.1106***
	(0.00002)	(0.00002)
Household size is equal to 3	-0.0118***	-0.0668***
	(0.00002)	(0.00002)
Household size is equal to 4	0.0002***	-0.1105***
	(0.00003)	(0.00003)
Household size is equal or greater than 5	0.0517***	-0.1225***
	(0.00002)	(0.00003)
Number of Children	0.0015***	-0.0050***
	(0.00001)	(0.00001)
Age	-0.3090***	-0.5805***
	(0.00013)	(0.00010)
Age2	-0.0161***	0.0313***
	(0.00001)	(0.00001)
Constants	-1.3888***	-0.4373***

	(0.00030)	(0.00026)
Employment Equation		
Initial Employed	0.6847***	0.6920***
	(0.00003)	(0.00002)
$Employed_{t-1}$	1.2073***	1.2334***
	(0.00004)	(0.00002)
$Excel_t$	0.2101***	0.0883***
	(0.00006)	(0.00002)
$Excel_{t-1}$	-0.0263***	-0.1173***
	(0.00004)	(0.00002)
Has Capital Income	-0.1514***	-0.0796***
	(0.00003)	(0.00002)
Lives in BC	0.0539***	0.0724***
	(0.00003)	(0.00003)
Lives in Ontario	0.1157***	0.1172***
	(0.00002)	(0.00002)
Lives in East	-0.1247***	-0.0959***
	(0.00005)	(0.00005)
Lives in Prairies	0.3297***	0.1147***
	(0.00003)	(0.00003)
Married	0.2190***	-0.2617***
	(0.00004)	(0.00003)
Owns a house	-0.0208***	-0.0222***
	(0.00003)	(0.00003)
Secondary school graduated	0.3053***	0.2384***
	(0.00003)	(0.00003)
Has some post-secondary	0.2185***	0.2918***
	(0.00003)	(0.00003)
Post-secondary graduated	0.3545***	0.4494***
	(0.00003)	(0.00003)
Immigrant	0.0664***	-0.0225***
	(0.00002)	(0.00002)
Household size is equal to 2	0.0242***	-0.0182***
	(0.00002)	(0.00002)
Household size is equal to 3	0.0896***	0.0177***
	(0.00002)	(0.00002)
Household size is equal to 4	0.1887***	0.0390***
	(0.00003)	(0.00002)
Household size is equal or greater than 5	0.0639***	0.1136***
	(0.00003)	(0.00003)
Number of Children	-0.0723***	-0.2050***
	(0.0001)	(0.00001)
Age	1.7526***	2.0981***
	(0.00011)	(0.00009)
Age2	-0.2333***	-0.2591***
	(0.00001)	(0.00000)
Constants	-3.8933***	-5.0739***
	(0.00022)	(0.00017)
atanhrho	-0.05552***	-.2683***
	(0.00003)	(0.00003)

Note: Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%. Reference groups are not Excellent, no capital income, Quebec, not married, no owning a house, less than secondary school, not immigrant, household size equal to 1.

Appendix 3.D

Table 3.D: Health Utility index (HUI3) by Self-Perceived-Health (SPH)

Variable Name	HUI3
Healthy	0.9209
Not Healthy	0.6379

Note: Statistics are computed by a pooled sample of all 9 cycles of NPHS.

The above table shows that individuals who are healthy, have a higher rate of health utility index (HUI3) compared to those who are not healthy (0.92 vs. 0.64). Thus, using an objective measure of health (HUI3) coincides with the subjective measure of (SPH). Consequently, this indicates that HUI3 can be used as good alternative for SPH in health equation.

References

Angrist, J., & Pischke, J.-S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion.

Au, D. W. H., Crossley, T. F., & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health economics*, *14*(10), 999-1018.

Baker, M., Stabile, M., & Deri, C. (2004). What do self-reported, objective, measures of health measure?. *Journal of human Resources*, *39*(4), 1067-1093.

Bambra, C., & Eikemo, T. A. (2009). Welfare state regimes, unemployment and health: a comparative study of the relationship between unemployment and self-reported health in 23 European countries. *Journal of Epidemiology & Community Health*, *63*(2), 92-98.

Bazzoli, G. J. (1985). The early retirement decision: new empirical evidence on the influence of health. *Journal of human resources*, 214-234.

Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S., & Rust, J. (2004). How large is the bias in self-reported disability?. *Journal of Applied Econometrics*, *19*(6), 649-670.

Better Health. (2012). Work-related stress. Retrieved from https://www.betterhealth.vic.gov.au/health/healthyliving/work-related-stress

Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: a production function approach. *World development*, *32*(1), 1-13.

Boaz, R. F., & Muller, C. F. (1990). The validity of health limitations as a reason for deciding to retire. *Health Services Research*, *25*(2), 361.

Bound, J. (1989). *Self-reported vs. objective measures of health in retirement models* (No. w2997). National Bureau of Economic Research.

Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics*, *6*(2), 179-202.

Breslaw, J. A., & Stelcner, M. (1987). The effect of health on the labor force behavior of elderly men in Canada. *Journal of Human Resources*, 490-517.

Bryson, A., Forth, J., & Stokes, L. (2014). Does worker wellbeing affect workplace performance. *Department of Business Innovation and Skills, London*.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, *17*(1), 77-90.

Cai, L. (2018). State dependence of labour force participation of married Australian women. *International Journal of Manpower*.

Campolieti, M. (2002). Disability and the labor force participation of older men in Canada. *Labour economics*, *9*(3), 405-432.

Canadian Center for Occupational Health and Safety. (2018). Workplace Stress – General. Retrieved from https://www.ccohs.ca/oshanswers/psychosocial/stress.html

Case, A., Paxson, C., & Islam, M. (2009). Making sense of the labor market height premium: Evidence from the British Household Panel Survey. *Economics letters*, *102*(3), 174-176.

Centers for Disease Control and Prevention. (2016). Workplace Health Promotion. Retrieved from https://www.cdc.gov/workplacehealthpromotion/model/index.html

Clark, A. E., & Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, *104*(424), 648-659.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, *3*, 3309-3416.

Dabboussy, M., & Uppal, S. (2012). *Work absences in 2011*. Statistics Canada.

Damian, J., Ruigomez, A., Pastor, V., & Martin-Moreno, J. M. (1999). Determinants of self assessed health among Spanish older people living at home. *Journal of Epidemiology & Community Health*, *53*(7), 412-416.

Delattre, E., Moussa, R. K., & Sabatier, M. (2019). Health condition and job status interactions: econometric evidence of causality from a French longitudinal survey. *Health economics review*, *9*(1), 3.

Disney, R., Emmerson, C., & Wakefield, M. (2006). Ill health and retirement in Britain: A panel data-based analysis. *Journal of health economics*, *25*(4), 621-649.

Doyle, A. (2019). Types of Workplace discrimination. Retrieved from https://www.thebalancecareers.com/types-of-employment-discrimination-with-examples-2060914.

Dwyer, D. S., & Mitchell, O. S. (1999). Health problems as determinants of retirement: Are self-rated measures endogenous?. *Journal of health economics*, *18*(2), 173-193.

Feinstein, L., Sabates, R., Anderson, T. M., Sorhaindo, A., & Hammond, C. (2006, March). What are the effects of education on health. In *proceedings of the Copenhagen symposium" measuring the effects of education on health and civic engagement*.

Fischer, J. A., & Sousa-Poza, A. (2009). Does job satisfaction improve the health of workers? New evidence using panel data and objective measures of health. *Health economics*, *18*(1), 71-89.

Gan, L., & Gong, G. (2007). *Estimating interdependence between health and education in a dynamic model* (No. w12830). National Bureau of Economic Research.

Gannon, B. (2005). A dynamic analysis of disability and labour force participation in Ireland 1995–2000. *Health economics*, *14*(9), 925-938.

Goldberg, P., Guéguen, A., Schmaus, A., Nakache, J. P., & Goldberg, M. (2001). Longitudinal study of associations between perceived health status and self reported diseases in the French Gazel cohort. *Journal of Epidemiology & Community Health*, *55*(4), 233-238.

Goodman, N. (2015). The impact of employment on the health status and health care costs of working-age people with disabilities. *Lead Center Policy Brief*.

Graham, C., Higuera, L., & Lora, E. (2011). Which health conditions cause the most unhappiness?. *Health economics*, *20*(12), 1431-1447.

Guo, T., & Hu, L. (2011). Economic Determinants of Happiness. *arXiv preprint arXiv:1112.5802*.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of health economics*, *28*(6), 1116-1125.

Harris, R., Tobias, M., Jeffreys, M., Waldegrave, K., Karlsen, S., & Nazroo, J. (2006). Racism and health: The relationship between experience of racial discrimination and health in New Zealand. *Social science & medicine*, *63*(6), 1428-1441.

Hum, D., & Simpson, W. (1996). Canadians with disabilities and the labour market. *Canadian Public Policy/Analyse de Politiques*, 285-299.

Hum, D., Simpson, W., & Fissuh, E. (2008). The Impact of Health on Labour Supply over the Life Cycle. *Canada: Department of Economics, The University of Manitoba*.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.

Ippoliti, F., Corbosiero, P., Canitano, N., Massoni, F., Ricciardi, M. R., Ricci, L., ... & Ricci, S. (2017). Work-related Stress, over-nutrition and cognitive disability. *La Clinica Terapeutica*, *168*(1), 42-47.

Kalwij, A., & Vermeulen, F. (2008). Health and labour force participation of older people in Europe: what do objective health indicators add to the analysis?. *Health economics*, *17*(5), 619-638.

Kerkhofs, M., Lindeboom, M., & Theeuwes, J. (1999). Retirement, financial incentives and health. *Labour Economics*, *6*(2), 203-227.

Khullar, D., Chokshi, D.A. (2018) .*Health, Income, & Poverty: Where We Are & What Could Help.* Retrieved from *https://www.healthaffairs.org/do/10.1377/hpb20180817.901935/full/*

Kreider, B. (1999). Latent work disability and reporting bias. *Journal of Human Resources*, 734-769.

Larsson, D., Hemmingsson, T., Allebeck, P., & Lundberg, I. (2002). Self-rated health and mortality among young men: what is the relation and how may it be explained?. *Scandinavian journal of public health*, *30*(4), 259-266.

Lindeboom, M., Llena Nozal, A., & Van der Klaauw, B. (2006). Disability and work: the role of health shocks and childhood circumstances.

Lindeboom, M., & Kerkhofs, M. (2002). Health and Work of the Elderly.

Lindeboom, M., & Kerkhofs, M. (2009). Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, *24*(6), 1024-1046.

Lindeboom, M., & Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of health economics*, *23*(6), 1083-1099.

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2004). Unemployment alters the set point for life satisfaction. *Psychological science*, *15*(1), 8-13.

Lundborg, P. (2008). The Health Returns to Education-What Can We Learn from Twins?. *Available at SSRN 1113685*.

Martikainen, P., Adda, J., Ferrie, J. E., Smith, G. D., & Marmot, M. (2003). Effects of income and wealth on GHQ depression and poor self rated health in white collar women and men in the Whitehall II study. *Journal of Epidemiology & Community Health*, *57*(9), 718-723.

McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology*, *90*(1), 53.

Morneau Shepell Report. (2015). The true picture of workplace absenteeism. Retrieved from https://www.morneaushepell.com/sites/default/files/assets/permafiles/62090/true-picture-workplace-absenteeism.pdf.

Office of Disease Prevention and Health Promotion. (n.d.). Nutrition and Weight Status. https://www.healthypeople.gov/2020/topics-objectives/topic/nutrition-and-weight-status

Oguzoglu, U. (2012). Dynamics of disability and work in Canada.

Pelkowski, J. M., & Berger, M. C. (2004). The impact of health on employment, wages, and hours worked over the life cycle. *The Quarterly Review of Economics and Finance*, *44*(1), 102-121.

Pharr, J. R., Moonie, S., & Bungum, T. J. (2011). The impact of unemployment on mental and physical health, access to health care and health risk behaviors. *ISRN Public Health*, *2012*.

Public Health Agency of Canada. (2005-2008). Economic Burden of Illness in Canada, 2005-2008**.** Retrieved from https://www.canada.ca/en/public-health/services/reports-publications/economic-burden-illness-canada-2005-2008.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population studies*, *29*(2), 231-248.

Qin, L., Chen, C. P., Liu, X., Wang, C., & Jiang, Z. (2015). Health status and earnings of migrant workers from rural China. *China & World Economy*, *23*(2), 84-99.

Stern, S. (1989). Measuring the effect of disability on labor force participation. *Journal of human Resources*, 361-395.

Stewart, N. (2013). Missing in action: Absenteeism trends in Canadian organizations. Conference Board of Canada.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of health economics*, *27*(3), 531-543.

The Sanofi Canada Healthcare Survey (2020). Future Forward Frontline Perspectives on the Future of Health Benefit Plans. Retrieved from https://www.sanofi.ca/-/media/Project/One-Sanofi-Web/Websites/North-America/Sanofi CA/Home/en/Products-and-Resources/sanofi-canada-health-survey/sanofi-canada-healthcare-survey-2020 EN.pdf?la=en&hash=F1C763AA6B2F32C0BF2E623851FD05FD.

Thomas, K. C., & Ellis, A. R. (2013). Patterns of healthcare use and employment among people with disabilities. *Disability and Health Journal*, *6*(2), 133-140.

Waddell, G., & Burton, A. K. (2006). *Is work good for your health and well-being?*. The Stationery Office.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, *20*(1), 39-54.

Workplace Wellness. (n.d.) Health-related work policies. Retrieved from https://www.workplacewellnessonline.ca/pdfs/Health-related%20workplace%20policies.pdf

World Health Organization. (1958). *The first ten years of the World Health Organization*. World Health Organization.

Yelln, E. H., & Trupin, L. (2003). Disability and the characteristics of employment. *Monthly Lab. Rev.*, *126*, 20.

Chapter 4. Returns to Education in a Dynamic Approach in Canada

Abstract

This study investigates the impact of education on wages using a dynamic Mincer equation. This study uses the recent panel of SLID survey (Survey of Labour and Income Dynamics) from 2005 to 2010 and analyses men and women separately. The standard approach to estimate the return to education does not consider earning persistence. Using a dynamic Mincer equation, results show that earning persistence plays an important role in estimating the return to education. Particularly, a static Mincer equation, gives an upward biased estimates of return to education. Moreover, in contrast to a static Mincer equation where the return to education is constant, a dynamic Mincer equation gives a return that depends on labour market experience.

Keywords: education, wages, dynamic Mincer equation

4.1. Introduction

Individuals improve their knowledge and obtain required skills to increase career success through education. In other words, individuals who have higher degrees have higher chance to be employed. In this regard, Berger and Parkin 2009, demonstrates that post- secondary graduates are more likely to be employed and they earn more than those who do not continue their studies past high school. In addition, the authors show that the gap between unemployment rates of young Canadians with higher and lower levels of education has widened over the past 35 years (1971-2005). More educated individuals have more opportunities to find better jobs with higher incomes. In addition, education helps individuals to adopt themselves in case they face any sudden changes in their workplace conditions (Pat, 2016). Highly educated Canadians are more likely to have high incomes. In 2010, about one in four (24.1%) of individuals with university degree were in top 10% of income recipients compared to one in ten of those with a postsecondary certificate or diploma below bachelor's degree (Statistics Canada, 2016). Those individuals with higher level of education also experience healthier lifestyles with lower smoking and obesity rates. Millar and Wigle (1986) demonstrate that the prevalence of risk factors including cigarette smoking, overweight, obesity, elevated diastolic blood pressure, physical inactivity, excessive alcohol consumption, elevated cholesterol level, diabetes mellitus and the conjoint use of oral contraceptives and cigarettes tended to be higher among Canadian men and women between the age of 20 and 69 with low level of education. Those individuals with a higher degree can also enjoy better job security and satisfaction. Strumpel (1971) shows that more educated people are much less likely to be unemployed or to be on strike.

Higher education has significant contributions to society as well. Higher educated workers typically paying more tax and they are less likely to rely on social public assistance (Education Pays, 2016). Another advantage of higher levels of education is higher levels of civic participation. People with higher degree were much more likely to be civically engaged. According to the 2013 General Social Survey, 78% of those with a university degree between the age of 25 and 64 were members or participants of a group in comparison with 56% of those with high school diploma and 41% of those without a high school diploma (Turcotte , 2015). In addition, among individuals between the age of 25 and 34, those without a high school diploma have a voting rate of 44% compared to 81% for university graduates (Turcotte, 2015).

Canada has high completion rates in postsecondary education. In 2012, 57.3 per cent of Canadian aged 25-34 had either a college or university certificate which was the third highest among OECD members (Parliamentary Budget Officer, 2016). Tuition costs is increasing rapidly across most of the country. On average, undergraduate pay 40 percent more in tuition than they did 10 years ago (Statistics Canada, 2016). According to a study by Van Pelt and Emes in 2015, during the period 2001/2 to 2011/12, the overall nominal public spending in Canada on government schools increased from $38.9 billion to $59.6 billion. In addition, education spending per pupil in government schools also had a growth of 63.2 percentage points ($7,250 to $11,835) per pupil. In 2004-5, total federal spending in postsecondary education was about $8.7 billion and it increased steadily to more than $12.8 billion in 2010-11 (PBO, 2016). As individuals consider education as an investment for their future, thus they decide to continue attending colleges or universities if the benefits of higher education exceed of their tuition costs. Therefore, identifying the

return to education can help individuals on investment decision purposes. Furthermore, the return to education can be useful for policy makers who decide on education expenditures and prioritize schooling levels and finance schooling programs.

The human capital theory proposed by Becker (1964) has been widely used to analyze the wage differentials among individuals with differences in their human capital such as education and experience. According to this theory, education is as an investment of knowledge or skills which can increase productivity of individuals and result in higher incomes. There is an extensive literature on using the Mincer earning function as a framework for measuring return to education (Card, 1999; Heckman et al. 2003; Trostel, 2005; Lemieux, 2006; García-Suaza et al. 2014; Hartog and Gerritsen, 2016; Mazza et al. 2018). According to the standard Mincer equation proposed by Mincer (1974), the observed wage is explained by schooling years, potential labour market experience and potential labour-market experience squared. The Canadian studies that have investigated the return to education are static or cross-sectional (Murphy, Riddell & Romer; 1998; Burbidge, Magee & Robb, 2002; Lemieux, 2006; Oreopoulos, 2006; Boudarbat, Lemieux & Riddell, 2010; Bourbeau, Lefebvre & Merrigan, 2012; Lemieux, 2014; Ren and Shannon, 2017; Aydede and Dar, 2017) and they have not estimated the return to education through a dynamic approach. In other words, they have not taken into account earnings persistence in estimating the return to education. Moreover, they have used an instrumental variable approach (IV) to solve the endogeneity of education. This study contributes to the Canadian literature by estimating a dynamic Mincer model through a system GMM method to obtain more precise model specification for the return to education. In this regard, this study uses the longitudinal Survey of Labour and Income

Dynamics (SLID) and estimates the dynamic model for men and women between the age of 25 and 64 separately.

The rest of this study is organized as follows: Section 4.2 provides literature review of Canadian and international studies; section 4.3 introduces Mincer earning function and its dynamic revision; section 4.4 presents an overview of the data and offers descriptive statistics; section 4.5 explains the methodology used in the study; section 4.6 describes the empirical results; section 4.7 concludes and provides some avenue for further research.

4.2. Literature Review

4.2.1. Canadian Literature

The findings on the returns to education are not conclusive, as they have used different data sources, time periods, measures for earnings, method of estimation and variables to estimate the return to education. Some studies mostly have focused on the evolution of returns to education over time (Freeman and Needels, 1993; Murphy, Riddell and Romer, 1998; Burbidge, Magee and Robb, 2002; Bar-or et al. 1995; Boudarbat, Lemieux and Riddell, 2010; Bourbeau, Lefebvre and Merrigan, 2012). However, there are studies in which the effects of education on earnings has been investigated.

For instance, Lemieux (2006) mentions some adjustments to the Mincer equation in order to be an accurate benchmark for estimating wage determination equations. The author suggests using a quadratic function in potential experience instead of just a quadratic. He believes that using a quadratic potential experience in Mincer equation, understates wage growth for younger worker. Moreover, it leads to spurious decline in

wages among older workers. He also suggests including a quadratic term in years of schooling to capture the convexity in the relationship between schooling and wages. Furthermore, he suggests allowing for cohort effects to capture the dynamic growth in returns to schooling among cohorts born after 1950. In fact, the author believes that the basic Mincer human capital earnings model is an accurate model in a stable environment where educational achievement grows smoothly across cohorts as it was periods studied by Mincer (1974). Hence he concludes that in a less stable environment, major shifts in the relative supply of different age-education groups can make important changes in the structure of wages that have to be considered in a standard Mincer equation by adding cohort effects, or explicitly modeling the relative supply and demand for different groups of workers.

Oreopoulos (2006) uses data from the 33% sample of the 1971 Census and the 20% samples from the 1981, 1986, 1991, 1996, and 2001 Censuses to investigate the effects of compulsory schooling on educational attainment and socio-economic outcomes. The sample includes all individuals between the age of 20 and 65 who are born in a Canadian province and who are 14 years of age between 1920 and 1990. The author estimates the returns to compulsory schooling on earnings and other socio-economic outcomes by instrumenting individual grade attainment with the compulsory school laws and compares least squares (OLS) estimates of the returns to grade attainment with the instrumental variables estimates (IV). The author shows that results from OLS and IV estimates of returns to schooling are almost similar, 12.7% and 12.1% respectively. He mentions that when sample is restricted to men or when earnings are used instead of income in the regression, the IV coefficients decrease by more than a third which there is

no obvious explanation for this result. Moreover, the author shows that when changes in compulsory school requirements are weighted equally by province and birth cohort, returns to schooling based on income are about 11.3% for men and women while the effects on earnings are lower than those on income.

Lemieux (2014) investigates the return to education through several possible channels. The first is "pure returns" channel in which education makes workers more productive for a given task. The second indicates that education helps workers to get higher-paying jobs and the final channel is that workers earn more when their job is matched to their field of study. Using data from the 2005 National Graduate Survey and the 2006 Canadian Census, he finds that the second and third channels account for close to half of the conventional return to education while the first channel (the general productivity effect) accounts for the rest of half. Furthermore, he concludes that the return to education varies greatly depending on occupation, field of study and the math between these two factors.

Ren and Shannon (2017) use data from the 2011 National Household Survey (NHS) and the Canadian Census for 1981, 1986, 1991, 1996, 2001, and 2006 to estimate the private rate of return to a Canadian bachelor's degree that allows for differences in earnings determination, taxes, and tuition by birth cohort. Samples are Canadian born aged 18–64 who have either high school or a university bachelor's degree. The authors show that earnings equations have shifted significantly across cohorts in ways that have increased the returns to education for more recent cohorts especially for men. The authors conclude that rate-of-return estimates based on a single cross-sectional data set underestimate the rates of return of recent cohorts especially for men. In other words, the

authors believe that rates of return based on single cross-sectional data assume that earnings of those who are currently old are good proxies for the future earnings of those who are currently young, thus ignoring the increase in the educational earnings premium experienced by recent cohorts. The authors suggest that since allowing for cohort effects increases estimated returns to education, postsecondary education is a good investment for recent birth cohorts.

Aydede and Dar (2017) use the 20% sample of the 2006 Canadian Census to investigate the return to education for internationally educated immigrants after controlling for their occupational matching in hosting labor markets. The authors use two continuous indices that quantify the matching quality of the native-born in both fields of study and educational degrees. The authors believe that these indices allow them to separate the effects of immigrants' occupational attainment and their foreign schooling quality on wage earnings by measuring immigrants' occupational match relative to that of native-born. The authors show that, in general, well-matched native-born and Canadian-educated immigrants earn more the higher their educational attainment in terms of degrees. While the payoffs to schooling for well-matched foreign-educated workers are severely discounted. The authors mention that except for graduate degrees, the payoffs to all other degrees, including bachelor's degree, are effectively discounted to zero. In addition, they indicate that although foreign-educated immigrants generally have higher educational attainments, their matching quality is low, especially for postsecondary degree holders. The authors conclude that the discounted earnings of internationally educated immigrants are due to their non-equivalent quality of source-country human

capital and cannot be solved through a better matching of skills to occupations by postmigration improvements polices.

4.2.2. International Literature

Thera are also numerous studies which investigate the return to education. For instance, Angrist and Krueger (1991) use an individual's quarter of birth as an instrument for education. They show that due to compulsory schooling laws, men who are born from 1930 to 1959 and earlier in the year have less schooling than those born later in the year and the estimation results are typically higher than the corresponding OLS estimates. Using National Longitudinal Survey (NLS) for young men sample, and taking college proximity as an instrument for schooling, Card (1993) finds that the results from IV estimates much above the corresponding OLS estimator. He mentions that as accessibility is much more important for the individuals who are on the margin of continuing education, college proximity have a bigger effect for children with less educated parents. Thus, he suggests an alternative model that uses interactions of college proximity with family background variables as instruments for schooling as well as college proximity as a control variable. The author finds that the IV estimation from this alternative model is somewhat lower than estimation using only college proximity as an instrument for schooling, but it is still about 30 percent above the OLS estimation.

Card (1999) analyzes the literature on the causal relationship between education and earnings. He concludes that the studies of identical twins, show a small upward ability bias (on the order of 10%). However, studies which using instrumental variables based on interventions to school system are 20-40% above the corresponding OLS

estimates. Furthermore, he mentions that those IV estimates of return to schooling based on family backgrounds are systematically higher than corresponding OLS estimates.

Andini (2007) uses the US National Longitudinal Survey of Youth for the period of 1980 to 1987 to investigate the impact of education on within-groups wage inequality by quantile-regression techniques. He modifies Mincer equation and adds past earnings as an additional explanatory variable and compares the dynamic model with the standard Mincer equation. Particularly, he shows that the variability of earnings at each decile explained by the dynamic Mincer model varies from 25 to 39 percentage points, while the variability explained by the static Mincer model varies from 5 to 10 percentage points. In addition, he suggests that in the static Mincer model, education does not reduce earnings dispersion. In other words, the return to schooling slightly increases over the wage distribution while in the dynamic model education may reduce wage dispersion.

Andini (2013a) uses data from the US National Longitudinal Survey of Youth (NLSY) for the period of 1980–1987 to estimate a dynamic Mincer wage equation. The author mentions that the Mincer assumption on the equality of net potential earnings and observed earnings works within the perfect-competition framework where the nominal wage equals the monetary value of the marginal labour productivity. However, within the imperfect-competition framework, the assumption is unlikely to hold. In this regard, the author brings the existence of wage bargaining at worker–employer level as one of the possible arguments for the Mincer assumption. The author believes that real-life labour markets are characterized by wage bargaining and there is a possibility of a margin-formation between observed earnings and net potential earnings. That indicates that observed earnings may not coincide with net potential earnings at any time. He shows

that a dynamic Mincer equation is the solution of a simple wage-bargaining model between a worker and an employer and that the standard static Mincer equation is a particular case in which the bargaining power of the worker is full (where the worker is able to earn all his/her net potential earnings and the employer is indifferent between employing and not employing). He concludes that the dynamic wage model provides a good fit of NLSY data. He also shows that individual wages are explained well by the dynamic model even if the wages are negotiated collectively and in the public sector. Furthermore, he suggests that once individual unobserved heterogeneity is taken into account, controlling for individual observed characteristics other than wage history and human capital may not radically improve the model specification.

Andini (2013b) uses data from the European Community Household Panel (ECHP) for Belgian male workers between the age of 18 and 65 for the period of 1994-2001 to estimate the returns to schooling. Unlike the Mincer assumption of equality of observed and net potential earnings, Andini suggests that observed earnings adjusts to net potential earnings over time. Consequently, this leads to a dynamic Mincer equation in which past observed earnings play an important role in explaining the observed earnings as an additional explanatory variable. The author concludes that the dynamic return to schooling in terms of observed earnings is dependent on labour market experience while static return to schooling in terms of observed or net potential earnings is constant over time. He also shows that disregarding earning persistence in the static Miner model, leads to upward-biased estimate of the potential return. Furthermore, he suggests that disregarding earning persistence also leads to a severely upward-biased estimate of the observed return at the beginning of the working life.

Heckman et al. (2018) estimate a robust dynamic model of schooling and its causal consequences for earnings, health, and smoking. The authors indicate that their model recognizes the sequential dynamic nature of educational decisions. They categorize the benefits of schooling at different levels into direct and indirect components arising from continuation values. The authors estimate substantial continuation value components of graduating high school and completing college for high-ability individuals. They suggest that schooling has valuable options for future schooling of high-ability individuals. While for low-ability individuals, they find a substantial direct effect of graduating high school and little continuation value. The authors show a positive sorting into schooling based on gains, especially for higher schooling levels. They conclude that schooling has strong causal effects on earnings, health, and healthy behaviors. In addition, they show that both cognitive and non-cognitive endowments affect schooling choices and outcomes at each level of schooling.

Patrinos, Psacharopoulos and Tansel (2019) use Turkish 2017 Household Labor Force Survey (HLFS) to estimate private and social returns to investment in education. The authors find the average rate of return to schooling to be 8.8 percent. Moreover, they show that women have higher returns to schooling than men. The authors use the 1997 education reform to instrument schooling. They conclude that the reform is associated with a higher return for women than for men. The authors find that the average rate of return to schooling in the public sector is higher than that in the private sector (7.9% vs. 6.5%). In addition, they demonstrate that the private return to vocational secondary education is higher than general secondary education (6.5% vs. 5.7%).

4.3. Mincer Earning Function

4.3.1. Static Mincer Earning Function and its Dynamic Revision

The earning function model presented by Mincer (1974) is a practical example extracted from the human capital theory. According to the theoretical foundations of Mincer's earnings function (1974) reported by Heckman et al. (2003), this section shows how the standard Mincer function is formulated. In addition, the dynamic revision of the Mincer function is discussed as well.

Mincer (1974) mentions that potential earnings at time t (E_t) depend on human capital investment in the last period. In other words, If an individual invests k_t of his/her potential earnings in training with a return of τ_t, this can be written as follows:

$E_{t+1} = E_t (1+ k_t \tau_t)$ (1)

By repeated substitution, $E_t = \prod_{j=0}^{t-1}(1 + \tau_j k_j)E_0$. Taking the log on both sides gives the equation as below:

$\ln E_t = \ln E_0 + \sum_{j=0}^{t-1} \ln (1 + \tau_j k_j)$ (2)

In equation 2, formal schooling (s) indicates years spent in full- time investment which starts at the beginning of life. In addition, the rate of return to schooling ($r_0 = ... = r_{s-1} = \beta$) and the rate of return to post- schooling investment ($r_s = \dots = r_{t-1} = \tau$) in terms of potential earnings, all are constant over time. By these assumptions, equation (2) can take the following form:

$\ln E_t = \ln E_0 + s \ln(1+\beta) + \sum_{j=s}^{t-1} \ln (1 + \tau k_j)$ (3)

and for small β and τ :

$\ln E_t \approx \ln E_0 + \beta s + \tau \sum_{j=s}^{t-1} k_j$ (4)

For finding a relationship between potential earnings and labour market experience(x), Mincer (1974) assumes that post schooling investment linearly decreases over time:

$K_{s+x} = \theta\ (1- x/T)$ (5)

where $x = t-s \geq 0$ and T is length of working life.

Thus, by the assumption 5, the equation 4 can take the following format:

$\ln E_t \approx \ln E_0 - \theta\tau + \beta s + (\theta\tau + \theta\tau/2T)\ x - (\theta\tau/2T)\ x^2$ (6)

As observed earnings (wages) is equal to potential earnings minus investment costs, then the following equation can be written:

$\ln w_t \approx \ln E_t - \theta\ (1- x/T)$ (7)

by substituting (6) in (7), the standard Mincer equation can be found:

$\ln w_t \approx \alpha + \beta s + \lambda x + \gamma\ x^2$ or $\ln E_t - \theta\ (1- x/T) = \alpha + \beta s + \lambda x + \gamma\ x^2$ (8)

where $\alpha = \ln E_0 - \theta\tau - \theta$, $\lambda = \theta\tau + \theta\tau/2T + \theta/T$, $\gamma = -\theta\tau/2T$

Although Mincer earning function has been widely used as framework to estimate the return to education, it has been criticized by many authors as well. For instance, Murphy and Welch (1990) mention that the human capital earning function gives a very poor approximation of the true empirical relationship between earnings and experience. They conclude that using the standard Mincer function, understates early career earning growth about 30 to 50 percentage points and overstates midcareer growth by 20 to 50 percentage points.

Using 1940-1990 Census data, Heckman et al. (2003) examine the empirical support for Mincer's earning specification and the key implications of the Mincer model: (i) log- earnings experience profiles are parallel across schooling levels; (ii) log-earnings age profiles diverge with age across schooling levels; (iii) the variance of earnings has a

U-shaped pattern. The authors show that data from the 1940-1950 Censuses supports the two implications of Mincer model (i) and (ii), the evidence for implication (i) is weaker for 1960 and 1970 while data from 1980 and 1990 do not support the model. For the 1940 Census year, the variance of log-earnings over the life-cycle is relatively flat for whites. It is similarly flat in 1950 with increasing variance at the tails. However, data for black and white men from 1960-1990 Censuses have the U-shaped pattern predicted by Mincer model. In addition, they conclude that log earnings do not increase linearly with schooling, and experience profiles for log-earnings are not parallel across schooling types in recent decades. Particularly, they suggest a more general dynamic analysis of earnings function that takes into account tuition costs, income taxes, non-linearity in schooling, non-separability between experience and schooling and uncertainty in calculating returns to schooling.

Using comparable microdata from the International Social Survey Program (ISSP), Trostel (2005) estimates marginal rate of return to schooling in 12 countries for working age men. He takes a simple non-linear extension of the standard Mincer wage equation and uses an education polynomial in log-wage equation instead of a linear term. He finds a significant systematic nonlinearity in the marginal rate of return. In other words, he finds substantial increasing return at low levels of education (primary and secondary education) and substantial decreasing return in higher levels of education. He concludes that linear estimates of the rate of return remarkably understates the maximum marginal rate of return around 12 years of schooling and substantially overstate the rates of return at both the low and high levels of education.

Belzil (2008) finds that the effect of schooling and experience are not separable in the Mincer wage equation. In other words, after conditioning on skill heterogeneity, he finds a positive correlation between accumulated schooling and individual specific returns to experience. Hence, he concludes that standard Mincer equation that are based on the separability assumption[11], ignores the positive benefit of education on future wage growth and overestimates the return to high school education to significant margin (as much as 15 percentage points).

In the above-mentioned studies, the authors adjust the standard Mincer equation in order to find a better model specification for the earning function to fit the data well. However, neither of them criticizes the Mincer assumption for the equality of potential earnings and observed earnings. In this regard, some studies assume that the Mincer earning function holds but question the Mincer assumption for the equality of potential earnings and observed earnings to find a better model specification for the observed earning data (Andini 2007; 2010; 2013a; 2013b). In this regard, Andini (2013b) suggests that observed earnings adjusts to net potential earnings over time. The author takes Flannery and Rangan (2006) approach, and assumes that observed earnings dynamically adjusting to net potential earnings according to a simple adjustment model as follows:

$$\ln w_t - \ln w_{t-1} \approx \rho (\ln npe_{it} - \ln w_{t-1}) \qquad (1)$$

where $\rho \in [0, 1]$ is the average speed of adjustment.

[11] it refers to Mincer's model assumption of separability between the schooling and experience components of earnings which indicates that log earnings profiles are parallel in experience across schooling levels.

By substituting the static Mincer equation ($\ln w_{it} = \ln np_{it} + e_{it} = \alpha_i + \beta s_i + \lambda x_{it} + \gamma x^2_{it} + e_{it}$) in (1), he reaches a dynamic Mincer equation in which past observed earnings play an important role in explaining the observed earnings as an additional explanatory variable. Thus, he introduces the dynamic models as follows:

$$\ln w_{it} = \rho\alpha_i + (1-\rho) \ln w_{i\,t-1} + \rho\beta s_i + \rho\lambda x_{it} + \rho\gamma x_{it}^{2} + e_{it} \quad (2)$$

or

$$\ln w_{it} = \eta_1 \ln w_{i\,t-1} + \eta_2 s_i + \eta_3 x_{it} + \eta_4 x_{it}^{2} + \eta_i + e_{it}$$

where $\eta_1 = (1-\rho)$, $\eta_2 = \rho\beta$, $\eta_3 = \rho\lambda$, $\eta_4 = \rho\gamma$, $\eta_i = \rho\alpha_i$

This study uses the dynamic Mincer model extracted by Andini (2013b) and estimates that for the Canadian data. This study takes the recent panel of Survey of Labour and Income Dynamics (SLID); panel 5 (2005-2010) and estimates the model for men and women aged 25 to 64.

4.3.2. Static and Dynamic Schooling Returns

Regarding the returns to schooling, the static Mincer model ($\ln w_{it} = \alpha_i + \beta s_i + \lambda x_{it} + \gamma x_{it}^{2} + e_{it}$), shows that the returns to schooling based on the observed or net potential earnings is constant ($\frac{\partial \ln w_t}{\partial s} = \frac{\partial \ln w_{s+x}}{\partial s} \approx \beta$) and does not depend on potential work experience. On the other hands, since the static model estimation does not consider the persistency of the observed earnings, it overestimates the return to schooling compared to the dynamic model estimation. In the dynamic Mincer equation ($\ln w_{it} = \rho\alpha_i + (1-\rho) \ln w_{i\,t-1} + \rho\beta s_i + \rho\lambda x_{it} + \rho\gamma x_{it}^{2} + e_{it}$), the return to schooling based on the observed wages is $\beta(x) = \rho\beta \sum_0^x (1-\rho)^x$. [12] In contrast to the static Mincer model, the return to schooling based on the observed wages is not constant and it depends on potential experience (x).

[12]. See the Appendix 4.A for how the dynamic return to schooling is computed.

Furthermore, it is lower than the schooling returns based on the static Mincer function ($\beta(x) < \beta$). However, the static and dynamic return to schooling can be equal with unlimited potential work experience ($\rho\beta \left[\frac{1}{1-(1-\rho)} \right] \approx \beta$), it is not applicable. In other words, since individuals work experience can not be unlimited, the return to schooling based on the observed wage in the dynamic Mincer model, can not be equal to the return to schooling based on the observed wage in the static Mincer model.

4.4. Data and Variables

In this study the data is taken from the recent Survey of Labour and Income Dynamics (SLID); panel 5 (2005-2010). The SLID is a household Survey that considers all individuals in Canada; however, residents of Yukon, the Northwest Territories and Nunavut, residents of institutions and people living on Indian reserves or in military barracks are excluded (Statistics Canada, 2015). The SLID sample is composed of two panels and each panel includes about 17000 households and about 34000 adults which are interviewed every year up to six years (Statistics Canada, 2015). From 1993, the first panel of SLID is introduced then the second panel (1996-2001) is started in 1996 overlapping panel 1 (1993-1997) for a three-year period and this pattern repeats every three years which leads to panel 3 (1999-2004), panel 4 (2002-2007) and panel 5 (2005-2010) (Statistics Canada, 2015). This survey interviews individuals every year up to 6 years and includes information on a number of human capital variables, labor force experiences and demographic characteristics such as education, family relationships and household composition (Statistics Canada, 2015). In this regard, Dean (2010) mentions some advantage of using SLID over Census data. For instance, the SLID has more information than the Census on educational attainment. Unlike the Census, the SLID asks

respondents directly about their total accumulated work experience. Furthermore, its earning data is very accurate since more than eighty percent of respondents give permission to Statistics Canada to use income tax from the Canada Revenue Agency (CRA). Deans suggests that the absence of a significant measurement error in the earning data on SLID increases the precision of the estimates. Particularly, measurement error in earnings can be non-random and consequently bias the right-hand side variables when it is used as dependent variable in an earning regression (Bound et al., 1989). Thus, using earnings data from the SLID can solve this bias. Furthermore, as system GMM method is designed for panels with fixed T and large N (Roodman, 2009), the SLID would be appropriate for estimating the dynamic Mincer model.

This study is restricted to the individuals between the age of 25 and 64 and men and women is analyzed separately. The variables include logarithm of hourly wage, schooling years and potential work experience. The wage measure is the composite wage provided in the SLID and it is calculated based on the implicit hourly wages for all paid-worker jobs, weighted using total hours paid for each (SLID electronic data dictionary, 2012). For salary workers this measure is calculated using earnings and hours worked. However, for hourly wage earners the measure is what is directly reported by the respondent. For workers with multiple jobs the measure is a weighted average based on the hours worked in each job (Dean, 2010). Wages are adjusted for inflation using the provincial all-items CPI for base year 2002.

In the SLID, the variable for schooling years, is defined as numbers of years of schooling completed by person (elementary, high school, post- secondary). Since in the standard Mincer equation, it is assumed that individuals first stop schooling and then start

working, individuals with time-varying schooling years are dropped to make sure that they are not studying while working. Consequently, this keeps individuals with time-invariant schooling years in the sample. Potential experience is defined based on Mincer assumption as age - schooling years - 6.

Table 4.1 summarizes the demographic characteristics of the sample for men and women.

Table 4. 1: Summary Statistics[13]

	Men			Women		
	Obs.	**Mean**	**Std.dv.**	**Obs.**	**Mean**	**Std.dv.**
Logarithm of hourly wage	20877	3.049	0.459	20281	2.846	0.445
Schooling years	20877	13.389	2.587	20281	13.548	2.315
Potential work experience	20877	24.362	10.906	20281	24.391	10.903

In the sensitivity analysis section (7.1), extra control variables such as marital status, union membership, public sector, province of residence and occupations are used as well. The dummy variable union indicates whether the individual is a union member or covered by a collective agreement. The dummy variable public controls if the individual works in the public sector. Occupational dummies include Business, Finance and Administrative; Natural and Applied Sciences; Health; Social Science, Education, Government Service and Religion; Art, Culture, Recreation and Sport; Sales and Service; Trades, Transport and Equipment Operators; Unique to Primary Industry; Unique to Processing, Manufacturing and Utilities. The study uses Management Occupations as the base.

[13] In this chapter, the sample weight provided for the SLID is used for the summary statistics and all the estimations.

4.5. Methodology

4.5.1. Model Specification

A common way to estimate a dynamic model is using the least square estimator (OLS). However, this leads to biased estimates of return to education known as ability bias. In this regard, Gunderson and Orepoulos (2010) mention that because educated people can have other characteristics such as innate ability, motivation, organizational skills, that are associated with higher earnings and these characteristics are not controlled for in the estimation. Then omitting them from the estimating equation may be reflecting the effect of these factors and leads to biased estimates of returns to education. Moreover, the lag of dependent variable ($\ln w_{i\,t-1}$) is correlated with the individual unobserved heterogeneity (α_i). As a result, the OLS estimator leads to biased and inconsistent estimates.

Using fixed effects method also gives biased estimates of the coefficient of the lagged dependent variable. Nickell (1981) demonstrates the mechanism in which the within unit mean (the average of each variable within the cross-sectional unit) subtracted from each observation, creates a correlation between regressor and error. This bias is more serious when T is small even with a large number of individual units. In other words, he shows when $N\rightarrow\infty$, the inconsistency of the lag of dependent variable estimate ($\eta\acute{1}$) is of order 1/T, which can be so sizable when T is small. Random effect approach also creates the same problem. In fact, the unobserved individual heterogeneity (η_i) enters every value of $\ln w_{it}$ which results in the correlation of the lagged dependent variable with the composite error term.

One possibility is to eliminate the individual unobserved heterogeneity (η_i) by taking the first differences of the dynamic model. However, this leads correlation between the difference lagged dependent variable and the difference error term. Anderson and Hsiao (1981) use the second or third lag of the dependent variable in the form of differences or lagged levels as instruments for the lagged dependent variable as long as the error term is independent and identically distributed (i.i.d) and in a case in which the error term has AR(1) process, this method still works by using third or fourth lags of the dependent variable. Arellano and Bond (1991) find Anderson-Hsiao estimator consistent but not efficient as it does not use all of the sample information. In other words, in a approach called differenced generalized method of moments (GMM), Arellano and Bond consider all potential orthogonality conditions between the lag value of the dependent variable and the first difference of the error term. The differenced GMM approach has some disadvantages as well. For instance, the differenced GMM estimators are likely to perform poorly in case the dependent variable is highly persistent (Blundell and Bond, 1998). In fact, the lagged levels of the dependent variable are only weakly correlated with subsequent fist differences, hence that leads to weak instruments for the first differenced equations.

Another disadvantage of using the difference estimator is that the first difference transformation removes both the time-invariant variables and individual fixed effect. Since the key variable in this study (schooling years) is time-invariant, using the difference GMM estimator is not applicable. Arellano and Bover (1995) and Blundell and Bond (1998) add additional moment restrictions and allow lagged first differences to be used as instruments in the levels equations as well which corrects for any bias that would

result from using the difference estimator. In this approach, which is called System GMM, the difference model used by Difference-GMM estimator is combined with equation in levels in a system. Thus, the dynamic model can be written as below:

$$\ln w_{it} = \eta_1 \ln w_{i\,t-1} + \eta_2 s_i + \eta_3 x_{it} + \eta_4 x_{it}^2 + \eta_i + e_{it}$$

$$\ln w_{it} - \ln w_{i\,t-1} = \eta_1 (\ln w_{i\,t-1} - \ln w_{i\,t-2}) + \eta_3 (x_{it} - x_{i\,t-1}) + \eta_4 (x_{it}^2 - x_{i\,t-1}^2) + (e_{it} - e_{i\,t-1}) \quad (4)$$

$$\text{or } \Delta \ln w_{it} = \eta_1 \Delta \ln w_{i\,t-1} + \eta_3 \Delta x_{it} + \eta_4 \Delta \ln x^2_{\ it} + \Delta e_{it}$$

In the above model, the explanatory variables including lag of hourly wage ($\ln w_{i\,t-1}$), schooling years (s_i), potential experience (x_{it}) and potential experience squared (x_{it}^2), all are assumed endogenous and they are correlated with the individual unobserved heterogeneity (η_i). Following Andini (2013a), the instruments for the level equation is the lagged differences in the logarithm of the hourly wage ($\Delta \ln w_{i\,t-1}$). The instruments for the difference equation, are the lagged levels of all the time-varying explanatory variables ($\ln w_{i\,t-2}$, $x_{i\,t-2}$, $x_{i\,t-2}^2$). In the system GMM estimation, since the internal instruments are used in the estimation, it is necessary to check the validity of the instruments used. In this regard, Hansen J test and Arellano and Bond (AR) test are common tests for validity of instruments. Hansen overidentification test is known as a test of instrument validity, although; it can be used as a test of structural specification. In this regard, Roodman (2009) mentions that in case important explanatory variables are omitted from the model, they become part of the error term. As a result, this makes the error term being correlated with the instruments where they might not be in the correct model. Arellano and Bond (AR) test is used for autocorrelation of the idiosyncratic disturbance term. By construction, the residuals of the differenced equation have serial correlation. If the null

hypothesis of no second order serial correlation in the differenced residuals is not rejected, then the original error term is serially uncorrelated. However, if AR (2) is significant, second lags of endogenous variables are not valid instruments, since they are correlated with the error term. Thus, longer lags of the endogenous variables have to be used as instruments in the estimation (Roodman, 2009).

4.5.2. System GMM and Alternative Methods

Apart from Andini's work (2013a, 2013b) that estimate the dynamic model through system GMM approach to investigate the return to education, there are also other studies that use this method and alternative ones for the dynamic estimations but focus on other fields. For instance, Kohn and Averette (2014a) take the system GMM estimator to estimate the effect of relationship status on health. They use this method due to health dynamics, reverse causality from health to relationship status, unobservable heterogeneity and the lack of variation in relationship status. The authors do not find any difference in the effect on health between cohabitation and marriage for either men or women using the system GMM estimator while they show that marriage is better for health by FE estimators. The authors mention that since there is little variation in relationship status over time, FE models only identify individuals who change relationship status. Hence, the health experience of these group may not be generalizable to the majority of observations who stay in the same relationship. Kohn and Averette also compare the results on the relationship coefficients across different specifications (OLS, FE, FE with lag and System GMM) and find that controlling for lagged health decreases the size of the effect of relationship status on health for both men and women. Therefore, they suggest including lagged health in any analysis of health.

There are also studies that use a two-stage method to estimate the dynamic models. For instance, Kohn and Averett (2014b) measure the effect of relationship status on health. The model they use has two stages. In the first stage, they estimate a mixed logit model of relationship status and obtain a vector of the estimated unobserved heterogeneity then in the second stage, they control the estimated unobserved heterogeneity and estimate the dynamic health model by using generalized least squares (GLS) to account for heteroscedasticity and autocorrelation among individuals over time. The authors show that there is a strong persistency in health, and unobserved heterogeneity plays an important role in relationship choices. The authors do not find any positive effect of marriage on health relative to being never married while they find a positive heath effect of cohabitation relative to marriage for both men and women over 45 but not for younger ages. Furthermore, they do find a negative health effects for younger divorced men. Kripfganz and Schwarz (2019) use a two-stage approach to estimate a dynamic gravity model for US foreign direct investment (FDI). Particularly, in the first stage they estimate the coefficients of the time-varying regressors then in the second stage they regress the first stage residuals on the time-invariant regressors. Moreover, they adjust the second-stage standard errors for the first stage estimation error to avoid any misleading inference. The authors show history dependence of real bilateral stock of US outward FDI and conclude that disregarding the dynamic nature of the model leads to a sizable overestimation of the effect of the time-invariant geographical distance variable.

The two-stage methods may have some advantages. For instance, Kripfganz and Schwarz (2019) mention that a two-stage method can have partial insurance against

model misspecification. In other words, they indicate that the first stage estimates are unaffected by the choice of instruments used to identify the coefficients of time-invariant regressors in the second stage. The authors believe that the two-stage approach is flexible on the choice of the first stage estimator and it can offer a general model specification test. Furthermore, the authors mention that for the second stage, the overidentification restrictions for the time-invariant regressors can be tested by a conventional Hansen (1982) test. The two-stage methods may have some disadvantages ae well. For instance, Andini (2013a) suggests that in case there are more than one endogenous regressors, using two-stage approach needs to specify several first-stage models for the binary endogenous variables which makes the analysis too complicated. He also mentions that two- stage methods make the estimates strongly dependent of the way the first stage models are specified. Moreover, he believes that the first-stage models may not quantify all the unobserved heterogeneity.

4.6. Empirical Findings

In the empirical literature, the common way to estimate the return to education is the ordinary least square estimation method (OLS). However, OLS estimator gives a biased estimate for the return to education known as ability bias (Griliches, 1977). In other words, individuals with higher ability tend to acquire higher education. Thus, there is a correlation between unobserved individual effect and education that leads to the overestimation of the return to education by the OLS approach. Furthermore, the correlation of lagged dependent variable with the error term (Nickel's bias, 1981) is not taken into account by OLS estimation. In addition, the OLS estimation does not control for unobserved individual characteristics that can be correlated with the observed

independent variables. In this regard, as discussed earlier in the methodology section, using system GMM method help reach a more precise estimation, since it controls persistency of lagged dependant variable, endogeneity of independent variables (including time-invariant regressors) and fixed effects which captures the effect of person specific unobserved ability.

Table 4.2 shows the results using OLS and system GMM method to estimate the dynamic Mincer model. The table also contains estimations with year effects. Particularly for the system GMM method, the autocorrelation test and the robust estimates of the coefficient standard errors assume no correlation across individuals in the idiosyncratic disturbances and including time dummies makes this assumption more likely to hold (Roodman, 2009).

An F-test of joint significance of year dummies shows that year dummies are jointly significant for both men and women and year effects have to be included in the OLS estimations. Results using System GMM with year effects show that the hourly wage is highly persistent, and the coefficient of the lagged hourly wage is positive and significant for both men and women (0.8093 vs. 0.7710). This means that individual wages on the current period are highly dependent on individual wages at previous period. Other control variables such as schooling years, potential experience and potential experience square have the expected signs, but they are not significant for both men and women.

As it was discussed earlier in the methodology section, the least square estimator with year effects overestimates the coefficient for lagged dependant variable (the lagged of hourly wage) for both men (0.9021 vs. 0.8093) and women (0.9058 vs. 0.7710) compared to the system GMM with year effect. Moreover, the schooling coefficient is

underestimated by the OLS estimation for men (0.0087 vs. 0.0280) and women (0.0097 vs. 0.0626) respectively.

In this regard, the p-value for Hansen overidentification test confirms that instruments are exogenous and valid for both men and women respectively (0.1 vs. 0.420). The AR (2) is significant for men and women (0.010 vs. 0.001), which means that the second lags of the hourly wage is not a valid instrument since it is correlated with error term, thus longer lags of hourly wage is used as instrument in the estimation (lag 3 or longer). In addition, to avoid instrument proliferation, the study uses "collapse" option that is provided in system GMM formula to restrict the instruments used in the estimation.

The results from system GMM without controlling for year effects are not considered since the p-value for the Hansen test of overidentification shows that the instruments are not valid for men and women respectively (0.002 vs. 0.008).

Consequently, the results from the system GMM with year effects suggest that the lagged dependent variable (the lagged of hourly wage) have to be included in the Mincer function. This means that in contrast to the static Mincer model where potential earnings and observed earnings are equal simultaneously with the average speed of adjustment ($\rho=1$), in the dynamic Mincer model, observed earnings adjusting net potential earnings overtime with the adjustment ratio (ρ) equal to 0.1907 and 0.2290 for men and women respectively. The results also confirm the earlier work by Andini (2013a) and Andini (2013b) that find persistency of hourly wage and reject the static Mincer assumption.

Table 4.2: Dynamic Mincer Model for Men and Women

Dependent variable: logarithm of hourly wage	Men	Women
OLS without year effects		
Constant	0.2199***	0.1663***
	(0.01347)	(0.01280)
Logarithm of hourly wage (-1)	0.9017***	0.9061***
	(0.00380)	(0.0036)
Schooling years	0.0087***	0.0098***
	(0.00071)	(0.00073)
Potential experience	-0.0006	-0.0004
	(0.00068)	(0.00061)
Potential experience squared	0.0000	0.0000
	(0.00001)	(0.00001)
Observations	14377	14010
OLS with year effects		
Constant	0.2211***	0.1677***
	(0.01362)	(0.01291)
Logarithm of hourly wage (-1)	0.9021***	0.9058***
	(0.00381)	(0.0036)
Schooling years	0.0087***	0.0097***
	(0.00071)	(0.00073)
Potential experience	-0.0006	-0.0004
	(0.00068)	(0.00061)
Potential squared	0.0000	0.0000
	(0.00001)	(0.00001)
Observations	14377	14010
System GMM without year effects		
Constant	-11.4829***	-3.6570**
	(3.8822)	(1.7177)
Logarithm of hourly wage (-1)	-0.0378	0.5785***
	(0.23463)	(0.08951)
Schooling years	1.0650***	0.3449**
	(0.32775)	(0.13881)
Potential experience	0.0317***	0.0082**
	(0.01000)	(0.00385)
Potential experience squared	-0.0002**	-0.0000
	(0.00011)	(0.00005)
Arellano-Bond Test for AR (1) (p-value)	(0.771)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.000)	(0.001)
Arellano-Bond Test for AR (3) (p-value)	(0.563)	(0.612)
Hansen test of overidentification (p-value)	(0.002)	(0.008)
Observations	14377	14010

System GMM with year effects		
Constant	0.2076	-0.2066
	(0.79941)	(1.24902)
Logarithm of hourly wage (-1)	0.8093***	0.7710***
	(0.04420)	(0.05474)
Schooling years	0.0280	0.0626
	(0.05891)	(0.08946)
Potential experience	0.0025	0.0015
	(0.00670)	(0.00728)
Potential experience squared	-0.0001	-0.0000
	(0.00009)	(0.00006)
Arellano-Bond Test for AR (1) (p-value)	(0.000)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.010)	(0.001)
Arellano-Bond Test for AR (3) (p-value)	(0.359)	(0.410)
Hansen test of overidentification (p-value)	(0.100)	(0.420)
Observations	14377	14010

Notes: Year dummies not reported. Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%

Table 4.B1 (Appendix 4.B) shows the different estimation methods of static Mincer model for both men and women. Compared to the dynamic Mincer model estimations in table 4.2, the coefficient for the schooling years is much higher in the static models for both men and women. In other words, since the static model estimations do not consider the persistency of the observed earnings, they overestimate the return to schooling compared to the dynamic model estimations. Figure 4.1 and figure 4.2 show the return to schooling using the different estimation methods. The y-axis shows the return to schooling based on the observed wages. In the static system GMM method that controls for the endogeneity of variables and individual unobserved heterogeneity, the return to schooling is constant and does not depend on the potential work experience and it is estimated at $\beta = 0.8450$ and $\beta = 0.9635$ for men and women respectively. In fact, the return to schooling using the static system GMM method is equal to the coefficient of schooling in the wage equation for men and women.

In the dynamic system GMM method, using the formula extracted in section 4.3.2, the return to schooling based on the observed wages is equal to $\beta(x)=\rho\beta\sum_0^x(1-\rho)^x$. Thus, the return to schooling for men at $\beta(0)$, equals to $\rho\beta = 0.0280$, and with unlimited work experience, it changes to $\beta(\infty) = 0.0280/0.1907= 0.1468$. This indicates that the return to schooling based on the different levels of potential work experience is not the same. In other words, when the potential work experience increases, the return to schooling in terms of observed wages increases as well. The estimated return to schooling for women with zero potential work experience equals to $\beta(0)= \rho\beta=0.0626$ and it increase with unlimited work experience to $\beta(\infty)= 0.0626/0.2290= 0.2734$. Consequently, women also experience an increasing return to schooling based on the different potential work experience levels. Thus, the static system GMM that ignores the persistency of the observed wages, overestimates the return to schooling for men ($\beta(\infty)= 0.1468< \beta= 0.8450$) and women ($\beta(\infty)= 0.2734< \beta= 0.9635$). The high value of the return to schooling by the static system GMM method in compare to the other static models with OLS or random effects, may results from controlling for the endogeneity of variables as well as individual unobserved heterogeneity (Andini, 2014).

This is also in line with Andini (2013b) in which ignoring earning persistence makes the potential return to schooling to be biased upward.

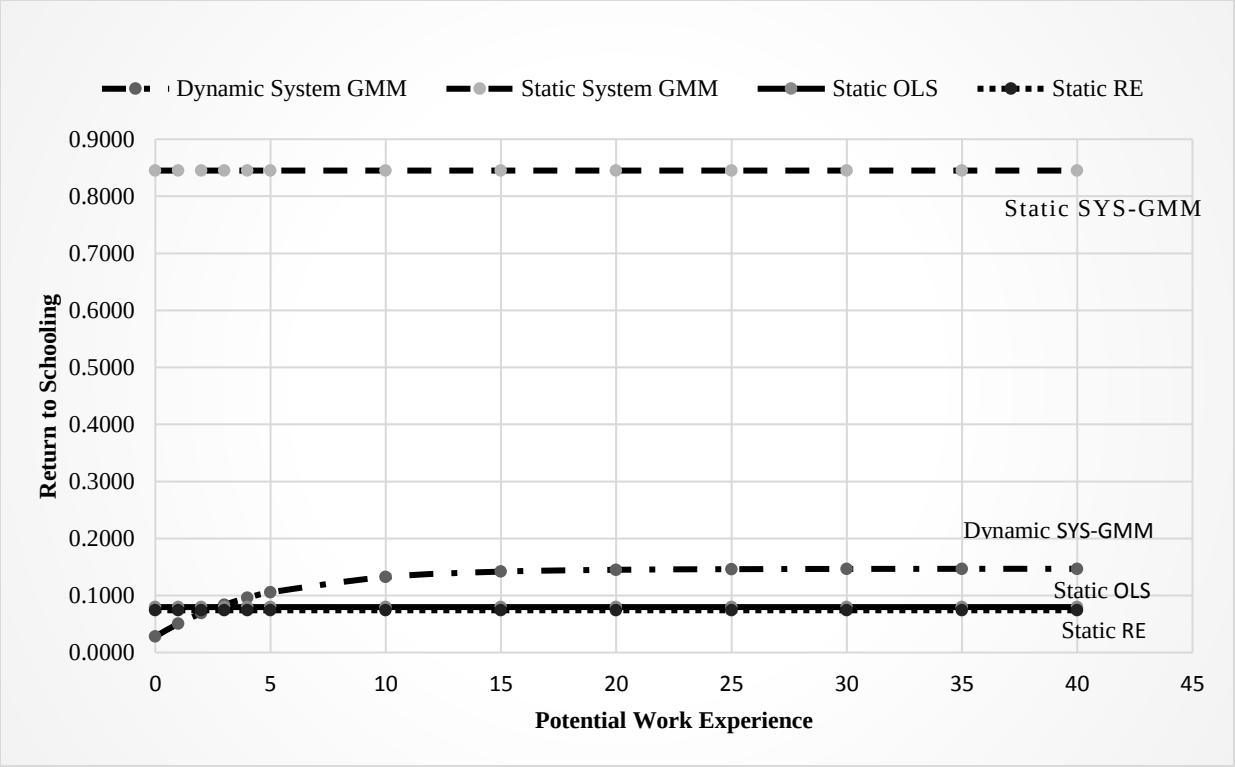

Figure 4. 1: Returns to Schooling for Men, Canada (2005-2010) SLID

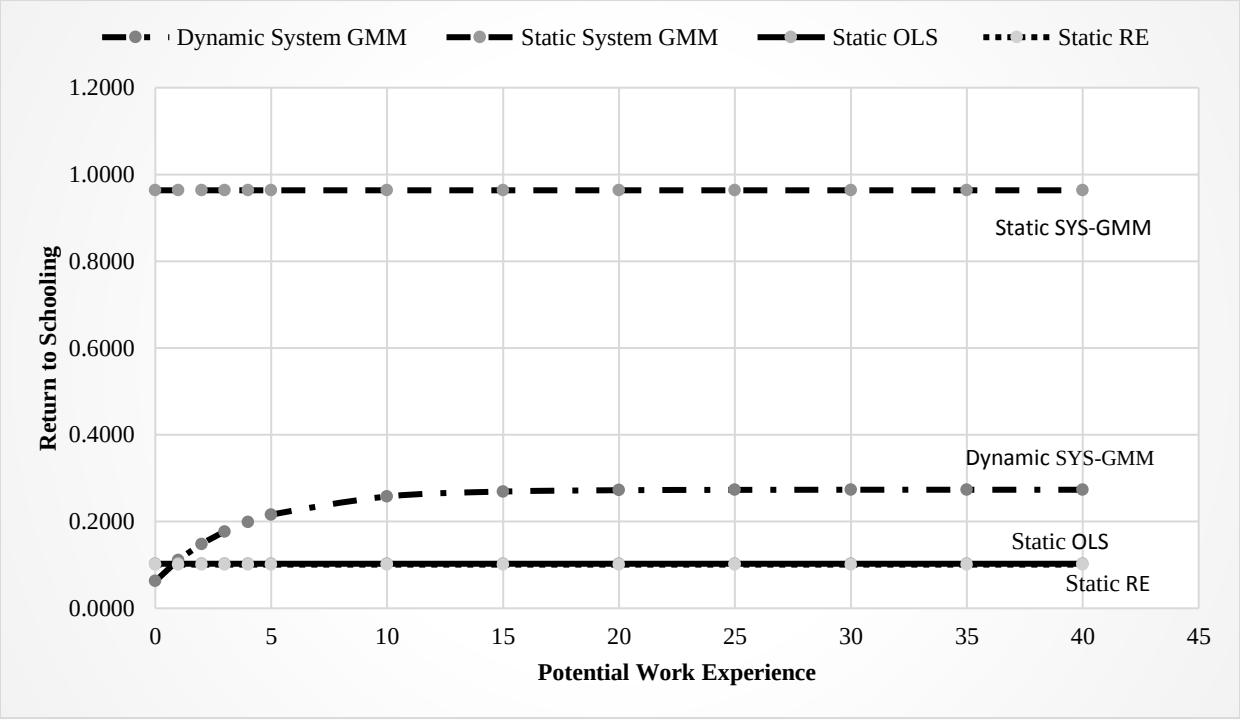

Figure 4. 2: Returns to Schooling for Women, Canada, (2005-2010) SLID

4.6.1. Sensitivity Analysis

In order to show the effect of endogeneity of schooling, the dynamic model is re-estimated by assuming that schooling is exogenous. Table 4.3 shows the dynamic Mincer model estimate using the system GMM method with exogenous schooling years. In the system GMM with year effects, results suggest that by taking schooling years as exogenous, the lag of hourly wage is positive and significant. However, it is slightly higher than its counterpart with the endogenous schooling years for both men (0.8311 vs. 0.8093) and women (0.8298 vs. 0.7710). The coefficient for the exogenous schooling years is significant and lower than its insignificant endogenous counterpart for men (0.0142 vs. 0.0280) and women (0.0178 vs. 0.0626). The estimated effect of potential experience is also positive and significant while it is higher compared to that of using endogenous schooling years for men (0.0053 vs. 0.0025) and women (0.0075 vs. 0.0015). Thus, even by taking schooling years as exogenous, the lag of hourly wage is positive and significant which confirms the earnings persistence have to be considered in the static Mincer function. Similar to the system GMM without year effects with endogenous schooling years, the results are also not valid with exogenous schooling years since the Hansen overidentification are significant for men and women respectively (p-value :0.005, 0.000 <0.05).

Table 4.3: Dynamic Mincer Model with Exogenous Schooling Years (S)

Dependent variable: logarithm of hourly wage	Men	Women
System GMM with year effects		
Constant	0.2872***	0.1764***
	(0.05391)	(0.06010)
Logarithm of hourly wage (-1)	0.8311***	0.8298***
	(0.03015)	(0.03195)
Schooling years	0.0142***	0.0178***
	(0.00260)	(0.00342)
Potential experience	0.0053*	0.0075**
	(0.00301)	(0.00350)
Potential experience squared	-0.0001*	-0.0001**
	(0.00006)	(0.00007)
Arellano-Bond Test for AR (1) (p-value)	(0.000)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.010)	(0.001)
Arellano-Bond Test for AR (3) (p-value)	(0.364)	(0.353)
Hansen test of overidentification (p-value)	(0.112)	(0.205)
Observations	14377	14010
System GMM without year effects		
Constant	0.2989***	0.1532***
	(0.05191)	(0.05602)
Logarithm of hourly wage (-1)	0.8299***	0.8555***
	(0.02750)	(0.02697)
Schooling years	0.0145***	0.0153***
	(0.00242)	(0.00299)
Potential experience	0.0043	0.0065*
	(0.00314)	(0.00381)
Potential experience squared	-0.0000	-0.0001*
	(0.00006)	(0.00007)
Arellano-Bond Test for AR (1) (p-value)	(0.000)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.008)	(0.001)
Arellano-Bond Test for AR (3) (p-value)	(0.367)	(0.309)
Hansen test of overidentification (p-value)	(0.005)	(0.000)
Observations	14377	14010

Notes: Year dummies not reported. Standard errors are in parenthesis. Significance levels: *10%, **5%, ***1%.

Another robustness check is to use other control variables such as marital status, union membership, public sector, province of residence and occupations in the dynamic estimation. Table 4.4 presents the system GMM with year effects for both men and women. Results show that even after controlling for other covariates, lag of the logarithm of hourly wage is still positive and significant for both men and women. However, other control variables are not significant. This confirms the earlier finding by Andini (2013a) that the extra control variables such as marital status, union membership, province of residence and occupations may be correlated with schooling variable. Thus, they may indirectly depend on the past wage which leads to most of the schooling dependent variables not to be statistically significant.

Although in the standard Mincer equation, potential experience is used based on the Mincer assumption, this may overestimate the actual work experience. Thus, the study also uses actual work experience as a robustness check in table 4.5. Results suggest that even by using actual experience as a proxy for the potential experience, lag of hourly wage is positive and significant, but it is slightly higher than that of using potential experience.

Table 4.4: Dynamic Mincer Model with Extra Control Variables

Dependent variable: logarithm of hourly wage	Men	Women
System GMM with year effects		
Constant	-4.2938	-2.8439
	(4.60626)	(3.77708)
Logarithm of hourly wage (-1)	0.5776***	0.4316***
	(0.16686)	(0.11271)
Schooling years	0.4403	0.3898
	(0.36192)	(0.25115)
Potential experience	0.0239	0.0027
	(0.02078)	(0.02373)
Potential experience squared	-0.0001	-0.0001
	(0.00011)	(0.00010)
Union	0.1025	-0.0971
	(0.10349)	(0.08221)
Public sector	0.0182	0.4384***
	(0.22672)	(0.13599)
Married	0.0356	-0.1467
	(0.09660)	(0.08969)
Occupations		
Business, Finance and Administrative	-0.0494	-0.2821
	(0.18531)	(0.17662)
Natural and Applied Sciences	-0.0760	-0.1383
	(0.16071)	(0.24343)
Health	-0.1308	-0.3163
	(0.25883)	(0.31079)
Social Science, Education, Government Service	-0.0091	-0.3795
and Religion	(0.33767)	(0.19551)
Art, Culture, Recreation and Sport	-0.0808	-0.1808
	(0.49092)	(0.36441)
Sales and Service	-0.3115	-0.2599
	(0.23179)	(0.17398)
Trades, Transport and Equipment Operators	-0.0063	-0.5787**
	(0.17620)	(0.25313)
Unique to Primary Industry	0.0854	-0.4736**
	(0.26721)	(0.23239)
Unique to Processing, Manufacturing and Utilities	-0.1686	-0.1825
	(0.18341)	(0.26211)
Lives in East	-0.5394	-0.7441
	(1.24318)	(0.60770)
Lives in Prairies	-0.9386	-0.6792
	(0.96841)	(0.75557)
Lives in Quebec	-1.0370	-1.4052
	(0.97956)	(1.18399)

Lives in BC	-1.9098	0.2831
	(1.17240)	(0.84711)
Arellano-Bond Test for AR (1) (p-value)	(0.000)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.003)	(0.002)
Arellano-Bond Test for AR (3) (p-value)	(0.572)	(0.693)
Hansen test of overidentification (p-value)	(0.701)	(0.504)
Observations	13747	13553

Notes: Year dummies not reported. Standard errors are in parenthesis. Significance levels: *10%, **5%, ***1%. References are: not a member of union or covered by a collective agreement, not working in the public sector, not married, management occupations, Ontario.

Table 4.5: Dynamic Mincer Model with Actual Work Experience

Dependent variable: logarithm of hourly wage	Men	Women
OLS without year effects		
Constant	0.1922***	0.1352***
	(0.01230)	(0.01032)
Logarithm of hourly wage (-1)	0.9138***	0.9097***
	(0.00401)	(0.00377)
Schooling years	0.0077***	0.0103***
	(0.00070)	(0.00071)
Actual experience	-0.0005	0.0004
	(0.00001)	(0.00048)
Actual experience squared	0.0000	-0.0000
	(0.00001)	(0.00001)
OLS with year effects		
Constant	0.1935***	0.1371***
	(0.01246)	(0.01047)
Logarithm of hourly wage (-1)	0.9142***	0.9094***
	(0.00401)	(0.00378)
Schooling years	0.0077***	0.0102***
	(0.00070)	(0.00071)
Actual experience	-0.0005	0.0003
	(0.00060)	(0.00048)
Actual experience squared	0.0000	-0.0000
	(0.00001)	(0.00001)
System GMM without year effects		
Constant	-4.8472	-10.5184**
	(4.86908)	(4.29070)
Logarithm of hourly wage (-1)	0.5571**	0.2038
	(0.28184)	(0.22254)
Schooling years	0.4546	0.9151***
	(0.41820)	(0.34668)
Actual experience	0.0124	0.0251***
	(0.01048)	(0.00748)
Actual experience squared	-0.0001	-0.0002**
	(0.00009)	(0.00008)
Arellano-Bond Test for AR (1) (p-value)	(0.052)	(0.259)
Arellano-Bond Test for AR (2) (p-value)	(0.003)	(0.000)
Arellano-Bond Test for AR (3) (p-value)	(0.654)	(0.446)
Hansen test of overidentification (p-value)	(0.039)	(0.002)
Observations	12312	12436
System GMM with year effects		
Constant	-0.0435	-0.6017
	(0.99797)	(0.95737)
Logarithm of hourly wage (-1)	0.8835***	0.8555***
	(0.04189)	(0.05531)

Schooling years	0.0233	0.0635
	(0.05786)	(0.06281)
Actual experience	0.0064	0.0118
	(0.00868)	(0.01130)
Actual experience squared	-0.0000	-0.0001
	(0.00008)	(0.00008)
Arellano-Bond Test for AR (1) (p-value)	(0.000)	(0.000)
Arellano-Bond Test for AR (2) (p-value)	(0.001)	(0.002)
Arellano-Bond Test for AR (3) (p-value)	(0.618)	(0.485)
Hansen test of overidentification (p-value)	(0.737)	(0.369)
Observations	12312	12436

Notes: Year dummies not reported. Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%.

4.7. Discussion and Conclusion

The purpose of this study is to investigate returns to education in a dynamic approach by controlling for earning persistence, endogeneity and individual unobserved heterogeneity. The study uses panel 5 (2005-2010) of Longitudinal Survey of Income Dynamics (SLID) for men and women between the age of 25 and 64. Results show that the hourly wage is highly persistent for both men and women. Thus, this rejects the standard Mincer assumption in which observed earning is equal to net potential earning. The results are also in line with earlier work by Andini (2013a) and Andini (2013b) that find persistency of hourly wage and reject the static Mincer assumption.

In addition, according to the dynamic Mincer model, the return to schooling shows an increasing trend at the beginning of the working life for both men and women while the static Mincer equation has a constant return to schooling overtime. Furthermore, results show that ignoring earning persistence in the static Mincer model, overestimates the return to schooling. The results also confirm findings by Andini (2013b) that the estimated return to schooling is constant and overestimated by the static Mincer model while the dynamic Mincer model shows an increasing return to schooling at the beginning of the working life.

For a robustness check, the study also takes into account other control variables such as marital status, union membership, public sector, province of residence and occupations. Results show that even after controlling for extra control variables, the lag of the logarithm of hourly wage is still positive and significant for both men and women while the extra control variables are not significant. This may indicate the correlation between these variables with schooling years which indirectly dependent on the past wage (Andini 2013a).

In contrast to Andini (2013b), where only the return to education for men was investigated, this study estimates the return to education for women. The results show that the return to schooling for women is greater than men. This is consistent with the majority of previous studies that find the higher return to schooling for women than men (Gwartney and Long, 1978; Carlson and Swartz, 1988; Trostel, Walker, and Woolley, 2002; Psacharopoulos and Patrinos, 2002). In this regard, there are some possible explanations for the higher return to schooling for women including an inverse relationship between years of schooling and discrimination, tastes, and circumstances factors; a male-female differential in the quality of educational attainment; occupational segregation of women into sectors where the returns to schooling are relatively high; biased estimates due to a failure to take account of sample selection; and biased estimates due to a failure to take account of the endogeneity of schooling or work experience (Dougherty, 2005). As noted by Dougherty (2005), schooling may have two effect on the earnings of women; a direct human capital effect in which schooling can increase women's skills and productivity similar to men and an indirect effect through attenuation of the adverse impact of discrimination, tastes and circumstances factors. Particularly, he shows that about half of the male-female differential in the returns to schooling is attributed to discrimination, tastes, and circumstances factors.

Regarding the higher return to education for women, government should invest more on women's schooling. For instance, to increase enrolment and completion rates, financial incentives such as cash transfers, scholarships and stipends are among the most effective interventions (Vanderkooy, 2018). In fact, these interventions can reduce gender

enrolment and achievement gaps and increase gender equality and empowerment at all education levels (Vanderkooy, 2018).

Results show that the static model overestimates the return to schooling since it ignores persistency of wages. In other words, the static model finds a constant return to schooling overtime. While according to the dynamic model, the return to schooling is increasing at the beginning of working life and it depends on the potential work experience. Therefore, using static models gives a wrong inference about the magnitude of the return to schooling. Consequently, policy makers should consider the dynamic return to schooling in schooling investment decisions. However, the direction of the schooling bias depends on many factors such as the number and type of covariates used in the wage equation, on whether the individual unobserved heterogeneity is controlled or not, on whether the endogeneity of variables are controlled or not and the number and the type of instrumental variables used in the estimation (Andini, 2014).

In order to demonstrate the effect of endogeneity of schooling, the dynamic model is re-estimated by assuming that schooling is exogenous. Results show that the logarithm of hourly observed wage is still positive and significant. Unlike the preferred estimation with endogenous schooling years, all the variables are significant and have the expected signs. Hence, by taking schooling years as exogenous, the lag of the dependent variable is slightly overestimated especially for women, but it still rejects the standard Mincer assumption in which potential earning considered to be equal to the observed earning.

Mincer assumption on potential work experience may overestimate the actual work experience Thus, the study uses the actual work experience as a proxy for the potential work experience as a robustness check. Results suggest that the lag of hourly

wage is positive and significant, but it is slightly higher compared to the one with potential work experience.

The purpose of this study is to provide a more precise estimation for the standard Mincer equation; however, there are still some limitations and some avenues for future research. In this study, schooling years is considered as a linear term in the estimations; however, there would be non-linearity in schooling years (Trostel, 2005; Lemieux, 2006). Thus, using schooling levels or qualifications instead of schooling years can be a good extension for further analysis.

Since the return to education varies greatly depending on occupation, field of study and the match between these two factors (Lemieux, 2014), using a job-match indicator that shows the relatedness of the occupation and filed of study can be helpful to reach a better model specification as a future research. Identifying the dynamic effect of education on within group wage inequality through quantile-regression method that focuses on the impact of schooling on the shape of the conditional wage distribution can be another interesting topic for further analysis.

Appendix 4.A

In order to estimate the dynamic return to schooling over the entire working life, Andini (2010) and Andini (2013b) write the equation for the dynamic model for each time period and compute the total return to schooling. For instance, at time s, the following expression for the dynamic model can be written as follow:

$$\ln w_s \approx \rho\alpha + (1-\rho)\ln\overline{w}_{s-1} + \rho\beta s + \rho\lambda 0 + \rho\gamma\, 0^2 \qquad (1)$$

Thus, the total return to schooling at time s would be:

$$\beta(0) = \frac{\partial \ln w_s}{\partial s} \approx \rho\beta \qquad (2)$$

and at time s+1, the model forms as bellow:

$$\ln w_{s+1} \approx \rho\alpha + (1-\rho)\ln w_s + \rho\beta s + \rho\lambda 1 + \rho\gamma\, 1^2 \qquad (3)$$

and the total return to schooling at time s+1 is:

$$\beta(1) = \frac{\partial \ln w_{s+1}}{\partial s} \approx \rho\beta + \rho\beta(1-\rho) \qquad (4)$$

Therefore, the formula for the dynamic return to schooling in terms of observed wages at time s+x is as follows:

$$\beta(x) = \frac{\partial \ln w_{s+x}}{\partial s} \approx \rho\beta\,[\,1 + (1-\rho) + (1-\rho)^2 + (1-\rho)^3 + \ldots + (1-\rho)^x\,] = \rho\beta \sum_0^x (1-\rho)^x. \qquad (5)$$

the above formula shows that the dynamic return to schooling based on the observed wages is not constant and it depends on potential work experience (x).

Appendix 4.B

Table 4.B1: Static Mincer Estimations for Men and Women

Dependent variable: logarithm of hourly wage	Men	Women
OLS with year effects		
Schooling years	0.0797***	0.1027***
	(0.00116)	(0.00125)
Potential work experience	0.0272***	0.0184***
	(0.00113)	(0.00109)
Potential work experience square	-0.0004***	-0.0003***
	(0.00002)	(0.00002)
Constants	1.5331***	1.1237***
	(0.02178)	(0.02320)
Observations	20,877	20,282
OLS without year effects		
Schooling years	0.0805***	0.1039***
	(0.00116)	(0.00125)
Potential work experience	0.0271***	0.0185***
	(0.00114)	(0.00109)
Potential work experience square	-0.0004***	-0.0003***
	(0.00002)	(0.00002)
Constants	1.5685***	1.1535***
	(0.02129)	(0.02289)
Observations	20,877	20,282
RE with year effects		
Schooling years	0.0741***	0.1007***
	(0.00212)	(0.00232)
Potential work experience	0.0325***	0.0220***
	(0.00144)	(0.00136)
Potential work experience square	-0.0005***	-0.0003***
	(0.00002)	(0.00002)
Constants	1.5485***	1.0907***
	(0.03519)	(0.03877)
Observations	20,877	20,282
RE without year effects		
Schooling years	0.0814***	0.1127***
	(0.00210)	(0.00230)
Potential work experience	0.0345***	0.0250***
	(0.00146)	(0.00137)
Potential work experience square	-0.0005***	-0.0003***
	(0.00002)	(0.00002)
Constants	1.4101***	0.8528***
	(0.03469)	(0.03792)
Observations	20,877	20,282

Static System GMM with year effects		
Schooling years	0.8450*	0.9635**
	(0.47809)	(0.44621)
Potential work experience	0.0712*	0.0773*
	(0.04279)	(0.04374)
Potential work experience square	-0.0004***	-0.0002***
	(0.00009)	(0.00007)
Constants	-9.5101	-11.8929*
	(7.19909)	(6.98868)
Number of observations	20,877	20,282
Number of groups	5459	5289
Number of instruments	15	15
Arellano–Bond test for AR (1) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (2) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (3) in differences (p-value)	(0.215)	(0.147)
Hansen test of joint validity of instruments (p-value)	(0.256)	(0.329)
Static System GMM without year effects		
Schooling years	0.9968***	1.0124***
	(0.21260)	(0.20876)
Potential work experience	0.0415***	0.0317***
	(0.00503)	(0.00406)
Potential work experience square	-0.0004***	-0.0002***
	(0.00009)	(0.00006)
Constants	-10.7935***	-11.47***
	(2.78433)	(2.81014)
Number of observations	20,877	20,282
Number of groups	5459	5289
Number of instruments	10	10
Arellano–Bond test for AR (1) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (2) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (3) in differences (p-value)	(0.233)	(0.180)
Hansen test of joint validity of instruments(p-value)	(0.000)	(0.002)
Static System GMM with year effects and exogeneous s		
Schooling years	0.0786***	0.1006***
	(0.00249)	(0.0027)
Potential work experience	0.0349***	0.0239***
	(0.00412)	(0.00352)
Potential work experience square	-0.0005***	-0.0004***
	(0.00008)	(0.00007)
Constants	1.4666***	1.1018***
	(0.05753)	(0.05788)
Number of observations	20,877	20,282
Number of groups	5459	5289

Number of instruments	17	17
Arellano–Bond test for AR (1) in differences(p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (2) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (3) in differences (p-value)	(0.228)	(0.142)
Hansen test of joint validity of instruments (p-value)	0.011	(0.000)
Static System GMM without year effects and with exogenous s		
Schooling years	0.0830***	0.1082***
	(0.00258)	(0.00294)
Potential work experience	0.0395***	0.0262***
	(0.00471)	(0.00427)
Potential work experience square	-0.0006***	-0.0003***
	(0.00009)	(0.00008)
Constants	1.3658***	0.9528***
	(0.06270)	(0.06550)
Number of observations	20,877	20,282
Number of groups	5459	5289
Number of instruments	12	12
Arellano–Bond test for AR (1) in differences(p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (2) in differences (p-value)	(0.000)	(0.000)
Arellano–Bond test for AR (3) in differences(p-value)	(0.374)	(0.336)
Hansen test of joint validity of instruments(p-value)	(0.000)	(0.000)

Notes: Year dummies not reported. Standard Errors are in parenthesis. Significance levels: *10%, **5%, ***1%.

References

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American statistical Association*, *76*(375), 598-606.

Andini, C. (2007). Returns to education and wage equations: a dynamic approach. *Applied Economics Letters*, *14*(8), 577-579.

Andini, C. (2010). A dynamic Mincer equation with an application to Portuguese data. *Applied Economics*, *42*(16), 2091-2098.

Andini, C. (2013a). How well does a dynamic Mincer equation fit NLSY data? Evidence based on a simple wage-bargaining model. *Empirical Economics*, *44*(3), 1519-1543.

Andini, C. (2013b). Earnings persistence and schooling returns. *Economics letters*, *118*(3), 482-484.

Andini, C. (2014). Persistence bias and schooling returns.

Angrist, J. D., & Keueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings?. *The Quarterly Journal of Economics*, *106*(4), 979-1014.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, *58*(2), 277-297.

Arias, O., Hallock, K. F., & Sosa-Escudero, W. (2002). Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. In *Economic applications of quantile regression* (pp. 7-40). Physica, Heidelberg.

Aydede, Y., & Dar, A. (2017). Is the lower return to immigrants' foreign schooling a postarrival problem in Canada?. *IZA Journal of Migration*, *6*(1), 4.

Bar-Or, Y., Burbidge, J., Magee, L., & Robb, A. L. (1995). The wage premium to a university education in Canada, 1971-1991. *Journal of Labor Economics, 13*(4), 762-794.

Becker, G. (1964). Human Capital, Nueva York. *National Bureau of Economic Research*, *8*(299), 493-517.

Belzil, C. (2008). Testing the specification of the Mincer wage equation. *Annales d'Economie et de Statistique*, 427-451.

Berger, J., & Parkin, A. (2009). The value of a degree: Education, employment and earnings in Canada. *The price of knowledge: Access and student finance in Canada*, 7-22.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, *87*(1), 115-143.

Boudarbat, B., Lemieux, T., & Riddell, W. C. (2010). The evolution of the returns to human capital in Canada, 1980–2005. *Canadian Public Policy*, *36*(1), 63-89.

Bound, J., Brown, C., Duncan, G. J., & Rodgers, W. L. (1989). *Measurement error in cross-sectional and longitudinal labor market surveys: Results from two validation studies* (No. w2884). National Bureau of Economic Research.

Bourbeau, E., Lefebvre, P., & Merrigan, P. (2012). The Evolution of the Returns to Education for 21-to 35-Year-Olds in Canada and across Provinces: Results from the 1991–2006 Analytical Census Files. *Canadian Public Policy*, *38*(4), 531-549.

Burbidge, J. B., Magee, L., & Robb, A. L. (2002). The education premium in Canada and the United States. *Canadian Public Policy/Analyse de politiques*, 203-217.

Card, D. (1993). *Using geographic variation in college proximity to estimate the return to schooling* (No. w4483). National Bureau of Economic Research.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics* (Vol. 3, pp. 1801-1863). Elsevier.

Carlson, L. A., & Swartz, C. (1988). The earnings of women and ethnic minorities, 1959–1979. *ILR Review*, *41*(4), 530-546.

Dean, J. (2010). The main culprit behind immigrant wage gaps in Canada: Lower returns to foreign schooling or work experience?. *Available at SSRN 2211293*.

DiBartolo, A. (1999). *Modern human capital analysis: estimation of US, Canada and Italy earning functions* (No. 212). LIS Working Paper Series.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men?. *Journal of Human Resources*, *40*(4), 969-988.

Flannery, M. J., & Rangan, K. P. (2006). Partial adjustment toward target capital structures. *Journal of financial economics*, *79*(3), 469-506.

Freeman, R. B., & Needels, K. (1993). Skill differentials in Canada in an era of rising labor market inequality. In *Small differences that matter: Labor markets and income maintenance in Canada and the United States* (pp. 45-68). University of Chicago Press.

García-Suaza, A. F., Guataquí, J. C., Guerra, J. A., & Maldonado, D. (2014). Beyond the Mincer equation: the internal rate of return to higher education in Colombia. *Education Economics*, *22*(3), 328-344.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society*, 1-22.

Gwartney, J. D., & Long, J. E. (1978). The relative earnings of blacks and other minorities. *ILR Review*, *31*(3), 336-346.

Hartog, J., & Gerritsen, S. (2016). Mincer Earnings Functions for the Netherlands 1962–2012. *De Economist*, *164*(3), 235-253.

Heckman, J. J., Humphries, J. E., & Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, *126*(S1), S197-S246.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2003). *Fifty years of Mincer earnings regressions* (No. w9732). National Bureau of Economic Research.

Kane, T. J., & Rouse, C. E. (1993). *Labor market returns to two-and four-year colleges: is a credit a credit and do degrees matter?* (No. w4268). National Bureau of Economic Research.

Kohn, J. L., & Averett, S. L. (2014a). Can't we just live together? New evidence on the effect of relationship status on health. *Journal of Family and Economic Issues*, *35*(3), 295-312.

Kohn, J. L., & Averett, S. L. (2014b). The effect of relationship status on health with dynamic health and persistent relationships. *Journal of Health Economics*, *36*, 69-83.

Lee, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics*, *141*(2), 1131-1158.

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer a pioneer of modern labor economics* (pp. 127-145). Springer, Boston, MA.

Lemieux, T. (2014). Occupations, fields of study and returns to education. *Canadian Journal of Economics/Revue canadienne d'économique*, *47*(4), 1047-1077.

Ma, J., Pender, M., & Welch, M. (2016). Education Pays 2016: The Benefits of Higher Education for Individuals and Society. Trends in Higher Education Series. *College Board*.

Mazza, A., Battisti, M., Ingrassia, S., & Punzo, A. (2018, April). Modeling Return to Education in Heterogeneous Populations. In *PAA 2018 Annual Meeting*. PAA.

Millar, W. J., & Wigle, D. T. (1986). Socioeconomic disparities in risk factors for cardiovascular disease. *CMAJ: Canadian Medical Association Journal*, *134*(2), 127.

Mincer, J. (1974). Schooling, Experience and Earnings. New York: Columbia University Press.

Murphy, K. M., Riddell, W. C., & Romer, P. M. (1998). *Wages, skills, and technology in the united states and canada* (No. w6638). National Bureau of Economic Research.

Murphy, K. M., & Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor economics*, *8*(2), 202-229.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.

Oreopoulos, P. (2006). The compelling effects of compulsory schooling: Evidence from Canada. *Canadian Journal of Economics/Revue canadienne d'économique*, *39*(1), 22-52.

Parliamentary Budget Officer (PBO) (2016). Federal Spending on Postsecondary Education. http://www.pbo-dpb.gc.ca/web/default/files/Documents/Reports/2016/PSE/PSE_EN.pdf.

Patrinos, H. A., Psacharopoulos, G., & Tansel, A. (2019). *Returns to Investment in Education: The Case of Turkey*. The World Bank.

Pat, W. (2016). To Stay Relevant, Your Company and Employees Must Keep Learning. *Harvard Business Review, 7*. https://hbr.org/2016/03/to-stay-relevant-your-company-and-employees-must-keep-learning.Psacharopoulos, G., & Patrinos, H. (2002). Returns to Investment in Education: A further update 'World Bank policy Research Working Paper 2881. Washington DC. *DC: The world Bank.*

Ren, F., & Shannon, M. (2017). Cohort Effects and the Returns to Canadian Undergraduate Education. *Canadian Public Policy*, *43*(3), 221-245.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The stata journal*, *9*(1), 86-136.

Statistics Canada. (2016). Education and occupation of high-income Canadians. Statistics Canada Catalogue no. 99-014-X2011003. Last modified: 2016-09-15 http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-014-x/99-014-x2011003_2-eng.cfm.

Statistics Canada. (2016). Tuition fees for degree programs, 2016/17. http://www.statcan.gc.ca/daily-quotidien/160907/dq160907a-eng.htm.

Statistics Canada. (2015). Survey of Labour and Income Dynamics (SLID)- a 2011 Survey Overview. Last modified: 2015-11-27. http://www.statcan.gc.ca/pub/75f0011x/75f0011x2013001-eng.htm.

Strumpel, B. (1971). Higher education and economic behavior. *A degree and what else*, 55-79.

Trostel, P. A. (2005). Nonlinearity in the return to education. *Journal of Applied Economics*, *8*(1), 191-202.

Trostel, P., Walker, I., & Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour economics*, *9*(1), 1-16.

Turcotte, M. (2015). Civic engagement and political participation in Canada. Statistics Canada Catalogue no. 89-652-X2015006. Last modified: 2015-11-30. https://www150.statcan.gc.ca/n1/pub/89-652-x/89-652-x2015006-eng.htm.

Vanderkooy, A. (May 28, 2018). Ensuring Education for Women and Girls at All Levels: A Feminist G7 Agenda. Retrieved from http://cidpnsi.ca/ensuring-education-for-women-and-girls-at-all-levels/

Van Pelt, D. N., & Emes, J. (2015). *Education Spending in Canada: What's Actually Happening?*. Fraser Institute.

Chapter 5. Conclusions

This dissertation includes three essays as follows: the first essay examines a dynamic effect of diabetes on employment in Canada. The second essay extends the dynamic analysis and uses a broader measure of health (self-reported health) and investigates its interaction with employment. The third essay explores a dynamic effect of education on hourly wage.

The first essay examines a dynamic effect of diabetes on employment in Canada. It takes data from 9 cycles of NPHS and uses a dynamic probit model with random effects to control for endogeneity of diabetes and persistency of employment. Results show that diabetes has positive and insignificant effect on employment for men. While, the effect of diabetes on employment for women is negative and significant. The results confirm the signs and significance of diabetes coefficients estimated by static studies such as Latif (2009); however, the numbers are much smaller (0.13 vs. 0.96 for men and -0.21 vs. 0-.44 for women). This may indicate that static and cross-sectional models overestimate the diabetes effect on employment. Particularly, a precise estimates of diabetes effect on employment would be helpful for policy makers, government or employers to know the economic burden and design the appropriate policies to reduce related costs. In this essay, the goal is to provide more accurate estimates of the effect of diabetes on employment; however, there are some limitations and more avenues for further research. For instance, the first essay has not differentiated between the main types of diabetes (type 1 and type 2) since NPHS has not specified the types of diabetes. This issue is important because the impact of diabetes on employment may vary depending on which types of diabetes are considered. As type 2 diabetes accounts for

about 90% of all individuals with diagnosed diabetes, this essay restricted the age group to 25-64 to moderate this problem.

This essay has examined the dynamic probit model using a 2SLS regression analysis. In this regard, the fitted value of the endogenous diabetes dummy variable from the first stage and family history of diabetes are used as instruments to put into the second step. Thus, one alternative would be using a recursive bivariate probit model in which employment and diabetes equations are considered jointly through the correlation of the error terms which results in more efficient estimates.

The second essay extends the dynamic analysis and uses a broader measure of health (self-reported health) and investigates its relationship with employment. The previous Canadian studies that have examined the effect of health on labour market outcomes are static or cross-sectional (Breslaw and Stelener, 1986; Au et al., 2005; Hum, Simpson and Fissuh, 2008) and they have used an instrumental variable approach (IV) to account for the endogeneity of health which does not give an efficient estimate for health effect. Thus, this study contributes to the literature using a simultaneous equation model to account for endogeneity of health along with persistency of health and employment to obtain more precise model specification for the interaction of health and labour market outcome. Results show that there is a high state dependency in employment and health for both men and women. This is consistent with earlier studies such as Hann and Myck (2009) that use German data for men aged 30-59 and find strong persistency in health and labour market risks. Moreover, there is a highly significant and positive effect of health on employment for both men and women. The coefficient estimates of the latent health in employment equation are positive and significant for both men and women respectively

(0.3246 vs. 0.1274). For the reverse effect, employment also has a positive and significant effect on health for both men and women. The coefficient estimates for the latent employment in health equation is positive and significant for both men and women respectively (0.2257 vs. 0.2207).

In the dynamic bivariate probit model, the lag of employment and the lag of health variable in the health and employment equations are also controlled respectively. Results show that in the employment equation, being healthy at previous period increases the probability of employment for both men and women; however, this effect is much higher for women than men. On the other hand, in the health equation, being employed at previous period has a negative and significant effect on the probability of health for both men and women.

Results demonstrate that the interactions between health and employment is different for different age groups. For instance, health is only endogenous for women 35 to 49 and men 50 to 64. Moreover, the effect of lag of employment on the probability of health is positive and significant for most age groups (although this effect is negative for women 35 to 49 and men 50 to 64 but it is not significant). On the other hand, the effect of lag of health on the probability of employment is positive and significant for most age groups (for women 35 to 49 and men 50 to 64, the effect is positive but not significant). Since the results show that there is a positive and significant interaction between health and employment, thus, health policies such as workplace programs that have positive and direct effect on health, can have positive and indirect effect on employment.

Using different instrumental variable approaches such as two stage predictor substitution (2SPS) and two-stage residual inclusion (2SRI) and compare the results with

those with a simultaneous equation model can be a good extension for further research. Particularly, it is expected that a simultaneous equation model gives more efficient results compared to a two-stage method since it considers the correlation of health and employment through their error terms.

The Canadian studies that have examined the return to education are static or cross-sectional (Murphy, Riddell & Romer; 1998; Burbidge, Magee & Robb, 2002; Lemieux, 2006; Oreopoulos, 2006; Boudarbat, Lemieux & Riddell, 2010; Bourbeau, Lefebvre & Merrigan, 2012; Lemieux, 2014; Ren and Shannon, 2017; Aydede and Dar, 2017) and they have not investigated the return to education through a dynamic approach. Particularly, they have not considered earnings persistence in estimating the return to education. Moreover, they have used an instrumental variable approach (IV) to solve the endogeneity of education. The third essay contributes to the Canadian literature by estimating a dynamic Mincer model through a system GMM method to obtain more precise model specification for the return to education. In this regard, it uses the longitudinal Survey of Labour and Income Dynamics (SLID) and estimates the dynamic model for men and women between the age of 25 and 64 separately.

Results demonstrate that the hourly wage is highly persistent for both men and women between the age 25 and 64. Consequently, the persistency of hourly wage results in rejection of the static Mincer assumption in which observed earning and net potential earning are equal. Earlier works such as Andini (2013a) and Andini (2013b) also find persistency of hourly wage and reject the static Mincer assumption. The results also show that the return to schooling is increasing at the beginning of the working life for both men and women compared to a constant return to schooling by static Mincer function. In

addition, the static Mincer function overestimates the return to schooling since it ignores earning persistence. However, Andini (2013b) estimates the return to education only for men, this study investigates the return to education for women as well. The results demonstrate that the return to schooling for women is greater than men. This is in line with the most previous studies that find the higher return to schooling for women than men (Gwartney and Long, 1978; Carlson and Swartz, 1988; Trostel, Walker, and Woolley, 2002; Psacharopoulos and Patrinos, 2002). In this regard, there may exist many reasons for the higher return to schooling for women such as an inverse relationship between years of schooling and discrimination, tastes, and circumstances factors; a male-female differential in the quality of educational attainment or occupational segregation of women into sectors where the returns to schooling are relatively high (Dougherty, 2005). However, male-female differential in the returns to schooling can be mostly due to discrimination, tastes, and circumstances factors (Dougherty, 2005).

Although this essay is intended to give a more precise estimation for the standard Mincer equation, there are some limitations and some avenues for further research. For instance, there would be non-linearity in schooling years (Trostel, 2005; Lemieux, 2006), thus, schooling levels or qualifications instead of schooling years can be a good extension for further analysis. Identifying the dynamic effect of education on within group wage inequality by quantile-regression in which the effect of schooling on the shape of the conditional wage distribution can examined, may account as another extension for further analysis.

In summary, all the above-mentioned essays are linked through a dynamic investigation to show the importance of persistency of labour outcomes. The results

demonstrate that labour outcomes such as employment and hourly wage are highly persistent. Moreover, according to the results in the first two essays, ignoring the persistency of employment results in overestimation of diabetes /health effects on employment. Regarding the third essay, ignoring the persistency of hourly wage leads to overestimation of return to education.

Since the sudden emergence of the coronavirus disease (COVID-19) pandemic, many individuals lost their jobs. Specifically, the labour market impact of the pandemic was significantly severe for workers more exposed to disease such as nurses and front-line workers (Beland et al. 2020). It seems that the negative effect of the COVID-19, can have long-term consequences on labour market. For instance, Barrero, Bloom and Davis (2020) find that about 42 percent of recent pandemic-induced layoffs will lead to permanent job loss. Moreover, they indicate that if the economic shutdown stays for many months, or if a serious pandemic turns into a recurring disease, there will be long-term consequences for the reallocation of jobs, workers, and capital across firms and locations.

Since, our results find the persistency of employment, thus, individuals who are unemployed at current time are more likely to be unemployed at future. Moreover, the emergence of the COVID-19 pandemic can amplify the persistency of unemployment. In this regard, it is suggested that policy makers expand paid leave for workers, preparing unemployment insurance benefit program and help businesses transit to full-time teleworking to decrease the adverse effect of COVID-19 on labour market outcomes (National Conference of State Legislatures, May 2020).

References

Andini, C. (2013b). Earnings persistence and schooling returns. *Economics letters*, *118*(3), 482-484.

Andini, C. (2013a). How well does a dynamic Mincer equation fit NLSY data? Evidence based on a simple wage-bargaining model. *Empirical Economics*, *44*(3), 1519-1543.

Au, D. W. H., Crossley, T. F., & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health economics*, *14*(10), 999-1018.

Aydede, Y., & Dar, A. (2017). Is the lower return to immigrants' foreign schooling a postarrival problem in Canada?. *IZA Journal of Migration*, *6*(1), 4.

Boudarbat, B., Lemieux, T., & Riddell, W. C. (2010). The evolution of the returns to human capital in Canada, 1980–2005. *Canadian Public Policy*, *36*(1), 63-89.

Bourbeau, E., Lefebvre, P., & Merrigan, P. (2012). The Evolution of the Returns to Education for 21-to 35-Year-Olds in Canada and across Provinces: Results from the 1991–2006 Analytical Census Files. *Canadian Public Policy*, *38*(4), 531-549.

Breslaw, J. A., & Stelcner, M. (1987). The effect of health on the labor force behavior of elderly men in Canada. *Journal of Human Resources*, 490-517.

Burbidge, J. B., Magee, L., & Robb, A. L. (2002). The education premium in Canada and the United States. *Canadian Public Policy/Analyse de politiques*, 203-217.

Carlson, L. A., & Swartz, C. (1988). The earnings of women and ethnic minorities, 1959–1979. *ILR Review*, *41*(4), 530-546.

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men?. *Journal of Human Resources, 40*(4), 969-988.

Gwartney, J. D., & Long, J. E. (1978). The relative earnings of blacks and other minorities. *ILR Review, 31*(3), 336-346.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of health economics, 28*(6), 1116-1125.

Hum, D., Simpson, W., & Fissuh, E. (2008). The Impact of Health on Labour Supply over the Life Cycle. *Canada: Department of Economics, The University of Manitoba*.

Latif, E. (2009). The impact of diabetes on employment in Canada. *Health economics*, *18*(5), 577-589.

Lemieux, T. (2014). Occupations, fields of study and returns to education. *Canadian Journal of Economics/Revue canadienne d'économique, 47*(4), 1047-1077.

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In *Jacob Mincer a pioneer of modern labor economics* (pp. 127-145). Springer, Boston, MA.

Murphy, K. M., Riddell, W. C., & Romer, P. M. (1998). *Wages, skills, and technology in the united states and canada* (No. w6638). National Bureau of Economic Research.

National Conference of State Legislatures. (May, 2020). COVID-19: Impact on Employment and Labor. Retrieved from https://www.ncsl.org/research/labor-and-employment/covid-19-impact-on-employment-and-labor.aspx.

Oreopoulos, P. (2006). The compelling effects of compulsory schooling: Evidence from Canada. *Canadian Journal of Economics/Revue canadienne d'économique, 39*(1), 22-52.

Psacharopoulos, G., & Patrinos, H. (2002). Returns to Investment in Education: A further update 'World Bank policy Research Working Paper 2881. Washington DC. *DC: The world Bank.*

Ren, F., & Shannon, M. (2017). Cohort Effects and the Returns to Canadian Undergraduate Education. *Canadian Public Policy*, *43*(3), 221-245.

Trostel, P., Walker, I., & Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour economics*, *9*(1), 1-16.

Trostel, P. A. (2005). Nonlinearity in the return to education. *Journal of Applied Economics*, *8*(1), 191-202.