University of Ottawa

Doctoral Thesis

Essays in Labour Economics

Author:
Zechuan Deng

Supervisor:
Dr. Louis-Philippe Morin

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Department of Economics
Faculty of Social Sciences
University of Ottawa

Declaration of Authorship

I, Zechuan Deng, declare that this thesis titled "Essays in Labour Economics" and the work presented in it are my own. Chapter one and chapter three of this thesis are done by myself. The second chapter of this thesis is done jointly with Dr. Pierre Brochu and Dr. Jonathan Créchet. My contribution is equal to theirs. I, hereby acknowledge the contribution of Dr. Pierre Brochu and Dr. Jonathan Créchet for the research related to chapter two.

I confirm that:

- This work was done wholly or mainly while in candidature for research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Zechuan DENG

Date: September 2021

Acknowledgements

Special thanks go to my thesis supervisor Louis-Philippe Morin for the solid academic training and continuous support. I am indebted to him for reading and editing previous drafts of this dissertation. I would like to express my highest gratitude to Louis-Philippe Morin for setting the bar high and therefore forcing me to surpass my limits.

I thank my thesis committee, composed of three internal examiners: Pierre Brochu, Jonathan Créchet and Minjoon Lee, for their suggestions, ideas and recommendations during my first and second thesis workshops and thesis defence. I also thank Mary Connolly for being on my committee as the external examiner. Lastly, I want to thank Carole-Anne Gratton and Diane Ritchot for their administrative help.

I am grateful to Paul Hartung, Anna Kemeny, Siqin Guo, Felix Fortin, Guy Gellatly, Sanmartin, Claudia, Elizabeth Richards, René Morissette and all my other colleagues from Statistics Canada for their continuous support in the past few years, so I can manage both work and school at the same time.

Lastly, I thank Sicong Ma, Yufan Hu, Philippe Kabore and my colleagues at the University of Ottawa for their moral and academic support throughout my years of study.

Dedicated to my wife and mother for believing in me and supporting me throughout my doctoral studies. In memory of Peter Klein, my dearest father, who was always there for me.

Contents

List of Figures

List of Tables

Abstract

The first chapter examines the causal effect of education on individuals' retirement decisions and changes in quality of life after retirement. When estimating the return to education, much of the existing literature assumes implicitly that individuals optimize over their career with a fixed year of retirement ("fixed-retirement-age" assumption). Using the provincial compulsory schooling laws as instruments to estimate the causal effects of educational attainment on individuals' retirement decisions and the change in the quality of life after retirement, I find that years of schooling have no significant impact on retirement decisions. Furthermore, among those who are already retired, years of schooling do not have any impact on their age at retirement and the change in the quality of life after retirement. Results in this paper support the "fixed-retirement-age" assumption used in empirical studies based on the classical return to education model where the exogenous assumption on retirement age might not be a bad approximation.

The second chapter relates to labour market flows and worker trajectories in Canada during COVID-19. We use the confidential-use files of the Labour Force Survey (LFS) to study the employment dynamics in Canada from the beginning of the COVID-19 pandemic through to mid-summer. Using the longitudinal dimension of this dataset, we measure the size of worker reallocation and document the presence of high labour market churning that persists even after the easing of social-distancing restrictions. As of July of 2020, many of the recent job losers – especially those who had been temporarily laid-off between February and April – have regained employment. However, this apparent strong recovery dynamics hides important heterogeneity, and large groups of workers, such as those who were not employed prior to the pandemic, face important difficulties with finding a job. Our results further suggest that gross job losses were higher among women and young workers during the shutdown and that older workers were more likely to leave the labour force when the economy reopened.

The third chapter analyzes the gender differences in early career labour market trajectories and wage growth in Canada. Using the Longitudinal Workers File (LWF) linked to the 2006 and 2016 Census, I find that, although some progress has been made on the gender gap on labour market trajectories for young workers in their earlier careers, women faced higher penalties for taking time out of the labour market compared with men. More specifically, regardless of the type of job separation, changed employer or occupation, and reason for separation, women's weekly wage annual growth rate was lower compared with their male counterparts. The results from regression analysis suggest that labour market trajectories affect women's and men's weekly wage annual growth rates differently. Women's weekly wage annual growth rate was more sensitive to temporary job separation compared with men, whereas men's weekly wage annual growth rate was more sensitive to permanent job separation. The Oaxaca-Blinder decomposition shows that the total number of permanent separations and the total number of permanent separations due to parental/maternity leave each explains about one-third of the cross-sectional gender differences in weekly wage annual growth rate observed for young workers in Canada.

General Introduction

This dissertation includes three essays on labour economics. All three chapters consider historical and current economic shocks that happened in Canada and examine their consequences in the short and long run. This research contributes to several strands of the literature.

First, I contribute to the literature on testing the "fixed-retirement-age" assumption on the return to education model based on empirical evidence. When estimating the return to education, much of the existing literature assumes implicitly that individuals optimize over their career with a fixed year of retirement ("fixed-retirement-age" assumption). However, there is not a lot of evidence about whether or not education affects retirement. In theory, individuals with a higher level of educational attainment would have higher lifetime income and thus more money before retirement, which could lead to a shorter career. At the same time, a more enjoyable occupation as the result of a higher level of educational attainment might lead to later retirement, whereas a less enjoyable occupation might lead to the opposite. Therefore, whether education increases or decreases the retirement age is an empirical question. Second, I contribute to the literature by estimating the causal effect of educational attainment on the change of individuals' quality of life after retirement, which is an essential element of long-term nonpecuniary effects of education that is crucial for future policy adjustment such as the age requirement for the Old Age Pension plan in Canada.

The first chapter examines the causal effect of education on individuals' retirement decisions and changes in quality of life after retirement. Specifically, I use the historical changes to the provincial compulsory schooling laws in Canada as an instrument to estimate the causal relationship between education and retirement decisions. The Ordinary Least Squares (OLS) regression results suggest that educational attainment is not partially correlated with the probability of being retired at age 65. But among those who are retired, a higher level of educational attainment is associated with an older retirement age, which is an indication of delayed retirement. Moreover, the OLS regression results suggest that a higher level of educational attainment is associated with a lower level of life enjoyment after retirement. However, the OLS estimates are subject to the issue of endogeneity and cannot be interpreted as a causal effect. On the other hand, the Two-stage Least Squares (2SLS) regression results suggest the effect of education on retirement decisions and change in the quality of life after retirement is small and statistically insignificant. Thus from the perspective of this paper, the "fixed-retirement-age" assumption in empirical studies based on the classical return to education model seems valid, and the exogenous assumption on retirement age might not be a bad approximation. And, there are no long-term causal effects of education on the change in the quality of life after retirement in Canada. In sum, these results suggest that the change in margin of the educational attainment through compulsory schooling laws does not impact retirement decisions and the change in the quality of life after retirement. My current attempts to investigate the channels of the main empirical results show that total annual income is one of the potential channels that could lead to delayed retirement through education; however, due to endogeneity, it cannot be interpreted as a causal relationship, and there should be other channels that offset the effect of income, which is a total sum up to a null effect of education on retirement decisions.

In the second chapter, the labour market flows, and worker trajectories in Canada during

COVID-19 are examined. We rely on the confidential-use files of the Labour Force Survey (LFS) to study the employment dynamics in Canada from the beginning of the COVID-19 pandemic through to mid-summer. The Canadian labour market underwent an unprecedented trajectory since the onset of the COVID-19 pandemic. The massive job losses due to the shutdown of the early phase of the outbreak were followed by a vigorous rebound upon the gradual reopening of the economy, as was the case for many countries. However, the discussions have largely focused on the behaviour of the economy in net terms, and little is known about the employment flow dynamics that have accompanied the virus propagation. Although highly informative about the impact of COVID-19, these net changes are likely to hide sizable worker reallocation flows across labour-market states and jobs, especially given the considerable magnitude of the COVID-19 shock. In particular, the net changes could mask important gross job losses that might be revealing the depth of the recession. Therefore, a clear understanding of the underlying labour reallocation process is critical to not only assess the severity of the shock but also to draw implications for the potential recovery of the next months. Our results show the presence of high labour market churning that persists even after the easing of social-distancing restrictions. As of July, many of the recent job losers, especially those who had been temporarily laid-off between February and April, have regained employment. However, this apparent strong recovery dynamics hides important heterogeneity, and large groups of workers, such as those who were not employed prior to the pandemic, face important difficulties with finding a job. Three factors appear to be key in accounting for the incomplete employment recovery of July: (1) the unusually high separation flows that characterize the labour market in the reopening phase; (2) the low reemployment probability of recent job losers who were classified as out of the labour force during the lockdown; and (3), the low job-finding rate of individuals who were out of work prior to the pandemic. Our results further suggest that gross job losses were higher among women and young workers during the shutdown and that older workers were more likely to leave the labour force when the economy reopened.

The last chapter explores the gender differences in early career labour market trajectories and wage growth in Canada. Previous studies based on the European labour market have shown that the gender pay gap is relatively minor when entering the labour market, yet widens quickly in early career, over the first 5 to 10 years, which highlights the importance of examining and decomposing the differences in early career labour market trajectories. Existing literature in Canada has primarily focused on the gender wage gap itself; however, since the differences observed in the gender wage gap start minor but diverge fast in later years, labour market trajectories could be an important factor to consider when tackling this difference. In addition, factors associated with a job move, such as the type of separation and post-movement labour market trajectory, remain mostly unclear and could differ substantially. Individuals with different job paths could have very different wage growth. To close these gaps, this chapter examines and decomposes the gender differences in labour market trajectories for young workers in Canada and their potential impact on wage growth over the first ten years of their careers. My results show that labour market trajectories affect women's and men's weekly wage annual growth rates differently. Where women's weekly wage annual growth rate is more sensitive to temporary job separation compared with men, and in comparison, men's weekly wage annual growth rate is more sensitive to permanent job separation. However, regardless of the type of job separation, changed employer or occupation, and reason for separation, women's weekly wage annual growth rate was lower compared with their male counterparts. These results suggest that women in their early career faced a higher penalty for taking time out of the labour market due to permanent job separation and permanent separation due to parental/maternity leave compared with their male counterparts in Canada.

Chapter 1

The Effects of Education on Retirement Decisions and Change in Quality of Life after Retirement: Evidence from Canada

1.1 Abstract

While theoretical and empirical findings on the pecuniary return to education are well-established, there are not many studies that examine the causal link between levels of education and the long-term nonpecuniary outcomes such as retirement (Oreopoulos and Salvanes, 2011). Using the 2013 to 2016 General Social Survey, this paper investigates the effects of education on retirement decisions and change in the quality of life after retirement. The conceptual framework of this paper uses a Two-Stage Least Square (2SLS) estimation strategy, with provincial compulsory schooling laws as instruments to estimate the causal effects of educational attainment on individuals' retirement decisions and change in the quality of life after retirement. The regression results suggest that years of schooling have no significant impact on retirement decisions. Among those who are already retired, years of schooling do not have any impact on their age at retirement and the change in the quality of life after retirement. This means the change in educational attainment induced by compulsory schooling laws does not impact retirement decisions and the change in the quality of life after retirement. Results in this paper support the "fixed-retirement-age" assumption used in empirical studies based on the classical return to education model where the exogenous assumption on retirement age might not be a bad approximation (e.g., see Becker, 1967).

1.2 Introduction

The monetary return to education has been studied extensively both in Canada and internationally. Many researchers have shown that a higher level of education leads to higher lifetime income (Oreopoulos, 2006) and proactive saving behaviour (Messacar, 2017). When estimating the return to education, much of the existing literature assumes implicitly that individuals optimize over their career with a fixed year of retirement (Becker, 1967). In other words, the retirement decisions are assumed to be exogenous when estimating the classical return to education model (Card, 2001). This means a fixed retirement age has optimality conditions implying log-earning expressions that can be estimated by OLS or 2SLS. However, there is not a lot of evidence about whether or not education affects retirement. In theory, individuals with a higher level of educational attainment would have higher lifetime income and thus more money before retirement, which could lead to a shorter career. At the same time, a more enjoyable occupation as the result of a higher level of educational attainment might lead to later retirement, whereas a less enjoyable occupation might lead to the opposite. Therefore, whether education increases or decreases the retirement age is an empirical question. Besides education, a change in career length can also significantly change all sources of retirement income. Thus, it's not sufficient to solely rely on lifetime income and saving to explain the standard of living after retirement in the absence of the effect of education on retirement decisions (Bronshtein, Scott, Shoven, and Slavov, 2018).[1] Knowing the effect of education on retirement will not only inform us on the validity of the "fixed-retirement-age" assumption but is also important for policymakers when it comes to future policy adjustment such as the age requirement for the Old Age Pension plan in Canada (Oreopoulos and Salvanes, 2011).

This paper makes two contributions. First, using data from the General Social Survey (GSS) and the historical changes to the provincial compulsory schooling laws in Canada, I attempt to determine the causal relationship between educational attainment and individuals' retirement decisions. Depending on the results, this could lead to a structural change to the classical return to education model to include retirement decisions as an endogenous factor rather than assuming it is exogenous.[2] Second, I use changes in the provincial compulsory school laws to estimate the effect of educational attainment on the change of individuals' quality of life after retirement, which is an essential element of long-term nonpecuniary effects of education (Oreopoulos and Salvanes, 2011).

My Ordinary Least Squares (OLS) regression results suggest that educational attainment is not partially correlated with the probability of being retired at age 65. But among those who are retired, a higher level of educational attainment is associated with an older retirement age, which is an indication of delayed retirement. Moreover, the OLS regression results suggest that a higher level of educational attainment is associated with a lower level of life enjoyment after retirement. However, the OLS estimates are subject to the issue of endogeneity and cannot be interpreted as a causal effect. On the other hand, the 2SLS regression results suggest the effect of education on retirement decisions and change in the quality of life after retirement is small and not statistically significant. Thus from the perspective of this paper, the "fixed-retirement-age" assumption in empirical studies based on the classical return to education model seems valid, and the exogenous assumption on retirement age might not be a bad approximation. And, there are no long-term causal effects of education on quality of life after retirement in Canada. My current attempts to investigate the channels of the main empirical results show that total annual income is one of the potential channels that could lead to delayed retirement through education; however, due to endogeneity, it cannot be interpreted as a causal relationship, and there should be other channels that offset the effect of income, which in total sum up to a null effect of education on retirement decisions.

[1] Bronshtein, Scott, Shoven, and Slavov (2018)'s empirical results suggest that deferring retirement increases all sources of retirement income, whereas saving more only increases the relatively small contribution of annuitized defined contribution balances. This means that the saving adjustment required to achieve a particular increase in retirement income is larger the later in the career that the adjustment takes place. In other words, saving more gets less powerful as the career progresses, but deferring retirement remains equally powerful.

[2] If education does affect retirement age, then the education parameter in the classical Mincer equation would not necessarily capture the return to education.

The rest of the paper is organized as follows. Section 2 provides a literature review. Section 3 discusses the key conceptual framework. Section 4 provides a summary of the datasets and descriptive statistics. Section 5 discusses the identification strategy. Section 6 presents the empirical results, while Section 7 presents a robustness check of the empirical results. Section 8 showcase my current attempts to investigate the channels of the main empirical results. Section 9 discusses the main results and Section 10 concludes.

1.3 Literature Review

Oreopoulos (2006) presents a comprehensive study on the effect of schooling on income where he uses changes in Canadian schooling laws to explore the effects of schooling on income and to measure the lifetime opportunity costs of dropping out of school early. Essentially, Oreopoulos constructs a dataset that tracks the historical changes of Canadian compulsory schooling laws from 1900 to 2000 by provinces. He uses some elements of the laws, including minimum school-leaving age in combination with child labour laws, as instruments for educational attainment. Based on his estimations, he finds that the introduction of tighter provincial restrictions on leaving school between 1920 and 1990 raised average grade attainment and incomes. Students compelled to attend an extra year of school experienced an average increase in annual income of about 12%. He also finds that an increase in schooling is associated with significant benefits regarding other socioeconomic outcome measures ranging from bilingualism, employment, and poverty status.

There are not many existing studies that directly link the return to education and retirement decisions, perhaps due to the lack of relevant data or because it is traditionally assumed that retirement decisions (retirement dates more specifically) are fixed in the classical return to education model (Card, 2001). In recent years, researchers started to develop a new focus around endogenous educational attainment choice and retirement decisions. Bloom, Canning, and Moore (2014) consider a lifecycle model with endogenous retirement age. They find that the optimal retirement age is delayed because of mortality decline but is reduced by productivity increase. For instance, when combining these two factors and using parameter values calibrated to the U.S. economy, they find that the effect of productivity increase is rough twice that of mortality decline, resulting in earlier retirement over time. Restuccia and Vandenbroucke (2012) studies the effects of these two changes on another lifecycle variable: years of schooling. They consider a lifecycle model with endogenous schooling duration and leisure hours and find that optimal schooling duration rises over time because of either mortality decline or productivity increase. Their computational exercise suggests that the increase in productivity was the main cause of the increase in schooling from 1870 to 1970. An interesting observation is that Bloom, Canning, and Moore (2014) endogenize retirement age without examining issues related to schooling, but Restuccia and Vandenbroucke (2012) endogenize schooling years while assuming retirement age is exogenously fixed.

A similar approach has also been taken in many related papers in the literature. Kalemli-Ozcan and Weil (2010) and d'Albis, Lau, and Sánchez-Romero (2012) endogenize retirement age but not schooling years, as in Bloom, Canning, and Moore (2014). On the other hand, Heijdra and Romp (2009) and Cervellati and Sunde (2013) endogenize schooling years but not retirement age, as in Restuccia and Vandenbroucke (2012). The advantage of this approach can be seen from Heijdra and Romp (2009)'s statement that "by zooming in on one decision at a time, simple and intuitive analytical insights are much easier to come by."

While the above point regarding the use of a parsimonious model in scientific investigation is appropriate in many situations, Cai and Lau (2017) think that a more comprehensive model will be helpful in this context. They state that years of schooling and retirement age choices are likely to be interrelated, assuming one of these two lifecycle choices as exogenous may miss some important economic factors. Based on Hazan (2009), economic intuition suggests that the optimal choice of years of schooling depends on the duration that the individual can reap the return of human capital accumulation. Given these economic reasons, Cai and Lau (2017) include both schooling and retirement choices in their study and conclude that mortality decline is a necessary and sufficient condition to have a negative effect on retirement age when years of schooling is exogenous and is only a necessary condition when years of schooling is endogenous.

Behrman, Mitchell, Soo, and Bravo (2012) develop a new measure of financial literacy based on the Chilean Social Protection Survey and show that financial literacy is still positively and significantly associated with wealth outcomes and thus early retirement after controlling for schooling. Messacar (2017) studies the effect of education on retirement saving behaviour in Canada. He shows that overall, individuals with lower levels of education are found to save less for retirement than those with higher levels of education, which could lead to a gap in retirement decisions. Both Messacar (2017) and Behrman, Mitchell, Soo, and Bravo (2012)'s findings suggest that there are other channels through which years of schooling can affect retirement decisions. Rooij, Lusardi, and Alessie (2012) present new evidence on financial literacy and retirement preparation in the Netherlands based on two surveys conducted before and after the onset of the financial crisis. They document that while financial knowledge did not increase from 2005 to 2010, in 2010, significantly more individuals reported having thought about their retirement. Using the information on financial conditions and financial knowledge of relatives as instrumental variables (IV), they find a positive causal effect of financial literacy on retirement preparation. An important finding in Rooij, Lusardi, and Alessie (2012) is that the level of educational attainment does not directly affect individuals' retirement decisions and preparation. Instead, the level of understanding of financial literacy is what matters. However, it is difficult to completely separate the effect of education from the effect of financial literacy. Depending on the education system, these two could be correlated. In fact, in most Canadian provinces, introductory economics, banking and investment courses are offered as optional courses in high school, with some provinces starting to offer such courses in the early 90s.

Compared to the small number of existing studies on education and retirement decisions, education and overall quality of life are well examined. However, there are still some gaps in the literature regarding the effect of education on the quality of life among the retired population, especially on how the quality of life has changed compared to pre-retirement. Most of the existing literature focuses on the level of quality of life before or after retirement, which is not the same concept of the change in the quality of life presented in this paper, and therefore a direct comparison is not feasible. However, results from these studies can still help to understand the potential determinants of the level of quality of life that might also apply in the case of change in the quality of life after retirement. Markides and Martin (1979) build a causal model based on path analysis to estimate the direct and indirect effects of selected predictor variables, including education on life satisfaction among the elderly. Using survey data collected in the U.S., they find that health and activity emerge as strong predictors of life satisfaction while income influences life satisfaction indirectly via activity. Similarly, Branch-Allen and Jayachandran (2016) uses a holistic approach to construct a causal model to identify major determinants of life satisfaction in Canada. They find that individuals who are female, younger, married, from a high socioeconomic status background, born in Canada, very religious, and demonstrated a high level of neighbourhood interaction have greater satisfaction with life. Møller (1988) conducted a case study of retired Zulu contract workers in South Africa who returned to a rural lifestyle from urban cities after retirement. They find that quality of life after retirement is influenced mainly by health, perceived financial security, participation in social activity, and positive sentiment towards retirement life. Interestingly, most of these studies find that compared to the other determinants, the direct effect of education on quality of life is not as strong and significant.

While the above studies focus on identifying determinants of quality of life, Powdthavee, Lekfuangfu, and Wooden (2015) estimates a structural equation model using data from Australia to obtain the direct and indirect associations between education and life satisfaction. They find the estimated direct effect of education on life satisfaction to be negative and statistically significant, while the total indirect effect is positive, sizeable and statistically significant. They conclude that the use of single-equation models regarding the effect of education on quality of life might lead to misleading conclusions. Meeks and Murrell (2001) develop a model of relationships between enduring attributes and indicators of successful ageing and find that education and trait negative affect are both directly related to the quality of life among older adults. For instance, they find that higher educational attainment is related to lower levels of trait negative affect in which lower negative effect results in better health and life satisfaction.

Overall, existing papers in the literature conclude two main patterns on the relationship between years of schooling and retirement decisions. First, the combined effects of a decline in mortality and productivity increases lead to higher levels of education and earlier retirement over time (Restuccia and Vandenbroucke, 2012). Second, a rise in schooling duration can affect the marginal benefit of continuing working through a higher human capital level and thus wage income, which could lead to a delay in retirement (Bloom, Canning, and Moore, 2014). Other pecuniary and nonpecuniary outcomes of schooling could also affect retirement decisions such as financial literacy (Behrman, Mitchell, Soo, and Bravo, 2012) and saving behaviour (Messacar, 2017). As for the effect of education on quality of life and quality of life after retirement, existing papers in the literature present mixed evidence where some studies suggest a weak direct effect of education. However, some recent studies suggest a positive and significant effect, both directly and indirectly. It has been many years after the publication of Oreopoulos's paper in 2006, and individuals who were affected by changes in the compulsory schooling laws are getting close to retirement age. With the new release of General Social Survey (GSS) data, I further expand on current literature by using the identification strategy adopted from Oreopoulos (2006) to investigate the causal effect of educational attainment on retirement decisions and the change in the quality of life after retirement. This will also inform us of the potential validity of an important assumption in the classical return to education model, which assumes that the retirement decisions are exogenous.

1.4 Conceptual Framework

As Card (2001) summarized, the model of endogenous schooling can be illustrated in the framework of a simple model that builds on Becker (1967). In such a model, individuals face a market opportunity locus that gives the level of earnings associated with alternative schooling choices and reaches an optimal schooling decision by balancing the benefits of higher schooling (which are reaped over the lifecycle) against the costs (which are born early on). Issues arise when researchers try to apply OLS to estimate the return to schooling, which includes common endogeneity issues such as measurement errors and self-selection. Some of these issues can be resolved using different identification strategies such as 2SLS with IV, which is the main identification strategy adopted in this paper. Another issue on the estimate of the return to education is that much-existing literature assumes implicitly that individuals optimize over their career with a fixed year of retirement (Becker, 1967). Card (2001) summarized and outlined various challenges and econometric issues related to the current return to education model and identification strategies raised by endogenous schooling decisions, including the assumption that individuals have a "fixed-retirement-age". In other words, the retirement decisions are assumed implicitly to be exogenous when estimating the classical return to education model. Therefore, whether education increases or decreases the retirement age is an empirical question. However, there is not a lot of evidence about whether or not education affects retirement. If there is a causal relationship between education and retirement, then retirement decisions should be included in the model as an endogenous factor rather than assume it is exogenous.

Overall, existing literature such as Hazan (2009) and Bloom, Canning, and Moore (2014) has suggested that a higher level of educational attainment would likely cause a delay in retirement which contradicts the "fixed-retirement-age" assumption in empirical studies based on the classical return to education model. To further expand and test this assumption, I will investigate whether there is a causal effect of schooling on retirement decisions using a commonly used IV strategy. The instruments will be based on schooling laws previously used in Oreopoulos (2006) and Messacar (2017). The IV for educational attainment in this paper relies on the change in compulsory schooling laws. The compulsory schooling laws in Canada have existed for more than 100 years. Many policymakers believe that maintaining the compulsory school attendance laws will lead to a higher level of lifetime outcomes for individuals and society benefits collectively because a higher level of educational attainment promotes good citizenship and economic development (Oreopoulos, 2006). The compulsory schooling laws were first introduced in Ontario in 1871 to increase school attendance rate as well as a commitment to help low-income families' children to avoid poverty as adults. Appendix Figure A1 from Oreopoulos (2006) summarizes the compulsory schooling laws in Canada over the last 100 years, highlighting provincial requirements concerning the maximum age for beginning school and the minimum age for leaving/dropping school. Note that this is a figure directly from Oreopoulos (2006) which is not included in this study. The underlying details of these laws are complex, and there exist some exceptions. Ontario, for example, was the first Canadian province to introduce compulsory schooling. By 1970, all of the Canadian provinces enforced a minimum school-leaving age of either 15 or 16. These laws remained unchanged for many years. However, starting with Prince Edward Island in 1980, those provinces that had set the minimum leaving age at 15 raised it to 16. For instance, Newfoundland did so in

1987, Quebec in 1988, British Columbia in 1990, and Nova Scotia in 1996.

1.5 Data and Descriptive Statistics

The main dataset for this paper is based on the 2013-2016 Canadian General Social Survey (GSS).[3] There are a few advantages to using the GSS. First, unlike the Census, the GSS is conducted on an annual basis which provides the most up-to-date information and measurement on various socioeconomic indicators of Canadians. In the GSS, there is a set of harmonized contents/questions that are collected continuously despite different main topics in each year which includes various variables on education and retirement that are not included in the Census. Therefore, it is possible to append the GSS from different years into one single dataset to analyze education and retirement decisions. Second, variables on income such as individual income and family income are directly linked from the T1 and T4 administrative data files from the Canadian Revenue Agency (CRA), which is more accurate than the traditional survey method that is used in the early Census. And lastly, in the 2016 GSS, the survey includes new questions on various aspects of life after retirement, which are not collected in the previous cycles of GSS and the Census.

For this paper, selected harmonized contents on education and retirement decisions are appended into one dataset using the 2013-2016 GSS. New variables related to the change in the quality of life after retirement are also selected out from the 2016 GSS to investigate other additional aspects of the effect of education on retirement. These variables are discussed in further detail in the next section.

To replicate Oreopoulos (2006)'s identification strategy, data related to compulsory schooling laws in Canada are merged with the GSS dataset by the province of birth and birth cohort. Note that, unlike Oreopoulos (2006), the minimum school leaving age is the only IV that is used in this study. However, based on my conversation and the do files from the author, the minimum school leaving age is the main IV that Oreopoulos (2006) used for the regression analysis. In addition, the other I.V.s, such as total minimum instruction, did not match well with the GSS data and therefore are omitted in this study. Like Oreopoulos (2006), I restrict my sample to respondents who were age 14 between 1931 and 1975, those most likely to be affected by the compulsory schooling laws. Self-employed individuals are excluded since their retirement decisions could vary significantly different from the paid employees. In addition, I only kept individuals who were 50 years or older at the time of the survey in the dataset to focus on the outcome of retirement. Different from Oreopoulos (2006), I include immigrants in the later section to increase the sample size for the robustness test. After adjustment, the total number of respondents left in the merged dataset is 33,509 excluding immigrants and 41,738 including immigrants.

Within these individuals left in the datasets, not all variables are consistently observed. For instance, some variables may not be asked for all respondents due to survey flow (valid

[3]The target population of GSS includes all non-institutionalized persons 15 years of age and older, living in the ten provinces of Canada. Each year the main topic of the GSS change, and certain topics rotates and are re-collected after a period of 4 years. For this paper, I am using the 2013 to 2016 GSS, which includes topics on social identity, victimization, time use and a brand new topic on Canadians at Work and Home in 2016.

skips). For example, respondents are asked if they were paid employees only if they indicated that they were working full-time during the past 12 months. Respondents can refuse to answer certain questions (refusal) or state that they do not know the answers (don't know), which would be treated as missing values. In addition to the missing values originating from the GSS, data from Oreopoulos (2006) on cohort-varying controls also contain missing values for some provinces and certain cohorts. For instance, the fraction of the working population in the manufacturing sector in a province when respondents aged 14 is unavailable between 1972 and 1975 across nine provinces. Therefore, when the dataset is merged with the GSS data by birth province and cohort, additional missing cells are generated. After excluding respondents with missing values in the key variables for regression analysis, 28,844 respondents are left in my sample, excluding immigrants and 30,465 including immigrants. In the 2016 GSS alone, which is used to investigate individuals' post-retirement life quality, the total respondents left in the dataset after adjustment is 3,887 excluding immigrants and 4,070 including immigrants.

First, to replicate Oreopoulos (2006)'s results, data are collapsed based on various demographic variables based on Oreopoulos (2006)'s specification to calculate the equal province and cohort weight. These variables are years of schooling, age, a binary variable for individuals that completed high school, a binary variable for individuals who are bilingual, a binary variable for individuals that are unemployed, a binary variable for individuals who are married, a binary variable for individuals who are divorced, a binary variable for individuals who are self-employed, minimum school-leaving age and lastly, cohort-varying controls. Cell means are then generated based on the survey weights for all variables by survey year, birth cohort, and province. The rest of the results on retirement decisions are estimated at the individual level. Table A.1 contains the descriptive statistics for native Canadians aged 14 between 1931 and 1975 observed in the 2013 to 2016 Canadian GSS. Similarly, Table A.2 contains descriptive statistics for both native Canadians and non-natives. In Oreopoulos (2006), immigrants are excluded because they do not have a province of birth in Canada. In this paper, I assume that immigrants' current province of residence is the same as the province of birth to increase the sample size for the robustness test. Besides, immigrants make a large proportion of the Canadian population, which, in my opinion, are worth to be investigated. Note that in this case, only those who arrived in Canada at the age of 16 or younger are kept in the sample since the compulsory schooling instrument won't apply to those who are 17 or older. However, I try to match the GSS dataset to be as close as possible to Oreopoulos (2006)'s dataset from the Census, which covers individuals born between 1911 and 1961. The population cohorts in my sample end up being different due to the data constraint of a limited match by birth province and cohort between the GSS data and Oreopoulos (2006)'s a dataset on compulsory schooling laws. For example, there are very few individuals in the GSS that were born between 1911 and 1930, which resulted in some observations being dropped after the merge. In other words, the GSS represents a portion of the population cohort in Oreopoulos (2006), and some differences are expected.

As shown in Table A.1, the sample population distribution for the native-born Canadians from the 2013 to 2016 GSS looks quite similar to 1971 to 2001 Canadian Censuses with a few noticeable differences (see appendix table 1 Oreopoulos (2006)). For instance, individuals in the GSS have much a higher exposure to some post-secondary schooling (0.493) compared to the Censuses (0.322). Another noticeable difference is the inflation-adjusted log household income wherein the GSS individuals have a mean of 10.184 vs. 9.633 from the Censuses. Overall, the population characteristics in the 2013 to 2016 GSS are quite similar to 1971 to 2001 Canadian Censuses. Similarly, in Table A.2 which also contains the non-native individuals in the GSS, the population characteristics look identical to Oreopoulos (2006) with a higher mean percentage of some post-secondary schooling exposure and higher log household income after adjusting for inflation. In Table A.1, the differences between population-weighted counts and equal-province cohort weighted counts are minor, which is consistent with Oreopoulos (2006). Although the population cohorts are different from Oreopoulos (2006), the overall population characteristics are relatively close. With similar population characteristics and identification strategy, I expect to reach similar results to Oreopoulos (2006) on the effect of years of schooling to log total income through the compulsory schooling laws. Note that the measurements of the change in the quality of life after retirement are only available in the 2016 GSS and therefore do not have enough sample size to calculate the means using the equal province/cohort weights.

1.6 Estimation Strategy

First to replicate Oreopoulos (2006)'s results, the base OLS regression model is:

$$\begin{aligned} Log\ (INC)_{pct} &= \beta_0 + \beta_1 SCH_{pct} \\ &+ \beta_2 Age_{pct} + \beta_3 Age^2_{pct} + \beta_4 Age^3_{pct} + \beta_5 Age^4_{pct} \\ &+ \beta_6 Female_{pct} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{pct} \end{aligned} \tag{1.1}$$

Where INC_{pct} is the average total annual income of individuals born in province p from birth cohort c and survey year t. SCH_{pct} is the average years of schooling for the individuals born in the province p from birth cohort c observed in survey year t. π_p, υ_c and ϕ_t are fixed effects for the province, birth cohort, and survey year, respectively. Age_{pct} is the average age for the group born in the province p from birth cohort c and survey year t. $Female_{pct}$ here is the proportion of females for the group born in province p from birth cohort c and survey year t. P_{pc} is a vector of cohort-varying controls from Oreopoulos (2006) which includes the fraction of population living in rural towns, the fraction of population working in the manufacturing sector, the per-student school expenditure, the number of secondary schools and the number of teachers. These controls are added in the model to control for observable factors that might affect individuals' years of schooling other than the compulsory schooling laws. ϵ_{pct} is the error term which is clustered at province and cohort level as in Oreopoulos (2006).

Similar to Oreopoulos (2006), the first and second stage models are the following:

$$\begin{aligned} SCH_{pct} &= \gamma_0 + \gamma_1 MSL_{pc} \\ &+ \gamma_2 Age_{pct} + \gamma_3 Age^2_{pct} + \gamma_4 Age^3_{pct} + \gamma_5 Age^4_{pct} \\ &+ \gamma_6 Female_{pct} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{pct} \end{aligned} \tag{1.2}$$

$$\begin{aligned} Log\ (INC)_{pct} &= \beta_0 + \beta_1 \widehat{SCH}_{pct} \\ &+ \beta_2 Age_{pct} + \beta_3 Age^2_{pct} + \beta_4 Age^3_{pct} + \beta_5 Age^4_{pct} \\ &+ \beta_6 Female_{pct} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{pct} \end{aligned} \tag{1.3}$$

Where MSL_{pc} is the minimum schooling leaving age when individuals are aged 14. $\widehat{SCH}_{pct}$ is the group's predicted years of schooling after estimating the first-stage equation. The regression analysis on log total annual income is estimated using the equal-province cohort weight as in Oreopoulos (2006).

Next, adpated to Oreopoulos (2006), the base OLS regression model to estimate the effect of years of schooling on retirement decisions and the change in quality of life after retirement is:

$$\begin{aligned} RET_{it} &= \beta_0 + \beta_1 SCH_{it} \\ &+ \beta_2 Age^2_{it} + \beta_3 Age^3_{it} + \beta_4 Age^4_{it} \\ &+ \beta_5 Female_{it} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{it} \end{aligned} \tag{1.4}$$

Note the analysis here on retirement outcomes is conducted at the individual level instead of the cell level. Due to the improved computing power and statistical software, it's not necessary to collapse the data to cell level like Oreopoulos (2006). RET_{it} are the retirement outcome which includes the binary measurement on retirement status and retirement age for those who are already retired observed in survey year t.[4] For the change in the quality of life after retirement, two binary variables are chosen from the 2016 GSS, which are the enjoyment of life before and after retirement and the financial standard of living after retirement. Both binary variables are assigned the value of one if the situation is better compared to pre-retirement and zero otherwise.[5] SCH_{it} is individuals' total years of schooling observed in survey year t. $Female_{it}$ here is the binary measurement on sex at birth for individual i observed in survey year t. ϵ_{it} is the error term which is clustered at province and cohort level as in Oreopoulos (2006). Note that Age_{it} is excluded from the model due to collinearity with the cohort fixed effects.

[4]Retirement status here is defined using the individuals' main activity during the past 12 months which contains the following options: 1. Working at a paid job or self-employed, 2. Looking for paid work, 3. Going to school, 4. Caring for children, 5. Household work, 6. Retired, 7. Maternity/Paternity or parental leave, 8. Long-term illness, 9. Volunteering or caregiving other than for children, and 10. Other. Retirement status is assigned a value of one of the respondents selected 6. Retired as their main activity in the past 12 months and zero otherwise. Retirement age is derived using the birth year and year of the most recent retirement. Note for individuals who retired more than once, only the most recent retirement instance is accounted to derive the age at retirement.

[5]Question on the enjoyment of life before and after retirement is the following: Compared to the year before you most recently retired, do you now enjoy life more, less, or about the same? Similarly, the question on the financial standard of living after retirement is the following: Compared to your expectations before you most recently retired, how would you describe your financial standard of living?

The first and second stage models are the following:

$$\begin{aligned} SCH_{it} = \gamma_0 &+ \gamma_1 MSL_{pc} \\ &+ \gamma_2 Age_{it}^2 + \gamma_3 Age_{it}^3 + \gamma_4 Age_{it}^4 \\ &+ \gamma_5 Female_{it} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{it} \end{aligned} \tag{1.5}$$

$$\begin{aligned} RET_{it} = \beta_0 &+ \beta_1 \widehat{SCH}_{it} \\ &+ \beta_2 Age_{it}^2 + \beta_3 Age_{it}^3 + \beta_4 Age_{it}^4 \\ &+ \beta_5 Female_{it} + P_{pc} + \pi_p + \upsilon_c + \phi_t + \epsilon_{it} \end{aligned} \tag{1.6}$$

Where $\widehat{SCH}_{it}$ is the individuals' predicted years of schooling after estimating the first-stage equation. The regression analysis on retirement outcomes is estimated using the survey weights, which are calibrated to the true population from the 2011 Census.

1.7 Results

Replication of Oreopoulos, 2006, Years of Schooling and Total Annual Income

First, I attempt to replicate Oreopoulos (2006)'s results using the GSS data. The purpose of the replication is to test whether or not the identification strategy from Oreopoulos (2006) would work with the GSS data. Unlike the Census, some observations were dropped after the data merge between the GSS and the data on compulsory schooling laws. Therefore, I first replicate Oreopoulos (2006) to help me to verify if the final data set is still valid for the same identification strategy, which has been proven to be very good in past literature. Table A.3 contains the first stage regression results. Among all three specifications, the effects of the compulsory schooling laws have positive signs, which are consistent with the results from (Oreopoulos, 2006). In the first specification where there are no other control variables, the compulsory schooling laws (minimum school-leaving age) indicate that for each additional year in school as required by the compulsory schooling laws, individuals' years of schooling increases on average by 0.688 years (about seven months). This result is statistically significant at the 1% level with an F-stat of 137.71. Similarly, in the second and third specification with cohort-varying controls and fixed effect, the compulsory schooling laws (minimum school-leaving age) has coefficient estimates of 0.472 and 0.411, respectively. Similar to the results from the first specification, the results in the second and third specifications are statistically significant at the 1% level with an F-stat of 110.04 and 66.70, respectively. Overall the results from the first stage are consistent with the results from Oreopoulos (2006).[6]

Table A.4 contains the second stage results of years of schooling on annual log income. Among all three specifications, the effects of the predicted years of schooling have positive signs, which are consistent with the results from Oreopoulos (2006). In the first specification, for each extra year of schooling, the average annual income increases by 9.2%. This result is statistically significant at the 1% level. In the second and third specifications with

[6] In Oreopoulos (2006), the coefficient estimate of minimum school-leaving is 0.130, and it is statistically significant at the 1% level with an F-stat of 70.47 in the full specification model.

cohort-varying controls and fixed effect, the estimates on years of schooling remain statistically significant at 5% and 10% level with coefficient estimates of 8.2% and 4.5% respectively. Compared to Oreopoulos (2006), the significance level and the magnitude of coefficient estimates here in the second stage are not as significant and strong. In Oreopoulos (2006), the coefficient estimate of years of schooling on annual income in the second stage is 10.3%, and it is statistically significant at the 1% level. Table A.5 contains the OLS regression results of years of schooling on annual log income. In the third specification with cohort-varying controls and fixed effect, the coefficient estimate suggests an increase of 7.4% on average annual income when years of schooling increase by one year. Compared to the 2SLS regression results, the OLS regression results are more substantial in magnitude. However, when compared to the OLS regression results of Oreopoulos (2006), the results here are much smaller in magnitude. For instance, in Oreopoulos (2006), the coefficient estimate of years of schooling on annual income in OLS regression is 12%, and it is statistically significant at the 1% level.

Overall, the coefficient estimates of years of schooling on annual income are not as significant and substantial in magnitude compared to Oreopoulos (2006), perhaps because I am looking at older individuals compared to Oreopoulos (2006) which focus on individuals at their primary working age. At large, the results in the second stage are consistent with the findings from Oreopoulos (2006), which suggests that there is a positive causal relationship between years of schooling and income.

Years of Schooling and Retirement Status

I have shown in the previous section that my estimated returns to education are in line with (Oreopoulos, 2006). In this section, I explore the possible existence of a causal relationship between years of schooling and retirement decisions. To start, I first look at the retirement status, which is measured through individuals' main activity during the reference months. Note unlike in the previous section, the following analysis on retirement decisions is conducted at the individual level instead of the cell level. Table A.6 contains the first stage results of the compulsory schooling laws in equation (5) on years of schooling. Among all three specifications, the effects of the compulsory schooling laws have positive signs, which are consistent with the results from (Oreopoulos, 2006). In the third specification with cohort-varying controls and fixed effect, the coefficient estimate of compulsory schooling laws (minimum school-leaving age) suggests for each additional year in school as required by the compulsory schooling laws, there is an approximately additional four months increase in individual's years of schooling. This result is statistically significant at the 1% level with an F-stat of 30.83.

Table A.7 contains the second stage results for the effect of years of schooling on retirement status. In the first specification, where there are no other control variables, for each extra year of schooling predicted through the compulsory schooling laws, the probability of being retired decreases by 14.7% percentage points. This result is statistically significant at the 5% level. The results in the second specification are similar but with a smaller magnitude (about 3% percentage points). In the third specification, however, the sign of the coefficient estimate on years of schooling flipped to positive, and it becomes statistically insignificant. This indicates that when cohort-varying controls and fixed effect are taken into consideration, there is no evidence of a causal relationship between years of schooling and retirement status.[7] Table A.8 contains the OLS regression results of years of schooling on retirement status. Compared to the 2SLS regression results, the OLS coefficient estimates of years of schooling on retirement status show a similar pattern.

Years of Schooling and Retirement Age

In the previous section, I concluded that there is no evidence suggesting a causal relationship between years of schooling and retirement status. In this section, I go a step further to explore the possible existence of a causal relationship between years of schooling and retirement age among retired individuals. Retirement age, in this case, is derived from the last year where individuals reported working full-time and their birth year. Note, only individuals who reported that they are retired were asked the year where they last worked full-time; therefore, the sample population here is a subset of the sample population on retirement status. Table A.9 contains the first stage results of the compulsory schooling laws on years of schooling in equation (5). Similar to the first stage regression results shown above, the coefficient estimate in the third specification is statistically significant at the 1% level with an F-stat of 20.52.

Table A.10 contains the second stage results for the effect of years of schooling on retirement age. Among all three specifications, the signs of the coefficient estimates suggest a positive relationship between years of schooling and retirement age. However, they are not statistically significant. This indicates that there is no evidence of a causal relationship between years of schooling and retirement age. [8] Table A.11 contains the OLS regression results of years of schooling on retirement age. Unlike the 2SLS regression results and the OLS regression results for retirement status, the coefficient estimates from all three specifications are statistically significant. For instance, the coefficient estimate of years of schooling in the third specification suggests an additional increase of 3.8 months (coefficient estimate of 0.32) on age at retirement when years of schooling increase by one year. However, the OLS regression results are not sufficient to conclude a causal relationship between years of schooling and retirement age due to issues of endogeneity.

Overall, based on the 2SLS regression results, I fail to reject the null hypothesis that years of schooling have no causal impact on retirement decisions (Neither on the probability of being retired nor on the retirement age among those who are retired). Therefore, the assumption on "fixed-retirement-age" (Becker, 1967; Card, 2001) seems valid. From the perspective of the long-term nonpecuniary outcome, educational attainment does not seem to impact individuals' long-term retirement decisions.

[7] I also rerun the second stage model using the logit model, and the results are consistent with the main results here using the linear probability model. Detail results and tables are available upon request.

[8] Similar to retirement status, I also rerun the second stage model on retirement age using the logit model, and the results are consistent with the main results using the linear probability model. Detail results and tables are available upon request.

Years of Schooling and Change in the Quality of Life after Retirement

Using the new content introduced in the 2016 GSS, in this section, I attempt to explore the possible existence of a causal relationship between years of schooling and change in the quality of life after retirement among those who are retired. Since the content was only available in 2016, the sample population here is a subset of the sample population on retirement status where the year of observation is restricted to 2016. Table A.20 contains the first stage results of the compulsory schooling laws on years of schooling in equation (5). Compared to the first stage regression results shown above, the coefficient estimate in the third specification is only statistically significant at the 10% level with a smaller F-stat of 16.21. Perhaps due to the fact here, I am only looking at one year of observation which is much smaller than the number of observations in the primary dataset.

Table A.21 contains the second stage results for the effect of years of schooling on the change in the enjoyment of life before and after retirement. Among all three specifications, the signs of the coefficient estimates suggest a negative relationship between years of schooling and the likelihood of enjoying life more after retirement. However, in the second and third specifications, the coefficient estimates are not statistically significant. This indicates that when the control variables and fixed effects are taken into consideration, there is no evidence of a causal relationship between years of schooling and whether life is more enjoyable or not after retirement.[9] Table A.22 contains the OLS regression results of years of schooling on the enjoyment of life before and after retirement. Unlike the 2SLS regression results, the coefficient estimates from all three specifications are statistically significant at 1%. For instance, the coefficient estimate of years of schooling in the third specification suggests an additional decrease of 1.9 percentage points on the likelihood of enjoying life more after retirement when years of schooling increase by one year. However, the OLS regression results are not sufficient to conclude a causal relationship due to issues of endogeneity.

Table A.23 contains the second stage results for the effect of years of schooling on the change of financial standard of living after retirement. Among all three specifications, the signs of the coefficient estimates suggest a negative relationship between years of schooling and the likelihood of having a good financial standard of living after retirement. However, the magnitude of the coefficient estimates are close to zero and are not statistically significant. This indicates that there is no evidence of a causal relationship between years of schooling and having a good financial standard of living after retirement.[10] Table A.24 contains the OLS regression results of years of schooling on the financial standard of living after retirement. Similar to the 2SLS regression results, the coefficient estimates from all three specifications are not statistically significant and with a small magnitude close to zero.

[9] I also rerun the second stage model on the change in the enjoyment of life before and after retirement using the logit model, however in this case the model failed to converge due to small sample. Detail results and tables are available upon request.

[10] Similar to the change in the enjoyment of life after retirement discussed above, I rerun the second stage model here on the change of financial standard of living after retirement using the logit model, however in this case the model failed to converge due to small sample. Detail results and tables are available upon request.

1.8 Robustness Check

First, to check the robustness of the regression results on retirement decisions, I re-estimate the first and second stage regressions across three different age groups to see if the null effect of education on retirement status and retirement age stays consistent. The selected age groups are age 50 to 70, age 55 to 65, and age 60 to 65, which are the common age ranges where individuals in Canada retire. Note that these results are imprecisely estimated and serve as an extra step to ensure the null effect of education on retirement decisions and change in the quality of life after retirement shown in the main results section. The results in the first stage regressions for all age groups pass the 'relevance' test with an f-stat of excluded instruments larger than 20. Table A.12 contains the second stage results for the effect of years of schooling on retirement status across the three age groups. Among all three age groups, the coefficient estimates of years of schooling stay positive but are not statistically significant. Similarly, As shown in Table A.13 the coefficient estimates of years of schooling on retirement age are not statistically significant across all three age groups. These results suggest that the null effect of education on retirement status and retirement age is robust and is not affected by choice of sample based on individual's age groups. Next, I expand the size of the sample by including immigrants in the dataset to check the robustness of the regression results on retirement decisions and the change in the quality of life after retirement. Note that in this case, only those who arrived in Canada at the age of 16 or younger are kept in the sample since the compulsory schooling instrument won't apply to those who are 17 or older.

In Oreopoulos (2006), immigrants are excluded from the dataset due to two main reasons. First, since the compulsory schooling laws were regulated on provinces and year individuals at age 14, a key assumption that Oreopoulos (2006) had to make is that before the age of 16, individuals were assumed to reside in the same province as their province of birth, which is uncertain and less likely for immigrants. Also, it was challenging to identify immigrants who arrived at age 16 or younger, which otherwise would be out of scope for the compulsory schooling laws. In this paper, for the robustness test, I include immigrants who arrived in Canada at the age of 16 or younger based on immigrant status and the year where they first entered Canada.[11]In the following section, for immigrants who were not born in Canada, the data from Oreopoulos (2006) on compulsory schooling laws are linked to GSS using province of residence (at the time of the survey) and year when respondent aged 14. One important assumption I have to make is for these individuals, their province of residence at the time of survey is the same as the province of residence at age 16 or younger.

Years of Schooling and Retirement Status, Including Immigrants

Table A.14 contains the first stage results of the compulsory schooling laws in equation (5) on years of schooling, including immigrants. Among all three specifications, the effects of the compulsory schooling laws have positive signs, which are consistent with the regression results above without immigrants. In the third specification with cohort-varying controls and fixed effect, the coefficient estimate of compulsory schooling laws (minimum school-leaving age) suggests an approximately additional 4.3 months increases in individual's years of schooling

[11]Although the analysis with immigrants here relies on strong assumptions, however, it can still provide valuable information given that a large portion of the sample in the GSS are immigrants.

when compulsory schooling laws increases by one year. This result is statistically significant at the 1% level with an F-stat of 29.75. Table A.15 contains the second stage results for the effect of years of schooling on retirement status, including immigrants. In the third specification, the sign of estimates on years of schooling flipped to positive, and it becomes statistically insignificant. This result is consistent with the main second stage regression results on retirement status without immigrants. In addition, the OLS regression results in Table A.16 also show a similar pattern to the results without immigrants.

Years of Schooling and Retirement Age, Including Immigrants

Table A.17 contains the first stage results of the compulsory schooling laws in equation (5) on years of schooling, including immigrants for those who are already retired. Among all three specifications, the effects of the compulsory schooling laws have positive signs, which are consistent with the regression results above without immigrants. In the third specification with cohort-varying controls and fixed effect, the coefficient estimate of compulsory schooling laws (minimum school-leaving age) suggests an approximately additional 4.5 months increases in individual's years of schooling when compulsory schooling laws increases by one year. This result is statistically significant at the 1% level with an F-stat of 20.38. Table A.18 contains the second stage results for the effect of years of schooling on retirement age, including immigrants. Like the main second stage regression results without immigrants, the coefficient estimates from all three specifications have a positive sign but are not statistically significant.

Years of Schooling and the Change in Quality of Life after Retirement, Including Immigrants

Table A.25 contains the first stage results of the compulsory schooling laws in equation (5) on years of schooling, including immigrants in the 2016 GSS. Among all three specifications, the effects of the compulsory schooling laws have positive signs, which are consistent with the regression results above without immigrants. In the third specification with cohort-varying controls and fixed effect, the coefficient estimate of compulsory schooling laws (minimum school-leaving age) suggests an approximately additional 2.5 months increase in an individual's years of schooling when compulsory schooling laws increases by one year. This result is statistically significant at the 10% level with an F-stat of 19.20. Table A.26 and Table A.28 contain the second stage results for the effect of years of schooling on the enjoyment of life and financial standard of living after retirement, including immigrants. Like the main second stage regression results without immigrants, the coefficient estimates from all three specifications have a negative sign but are not statistically significant. The OLS regression results in Table A.29 also show a similar pattern to the results without immigrants.

Overall, the 2SLS regression results including immigrants on both retirement decisions and the change in the quality of life after retirement do not change the original conclusion. This indicates that the main 2SLS regression results are robust.

1.9 Investigation of channels of the Null Effect of Education on Retirement Decisions

In this section, I attempt to explore the possible channels to investigate what is behind the null effect of education on retirement, which my main results suggest. First, I selected various socioeconomic factors from the GSS that could impact retirement decisions through education which includes: occupation, marital status, general health condition, mental health condition, life satisfaction, and log total annual income.[12] next, I use the same 2SLS identification strategy to estimate whether education has any causal impact on these factors. If there is evidence that suggests a causal link, using a standard OLS, I then estimate the factors' effect on retirement decisions (retirement status and retirement age). Table A.30 contains the second stage results for the effect of years of schooling on the first five occupation categories, including management, administration, science, health and education. As the table suggests, among all five occupations, the coefficient estimates are not statistically significant, with a weak F-stat of excluded instruments from the first stage regression of around 5.90. Similarly, in Table A.31, the coefficient estimates for the last five occupation categories, including arts, sales, trades, agriculture and manufacturing, are not statistically significant. This suggests that for those individuals that are currently working in the sample, years of schooling do not have any causal impact on their probability of being working in any specific occupations. Therefore, from the perspective of this paper, occupation is not one of the channels that could impact retirement decisions (retirement status in this case) through education. One drawback here is that the GSS does not collect the occupation information for those who are already retired, and therefore, I cannot estimate whether or not occupation has any impact on their timing of retirement (retirement age) through education. Next, Table A.32 contains the second stage results for the effect of years of schooling on the other selected socioeconomic factors, including marital status, general health condition, mental health condition, overall life satisfaction and lastly log total annual income.[13] Among all five outcome variables, only log total annual income have a statistically significant coefficient estimate at the 10% level, which is consistent with the results in the previous section of replication of Oreopoulos (2006) using the collapsed sample. This suggests that among all the selected socioeconomic factors in this

[12]Note The occupation question in the GSS is only asked individuals who identified themselves as currently working or worked anytime during the past 12 months at the time of the survey. Therefore for those who are already retired, we do not have any information regarding their occupation before retirement. The occupation variable includes ten categories which are based on the 2016 National Occupational Classification; the list includes 1. management occupations, 2. business, finance, and administration occupations, 3. natural and applied sciences and related occupations, 4. health occupations, 5. occupations in education, law and social, community and government service, 6. occupations in art, culture, recreation and sport, 7. sales and service occupations, 8. trades, transport and equipment operators and related occupations, 9. natural resources, agriculture and related production occupations and 10. occupations in manufacturing and utilities.

[13]Marital status include the following categories: 1. married, 2. living common-law, 3. widowed, 4. separated, 5. divorced and 6. single, never married. For the regression analysis in this section, both married and living common-law are selected to construct the binary marital status variable. Both the general health condition and mental health condition variables are self-rated questions that consist of the following categories: 1. excellent, 2. very good, 3. good, 4. fair, and 5. poor. The first three answers are used here to construct the binary variables for the general health and mental health condition. Variable on life satisfaction is measured on a scale between 0 and 10, and the original question asks the individuals to rank their feelings about life as a whole from 0 very dissatisfied to 19 very satisfied.

paper, total annual income is the only channel that has an impact on retirement decisions through education. The coefficient estimate of 0.031 indicates that one extra year of total years of schooling increases the total annual income by 3.10%, which translates to a coefficient estimate of 6.20% less likely to be retired and an extra 1.58 years increase in retirement age for those who are retired; as shown in Table A.33. From the perspective of income and substitution effects, the results in this section suggest a dominant substitution effect of education through total income on retirement decisions. This means that, on average, individuals chose to work longer to obtain the extra amount of income as a result of their educational attainment rather than the feeling of having enough savings to retire early to enjoy leisure. Although it cannot be interpreted as a causal relationship due to endogeneity that can cause by potential omitted variable bias, these results seem to suggest a delayed timing of retirement as a result of higher income through a higher level of education. However, this conflicts with my main findings, which suggest a null effect of education on retirement decisions. Therefore, there should be other channels that offset the effect of income that is yet to be discovered, perhaps family composition and degree of community involvement which might be affected by total years of schooling that could lead to early retirement. Additional studies are needed to continue the discussion.

1.10 Discussion

Unlike most of the relevant existing literature suggests (such as Cai and Lau, 2017 and Hazan, 2009), results from this paper suggest no causal relationship between education and retirement decisions. In theory, individuals with a higher level of educational attainment would have higher lifetime income and thus more money before retirement, which could lead to a shorter career. At the same time, a more enjoyable occupation as the result of a higher level of educational attainment might lead to later retirement, whereas a less enjoyable occupation might lead to the opposite. My results suggest that, in total, the two forces which work on retirement decisions in the opposite direction even out. One explanation could be that perhaps when making an optimal choice on educational attainment, individuals already formed their expectations on retirement through other channels such as expected lifetime income and savings. Thus, when educational attainment is reflected as years of schooling, it has no significant impact on their retirement decisions.

Results from this paper suggest there is no causal relationship between education and the change in the quality of life after retirement. A direct comparison of this result to the existing literature is difficult as most of them focus on the effect of education on the level of quality of life after retirement (e.g. such as Møller, 1988, Branch-Allen and Jayachandran, 2016 and Powdthavee, Lekfuangfu, and Wooden, 2015). That being said, interestingly enough, studies that are done based on data from developed countries seem to find education to be less critical in post-retirement life quality. Perhaps in developed countries like Canada, where social welfare policy is fully established and well implemented, the effect of education on the change of overall quality of life after retirement is likely smoothed out and instead is reflected through other channels such as physical and mental health conditions. Additional study is required to expand the discussion further.

There is still much work to be done on the investigation of channels for the main empirical results. My current attempts show that total annual income could lead to a delayed retirement through education, but the effect can become statistically insignificant once I adapt to a different model and estimation strategy that are free from endogeneity. Based on the main findings in this paper, there could be other channels that offset the effect of income, which in total sum up to a null effect of education on retirement decisions. It is also worth mentioning that education could affect mortality and, therefore, the sample of 'survivors' which is analyzed here. However, additional data would be required to investigate the potential effect of education on longevity (or mortality) and its implications for the interpretation of the findings in this study.

1.11 Conclusion

Using the 2013 to 2016 General Social Survey, this paper attempts to estimate the causal effect of education on individuals' retirement decisions and the change in the quality of life after retirement. The OLS regression results suggest that for those individuals that experienced changes in the compulsory schooling law between 1931 and 1975, educational attainment is not partially correlated with the probability of being retired at age 65. But among those who are retired, a higher level of educational attainment is associated with an older retirement age, which is an indication of delayed retirement. Moreover, the OLS regression results suggest that a higher level of educational attainment is associated with a lower level of life enjoyment after retirement. However, the OLS estimates are subject to the issue of endogeneity and cannot be interpreted as a causal effect. On the other hand, the 2SLS regression results suggest that for individuals who gained extra years of total schooling through changes in the compulsory schooling law between 1931 and 1975, the effect of education on retirement status, retirement age, enjoyment of life after retirement and financial standard of living after retirement is weak and insignificant. This means the change in educational attainment induced by compulsory schooling laws does not impact retirement decisions. Thus from the perspective of this paper, the "fixed-retirement-age" assumption in empirical studies based on the classical return to education model seems valid, and the exogenous assumption on retirement age might not be a bad approximation. Lastly, there is no evidence that suggests any long-term causal effects of education on the change in the quality of life after retirement in Canada. My current attempts to investigate the channels of the main empirical results show that total annual income is one of the potential channels that could lead to delayed retirement through education; however, due to endogeneity, it cannot be interpreted as a causal relationship, and there should be other channels that offset the effect of income, which in total sum up to a null effect of education on retirement decisions.

Chapter 2

Labour Market Flows and Worker Trajectories in Canada During COVID-19

2.1 Abstract

We use the confidential-use files of the Labour Force Survey (LFS) to study the employment dynamics in Canada from the beginning of the COVID-19 pandemic in 2020 through to mid-summer. Using the longitudinal dimension of this dataset, we measure the size of worker reallocation, and document the presence of high labour market churning, that persists even after the easing of social-distancing restrictions. As of July 2020, many of the recent job losers – especially those who had been temporarily laid-off between February and April – have regained employment. However, this apparent strong recovery dynamics hides important heterogeneity, and large groups of workers, such as those who were not employed prior to the pandemic, face important difficulties with finding a job. Three factors appear to be key in accounting for the incomplete employment recovery of July: (1) the unusually high separation flows that characterize the labour market in the reopening phase; (2) the low reemployment probability of recent job losers who were classified as out of the labour force during the lockdown; and (3), the low job-finding rate of individuals who were out of work prior to the pandemic. Our results further suggest that gross job losses were higher among women and young workers during the shutdown, and that older workers were more likely to leave the labour force when the economy reopened.

2.2 Introduction

The Canadian labour market underwent an unprecedented trajectory since the onset of the COVID-19 pandemic. The massive job losses due to the shutdown of the early phase of the outbreak were followed by a vigorous rebound upon the gradual reopening of the economy, as was the case for many countries. As of July, employment was 1.3 million below its pre-COVID level, despite 3 million jobs lost during the spring (Canada, 2020). This unusual trajectory has been extensively analyzed (e.g., Lemieux, Milligan, Schirle, and Skuterud, 2020; Jones, Lange, Riddell, and Warman, 2020; Canada, 2020). However, the discussions have largely focused on the behaviour of the economy in *net* terms, and little is known about the employment flow dynamics that have accompanied the virus propagation. Although highly informative about the impact of COVID-19, these net changes are likely to hide sizable worker reallocation flows across labour-market states and jobs, especially given the considerable magnitude of the COVID-19 shock. In particular, the net changes could mask important gross job losses that might be revealing of the depth of the recession. Therefore, a clear understanding of the underlying labour reallocation process is critical to not only assess the severity of the shock, but also to draw implications for the potential recovery of the next months.

This paper uses the confidential-use files of the Labour Force Survey (LFS) to study the dynamics of employment in Canada from the early onset of COVID-19 to mid-summer.[1] This paper is, to our knowledge, the first to analyze the impact of the pandemic using these data.[2] We propose a novel characterization of the trajectory of the Canadian economy in COVID-19 times, based on a detailed analysis of the worker reallocation process of the pandemic months. Our analysis takes advantage of the longitudinal dimension of the confidential version of the LFS.[3] Specifically, this dataset is the only one currently available that allows us to build panels following workers' trajectories for six consecutive months over the COVID period. More precisely, these panels can span the month prior to the start of the outbreak (and the ensuing shutdown) through to mid-summer, after the lifting of most of social-distancing restrictions. Another key advantage of this data is that it allows us to examine the evolution of non-response in the LFS since the beginning of the pandemic.[4] The virus has severely constrained the data collection activities of statistical agencies, resulting in unusually high (survey) non-response rates in many countries. This raises concerns about the reliability of COVID-time statistics.[5] Evaluating the scale of the non-response problem, and most importantly, its implications for labour-force estimates, is critical for assessing the impact of COVID-19. This paper, by documenting the evolution of this phenomenon in Canada using the confidential-use LFS, takes a step in this direction.[6]

In our analysis, we measure the size of gross worker flows in and out of employment since the start of the outbreak. This is necessary for an informed evaluation of the magnitude of the COVID-19 shock: since the gross employment outflows are partially offset by the inflows, the net employment changes, albeit dramatic, can only give a lower bound on the number of jobs lost during the shutdown. Gross flows are also key for assessing the strength of the recovery, as the high net employment gains of the reopening might hide large gross

[1] Specifically, we use Statistics Canada's internal-use files. These restricted-access/confidential-use LFS files are, to the best of our knowledge, very similar (if not identical) to the confidential-use LFS files that are made available to researchers in the Research Data Centres (RDCs). In both cases, one can take advantage of the longitudinal dimension of the LFS data, a key requirement of our analysis.

[2] Exceptions are two short Statistics Canada papers (i.e. Chang, Morissette, and H. Qiu, 2020; Hou, Picot, and J. Zhang, 2020), that are part of the "COVID-19 Data insight for better Canada" series, meant to provide timely insight on the current crisis. Additional details are provided in the literature review below.

[3] The variables needed to construct panels in the LFS are not made available in the public-use files. See the data section for more details.

[4] Specifically the information on when a person enters the LFS and whether her/his labour market information was fully imputed is suppressed in the public-use files.

[5] See U.S. Bureau of Labor Statistics (2020b) for the U.S. case or https://ilostat.ilo.org/topics/covid-19/covid-19-impact-on-labour-market-statistics for a more general discussion.

[6] Another difference with the public-use files of the LFS is that the confidential files provide 4 digit National Occupation Classification (NOC) codes, versus up to 2 digits for the public version. Given the important focus of the literature on the effect of COVID-19 across occupations (e.g. Béland, Brodeur, Mikola, and Wright (2020), Gallacher and Hossain (2020)), this might be a key advantage. This paper makes relatively limited use of the detailed NOC codes, but further research could take better advantage of this feature of the confidential-use LFS files for a thorough analysis of the distributional impact of the virus, and its implications for labour reallocation across occupations.

outflows. We find that along the pandemic path, the economy experienced considerable gross job losses: the 20 to 64 year-olds lost 1.3 million jobs between February and March, and a further 2.3 million between March and April. Job losses remained persistently high even after the reopening: between April and May, 1 million workers 20 to 64 years of age lost their jobs. In comparison, we find that the cumulative monthly employment outflows between February and May 2009 (i.e. during the Great Recession) were around 1.6 million for the same population. These figures are indicative of the remarkable severity of the shock and also reveal high excess reallocation flows or high labour market "churn", which suggests unusual worker mobility patterns.[7] Understanding the consequences of these patterns for the upcoming recovery requires further analysis of the COVID-19 employment dynamics.

Thus, our paper examines the composition of the COVID employment flows. Although an unusually large number of early-pandemic job losers were classified as temporary layoffs, around two-thirds of the outflows went towards non-participation and to a lesser extent, search unemployment.[8] This raises important questions regarding the recovery dynamics, as it suggests that many of the job losses were permanent. Moreover, the recovery will also crucially depend on the entry rates of workers that were non-employed *before* the advent of the virus. It is key, therefore, to examine the trajectories of individuals during the pandemic. We do so by using our LFS panel to analyze labour-market transitions. Approximately two-thirds of the workers who were employed in February (i.e. the month just before the propagation of the virus) but non employed in April (when social distancing restrictions were peaking) regained employment in July. As a result, almost nine out of ten workers who were employed in February have a job in the summer.

Given the unprecedented severity of the COVID shock, these numbers can be seen as reassuring regarding the strength of the rebound. But the reemployment dynamics hide significant heterogeneity. The job losers classified as not in the labour force (NILF) in April have low reemployment rates; this, taken with the fact that a large part of employment outflows was towards non-participation, could be taken as a worrying sign. Another concerning aspect is that the job-finding rates of individuals not employed in February were negatively impacted by the pandemic in the reopening months. This indicates that this group has been largely excluded from the sizable employment gains of the rebound.[9] In sum, our flow analysis suggests that the incomplete employment recovery of July can be explained by three important

[7] In the worker flows literature, excess worker reallocation flows refer to offsetting flows in and out of employment, defined as the sum of reallocation flows (i.e. inflows and outflows), minus net employment changes, i.e. the flows that are "hidden" by the net change (e.g. Davis and Haltiwanger, 1992). Note that certain papers distinguish between *churning* and *excess* worker flows, the former being defined as excess worker reallocation minus excess job reallocation (e.g. Burgess, Lane, and Stevens, 2000). In what follows, we will, to avoid excessive jargon, use "churning" instead of "excess reallocation".

[8] The high share of temporary layoffs in outflows is in line with the surge in the number of unemployed workers classified as on "temporary layoff" (e.g. Jones, Lange, Riddell, and Warman, 2020; Canada, 2020). As a result, the share of unemployed workers classified as engaged in "job search" shrunk. This increase in temporary unemployment is one of the important features of the COVID-19 slowdown and it is not surprising that this is reflected in the outflows. However, although temporary layoffs represent a higher share than usual of the job losses (about one-third), around two-thirds of the outflows are towards labour-force status with no apparent link with an employer (i.e. job search and non-participation).

[9] For instance, a NILF individual who reported wanting a job in February was two times less likely to be employed in July than a similar worker in 2019 (16% versus 32%).

factors: the low reemployment probabilities of the numerous NILF job losers; the depressed hiring dynamics of those who where jobless before the outbreak; and the persistently high job losses of the reopening months.

Due to the high churn observed in COVID times, it is important to assess the distributional impact of the virus. We therefore examine labour-market histories conditional on socio-demographic and job characteristics. Many papers have already addressed this question using Canadian cross-sectional data (e.g. Lemieux, Milligan, Schirle, and Skuterud, 2020; Béland, Brodeur, Mikola, and Wright, 2020). But again, net changes can hide large gross job losses within groups, and therefore, might result in an understating of the distributional consequences of the shock. We find that youths, the low-educated, and workers in low-tenured and non-unionized employment were more likely to lose their jobs during the shutdown. Moreover, our results suggest a higher risk of job loss *and* a lower reemployment probability (after the reopening) for women as compare to men. The analysis also indicates an uneven impact across age groups, which is reflected in participation decisions. On the one hand, job losers who are 50 to 64 years of age were more likely to leave the labour force upon the reopening of the economy and on the other, there are signs of lower entry into the labour force for 20 to 29 year-olds. These results call for further analysis of the impact of the virus on retirement and education decisions.

The COVID-19 social-science research has been evolving at an extraordinary pace. Within this already large literature, our work can be related to papers that study the labour-market impact of the pandemic in North America. Several papers have analyzed the initial impact of COVID-19 in Canada (i.e. over the spring months). Lemieux, Milligan, Schirle, and Skuterud (2020) document striking declines in employment and especially hours worked (-32% between February and April among the 20-64 year-olds). They also show that the decline in hours disproportionately impacted low-earning workers, which is consistent with the findings of Koebel and Pohler (2020). There is also evidence that COVID has had a significantly negative impact on youths and the low-educated (e.g. Béland, Brodeur, Mikola, and Wright, 2020), the self-employed (e.g. Béland, Brodeur, Mikola, and Wright, 2020), and widened the gender employment differences among parents with young children (e.g. Qian and Fuller, 2020). Finally, Jones, Lange, Riddell, and Warman (2020) examine the early reopening dynamics, and present contrasting signs about the strength of the rebound: in particular, many of the temporary laid-off went back to work in May, but the number of job searchers increased.[10]

Most of the existing research on the Canadian labour market relies on the public-use files of the LFS.[11] Yet, an important limitation of these data is that one cannot follow individuals

[10] In addition, a substantial literature has been interested in the impact of COVID across occupations, in line with the widespread belief that the task content of jobs is critical when determining the effect of the pandemic. Key dimensions include: the possibility to perform tasks remotely (e.g. Gallacher and Hossain, 2020), the frequency of close contact with coworkers and the public (e.g. Béland, Brodeur, Mikola, and Wright, 2020) and belonging to an essential industry (e.g. Jones, Lange, Riddell, and Warman, 2020).

[11] As previously discussed, exceptions are two papers from the "COVID-19 Data insight for better Canada". Chang, Morissette, and H. Qiu (2020) provide employment separation rate estimates from the start of the outbreak and Hou, Picot, and J. Zhang (2020) examine transition rates in and out of unemployment, with a focus on immigrants.

over time, resulting in an inability to analyze labour-market flows and trajectories over the pandemic period. As such, little is known so far about the COVID-19 gross employment dynamics in Canada. Our paper intends to fill this gap. In addition, since the confidential LFS provides information regarding data imputation and entry and exit of households into LFS samples, we are able to propose a novel and detailed picture of the evolution and impact of non-response rates during COVID.

The U.S. literature on the labour-market impact of COVID is already substantive. Within this literature, our work primarily relates to papers that analyze labour mobility. Since the public-use files of the Current Population Survey (the U.S. counterpart of the LFS) allows for the possibility of following individuals over time, there is already a substantial body of research on this topic, as opposed to Canada.[12] Cowan (2020) analyzes labour market transitions between February and March across subgroups of workers and finds that women, visible minorities and the low-educated had a higher job loss probability. Forsythe (2020) who analyzes labour market flows over the same period finds that the hiring rate remains unaffected by the early phase of the virus — an intriguing result that we also observe in our analysis. Cheng, Carlin, Carroll, Sumedha, Lozano Rojas, Montenovo, Nguyen, Schmutte, Scrivner, Simon, Wing, and Weinberg (2020) also study the reopening phase and their results suggest a strong but fading reemployment dynamics upon the easing of COVID-19 restrictions across U.S. states. In addition, Şahin, Tasci, and Yan (2020) estimate the historical relationship between unemployment claims and unemployment outflows as to generate unemployment projections for the next months.[13]

While the U.S. Current Population Survey (CPS) allows for the building of panels of up to four months in length,[14] the ability to follow individuals over *six* months is a particular strength of the LFS. Therefore, the Canadian LFS makes it possible to follows trajectories from before the start of the outbreak, through the shutdown, and into the advanced phases of the economy's reopening. Since, in many cases, reopening policies are gradual (or, at least, have gradual effects) it is important to follow individuals long enough to assess the effect of such policies.[15] Moreover, the fact that the LFS is available at a monthly frequency is another key advantage; in many European countries, for instance, labour force survey data are quarterly. These features of the LFS, combined with the fact that governments implemented strict and pervasive constraints on economic activity, makes the Canadian case an informative one for assessing the economic impact of the virus in high-income countries.

[12] See Madrian and Lefgren (2000) and Rivera Drew, Flood, and Warren (2014) for discussions on the longitudinal matching of individual observations in the Current Population Survey.

[13] Moreover, Coibion, Gorodnichenko, and Weber (2020) document a stark decline in participation due to the pandemic, which can be accounted for by early retirements.

[14] The CPS is based on a 4-8-4-month rotating sample design: households in a given rotation are interviewed for four consecutive months, and then leave the sample for eight months, before being interviewed four additional consecutive months.

[15] If the LFS had a similar rotating scheme to the CPS, we would only have been able to follow an individual from February to May, when only part of the shutdown restrictions had been eased (or alternatively, from the shutdown to, say, June or July). However, it is possible to use information on unemployment duration to look at longer histories. But this has limitations, as many workers might have transited into non-participation. Moreover, such an approach makes it difficult to consider job characteristics, which are typically available when the worker is employed (or under the form of recall information).

The rest of the paper is divided as follows. Section 2.3 presents our data. In Section 2.4 we explore how the COVID-19 pandemic impacted the LFS survey. Section 2.5 provides a discussion of our labour market flows results, whereas Section 2.6 focusses on the worker trajectory findings. Finally, Section 2.7 concludes.

2.3 Data

For our empirical analysis, we rely on the internal-use files of the Canadian Labour Force Survey (LFS).[16] Although we use LFS data going back to 2018, our main focus is on the February 2020 through July 2020 period.

The LFS is a monthly household survey that collects labour market information of the Canadian population.[17] It is the official source of employment and unemployment data. The LFS is also one of Statistics Canada's larger surveys, as it interviews approximately 54,000 households every month.

The high frequency of the data (i.e. monthly) and its tight production deadline (from enumeration to public release),[18] ensures that the LFS is a timely source of data for examining the impact of COVID-19 on the Canadian labour market. The rotating panel feature of the LFS is key to our analysis. The LFS follows households for six consecutive months, with one-sixth of the households being replaced every month. As such, we can construct mini panels up to six months in length.[19] This allows us to follow individuals through the COVID-19 shutdown and subsequent re-opening of the economy.

We restrict our attention to civilian workers aged 20 to 64, and exclude those living in the territories.[20] We impose the age restrictions so as to compare our findings with existing evidence regarding the COVID-19 shutdown (e.g. Lemieux, Milligan, Schirle, and Skuterud, 2020; Jones, Lange, Riddell, and Warman, 2020). Finally, we focus our attention on Canada's ten provinces (i.e. exclude the territories) for data-access reasons; the internal-use files that we rely upon do not have information on the territories.[21]

[16]As discussed in the introduction, these restricted-access files are, to the best of our knowledge, very similar (if not identical) to the confidential-use files that are provided in the RDCs.

[17]Specifically, the target population of the LFS consists of the "*...civilian, non-institutionalized population 15 years of age or older. It is conducted nationwide, in both the provinces and the territories. Excluded from the survey's coverage are: persons living on reserves and other Aboriginal settlements in the provinces, full-time members of the Canadian Armed Forces, the institutionalized population, and households in extremely remote areas with very low population density.*" (Canada, 2018).

[18]For example, enumeration for the July 2020 data started July 19, the first day following the July 12 to 18 reference week, and the data was publicly released on August 7, 2020.

[19]This requires the use of the confidential-use files, as the variables needed to construct the mini panels are suppressed in the public-use files. The details regarding the construction of the mini panels are left to the appendix (i.e. Appendix B.1). See Brochu (2021) for more information on the distinction between public-use and confidential-use LFS files.

[20]The LFS asks socio-economic characteristics of all adults (15 years and up) that live in targeted dwellings, including full-time members of the Canadian armed forces. However, no labour market information is gathered for the latter group, as they are not part of the target population.

[21]It should be noted that the public-use files also only cover the ten provinces. The same holds true for the confidential-use files that are typically made available in the RDCs. LFS data for the territories is considered a different product (datatset), and as such, it would require a separate RDC application. Finally, it must

Given that our focus is on labour market flows and transitions, our analysis relies mainly on mini panels. As such, we must provide additional information regarding the periods covered by our panels and any additional panel-related restrictions that we impose. The monthly span and frequency of the panel vary depending on the flow/transition of interest. For example, when we focus on labour market flows from February 2020 to March 2020, we rely on the two-month balanced panel covering the February 2020 - March 2020 period. Say, instead, we want to look at the probability a worker is employed in June, conditional on her/him being employed in February but out of work in April. In that case, we use a balanced panel that spans the 5-month period, but only includes three months of data (i.e. February 2020, April 2020 and June 2020). We do so for sample size reasons. Using the full 6-month mini panel (February 2020 through July 2020) for all of our analyses, would force us to rely on only one rotation - the rotation that entered the survey as of February. This would dramatically reduce our sample size (i.e. 1/6 of its original size).[22]

Requiring a balanced panel means that we drop observations (individuals) for three reasons: first, we drop individuals whose rotation group rotated-out part-way through the panel; second, we remove individuals that join an existing household part-way through the panel, as identified by the new birth variables; and third, individuals who are absent in the data for some, but not all months, for reasons other than mentioned above are also dropped.

Of the three reasons, the first is the most restrictive. For example, when looking at a panel that spans four months, one must drop three rotations out of six, i.e. we lose half the sample–even before imposing any additional restrictions. The second restriction is the least restrictive; meaning that very few observations are lost. This is due to the short time-span of our panels, and may also be influenced by reduced mobility brought about by COVID-19. The third and final reason encompasses multiple possibilities: an individual moved out before the end of the mini panel (but whose household did not rotate out), perhaps there were "unusual circumstances" that resulted in no information being recorded,[23] or even the matching across months was poor.[24] Although one cannot separate out the importance of each, one can say that as a group they represent a small proportion of the sample.[25]

Given that the LFS is not a panel per se, one must rely on a series of variables to match an individual's information over time (see the appendix for more details). We thus verify that we are indeed following the same person over time. We follow the lead of Rivera Drew, Flood, and Warren (2014) and drop observations for which there are inconsistencies in age (in years) and gender across time. These variables should not change over time (other than becoming one year older) as the age and gender questions are asked of when an individual first enters the survey.[26] We lose very few observations due to this restriction.

be recognized that the official employment and unemployment statistics (which are based on LFS data) only apply to the Canadian provinces.

[22]The composition of individuals would change across time for all other rotations. For example, those that are part of the January rotation would rotate out of the survey as of June, and would be replaced by a new set of households as of July.

[23]This could be due to a household changing its mind and not wanting to share information any further, or the LFS not being able to find a "donor" for possible record imputation, to name just a couple of reasons.

[24]This could be due to a problem with the matching variables.

[25]For the two consecutive month panels, for example, it is only a couple of percentage points.

[26]The LFS records the exact date of birth (i.e. day, month and year). As such, a person's age could change

2.4 Non response and imputation during COVID-19

The COVID-19 pandemic and the ensuing policy restrictions thoroughly disrupted economic activity of many countries. Statistical agencies were not exempted, as their collection activities were severely constrained. This has, to a certain extent, been documented for the U.S., (e.g. CPS: U.S. Bureau of Labor Statistics, 2020b), but very little is known for Canada (e.g. LFS), which raises important concerns about our current assessment of the impact of COVID-19 on the Canadian labour market. This section attempts to fill the gap in the literature.

We start by first documenting the LFS non-response problem. Figure B.1 shows non-response rates (for the LFS) covering the January 2019 through June 2020 period. One can see a dramatic rise starting in March 2020, with the non-response rate reaching 28.2% by June 2020. To put these recent numbers into perspective, the monthly non-response rate had only exceeded 13% twice (14% in July 2018; 13.4% in August 2018) since the start of the modern day LFS, i.e. since 1976.[27]

To understand the implications of such a rise, we take advantage of information on when respondents entered the survey, and whether their labour market information was imputed, information that is suppressed in the public-use files.[28] If one looks at Figure B.2, one can see that prior to the COVID-19 shock, approximately 11,000 individuals per month entered the survey as part of the incoming rotation group, i.e. people in the first month of their six-month window.[29] By March 2020, however, this number had dropped by about a third as compared to the previous month. The number did start to recover as the country started to re-open, but it has never returned to its pre-COVID levels.

At first glance, it would appear that non response is mainly an incoming rotation issue, since the number of observations in the March sample fell by a similar amount (see Figure B.3), i.e. it was just hard to get a hold of "new" respondents. In fact, the story is more nuanced. The suspension of all field collection activities resulting from the COVID-19 shutdown not only made it much more difficult to initially contact new LFS units (households), it also made it difficult to follow up with units that were hard to reach by telephone.[30] This is confirmed in the data. If a household does not respond in the current month, but had responded in a previous one (and there is no indication that the residents of the households had moved out), the LFS will carry forward their socio-demographic information, but impute their *current* labour market information. The LFS follows a hot-deck procedure, where it uses previously collected information (socio-demographic and also some labour-market related) to find a current "donor".[31] As such, the drop in the overall sample size does not give justice to the true extent of the non-response problem experienced during these COVID times. Fortunately, the

from one month to the next, i.e. increase by one year.

[27]See Brochu (2021) for a detailed discussion of LFS non-response rates.

[28]The conclusions that we draw below are not sensitive to these sample restrictions.

[29]This number does not measure all new participants to the survey. Some individuals enter part-way through the six-month window, but this is a relatively small fraction of new entrants.

[30]This information is based on email conversations with a Statistics Canada methodologist during the late July / early August 2020 period.

[31]It uses the same WRI approach for the case where it cannot reach some, but not all household members. Given the option of proxy response, this is of second order importance. See Statistics Canada (2017) for a more detailed description of the hot-deck procedure, and the exact list of variables used to find a donor.

confidential-use LFS files can identify individuals for which all labour market information was imputed (which Statistics Canada refers to as whole record imputation (WRI)).

From Figure B.3, one can see that WRI (the gap between the two curves) has increased since February, which is indicative that non response is not just an incoming rotation issue.[32] If one looks at March, for example, one sees that the drop in the sample that excludes WRI (which accounts for both non response of new households, and of those households that had answered in the past), is in absolute terms, two-thirds larger than the drop in the incoming rotation as shown in Figure B.1 (which only accounts for non response of new households).

An important question to address is whether the increase in non response, as measured by the rise in WRI, impacts the data. To look at this issue, we present in Table B.1 five labour-market summary statistics which have been extensively analyzed in the COVID literature (i.e. employment rate, unemployment rate, participation rate, temporary layoffs as a share of the labour force, and employed but absent from work as a share of total employment) for samples with and without WRI. We show numbers for February, April and June, and also for the different panels that we rely upon in various parts of our analysis.[33] For comparison sake, we also show numbers using cross-sectional data. Focussing on the first column of results, the pre-COVID period, one can see that the February numbers (e.g. February employment rate) are similar from one panel to the next, and also with the ones that rely on cross-sectional data. The same holds true when we look at the April and June rates. This is not surprising given that dropping individuals when they rotate out of the LFS is the single most important reason for why our sample shrinks when the length of the panel increases.

What is of particular interest, however, is that dropping individuals whose labour-market information was fully imputed has very little effect on our findings (see Table B.1). A similarity between the two sets of numbers does not necessarily mean that the imputation procedure got it right, but it is reassuring nevertheless.[34] There are institutional and historical considerations that would lend credibility to this interpretation: first, labour force status and occupation (in addition to socio-economic characteristics) are used to find a donor for imputation. That occupations be relied upon is critical in the present context given that COVID-19 affected occupations very differently (Lemieux, Milligan, Schirle, and Skuterud, 2020). Second, although historically high by LFS standards, a non-response rate in the mid- to late-20s is not unchartered territory for statistical agencies. More precisely, Statistics Canada (and other statistical agencies around the world) have faced rising non-response rates over the last 20-30 years and they have experienced rates in this ballpark (and higher) for other surveys (e.g. Canadian General Social Survey).[35] Finally, the LFS is a long-running survey which means they have had experience with disruptions to field operations, at least at the regional

[32]It is important to recognize that the LFS cannot carry out WRI for non-respondents that are part of the incoming rotation group, as there is no prior information that can be used to find a donor.

[33]For example, the February employment rate when using our balanced two-month panel spanning the months of February and March is 76.6%. When using a three-month panel spanning a five-month period (i.e. February, April and June data), the February rate stands at 77.4%.

[34]It is similar to the case when carrying out regression analysis with and without weights. It is comforting when the results are similar, in that the findings are not dependent on whether the statistical agency calculated the weights correctly.

[35]See Barrett, Levell, and Milligan (2013) and Green and Milligan (2010) for a discussion of this worldwide phenomena, and its implication for researchers.

level, whether they be weather related (e.g. Quebec ice storm (1998); Hurricane Juan hiting Atlantic Canada (2003)) or labour strife (e.g. interviewer strike (2003)).

In sum, this section attempts to provide a better understanding of the impact of COVID-19 on the LFS, and in the process, explore whether the substantive rise in full-record imputation that arose affected the data as measured by some important summary statistics for the current crisis. Further work is needed to fully address this issue, including documenting the characteristics of those whose labour market information was fully-imputed during COVID-19. In a follow-up study, Brochu and Créchet (2021) found that the non-response rate is not a significant issue on the LFS estimates for those who already entered the survey rotation. However, it is a more concerning issue to reach out to respondents in the first place. And overall, the potential bias in the LFS caused by the high non-response rate during COVID-19 time does not affect the LFS estimates at the macro level; however, it could be an issue if it is used to study specific domain where the weighting and survey adjustment process does not cover such as specific demographic groups.

2.5 Labour market flows

2.5.1 Worker flows in and out of employment

The Canadian labour market underwent dramatic changes since the beginning of the pandemic. Severe employment losses that occurred between February and April, were followed by a strong rebound between April and July. Yet, employment is still significantly below its pre-COVID level in early 2020 and 2019. While most of the discussion has focused on employment changes in net terms, little is known thus far about gross worker flows that underlie these changes. We use the longitudinal dimension of our LFS data to analyze monthly worker flows in and out of employment over the February to July period. This allows us to draw a picture of worker movements from the month just before the propagation of the virus, up to the late reopening stages of the economy.

Using our two-month mini panels, we estimate the total number of workers that flow between employment and non-employment from period $t-1$ to period t, and the employment stock in period $t-1$. Our flow rate for period t is the ratio of these two estimates. Expressing total flows between month $t-1$ and t, in terms of employment as of period $t-1$, is done to remain consistent with the existing studies that have analyzed the impact of the pandemic on employment changes (e.g. Lemieux, Milligan, Schirle, and Skuterud (2020)). For instance, the separation rate for April would be the number of workers flowing out of employment between March and April expressed in terms of total employment in March. More details are provided in Appendix B.1.1.

The upper panel of Table B.2 reports gross monthly flows in and out of employment for 2020 (and 2019 for comparison). Unsurprisingly, there were unusually high separation flows over the February to April 2020 period. The February-March and March-April gross outflows are 7.5% and 13.9%, respectively. In comparison, the outflow rates for the same periods in 2019 (and for 2018, as found in appendix Table B.13) are much lower, falling in the 2-3% range. Interestingly, inflow rates (i.e. hiring) during these months are similar to those of 2019 (and 2018), which is perhaps unexpected given the strictness of the restrictions imposed on the economy at this time. This evidence suggests that the dramatic COVID-led job losses are not due to a decline in hirings, but solely driven by massive flows out of employment. As such, the impact of the COVID-19 shutdown is likely much larger than previously believed. In absolute terms, the cumulative outflows between February and April represent 3.6 million job losses (versus 2.5 million for the net job losses).

A second takeaway of this table is that the reopening months are also characterized by high outflows: for May-June, the separation rate stood at 6.5%, and around 5% for June-July. Therefore, the rebound observed over the months of May and June masks substantial "churning", as measured by the very high worker reallocation rates observed in these months, as compared to the previous years (see Tables B.2 and B.13). In effect, it hides job losses that are remarkably important. Moreover, there is a discernible asymmetry between the size of the outflows of the early pandemic phases, and that of the inflows seen during the reopening. This asymmetry, combined with the persistently high outflow rates, both contribute in explaining why employment has yet to recover to its pre-COVID level.[36]

We complement this analysis by examining the movements due to absence from work. It has been argued that these absences have surged during the shutdown, partly due to the misclassification of temporarily laid-off workers as being employed but absent from work (U.S. Bureau of Labor Statistics, 2020a; Jones, Lange, Riddell, and Warman, 2020, e.g.). Our estimates of employment flows could be contaminated by this misclassification issue as well, which may understate the size of the labour movements due to COVID. To address this concern, we report in the bottom panel of Table B.2 estimates of flows in and out of the LFS category "employed, at work". However, we enlarge this grouping to include employment absences that are least likely to be related to the outbreak (i.e. vacation, parental leave and labour conflicts). As such, our "employment, at work" outflows focus on transitions to non-employment, and to employment absences due to illness, caring for relatives and no availability of work. Inflows are similarly defined. As expected, this approach tends to produce larger flow estimates. The broad picture, however, is similar to that of the employment-flow dynamics. In particular, the at-work dynamics features high separation rates starting in February and persisting in the late spring and early summer (as compared to 2019).

2.5.2 Composition of worker flows

Further analysis of these unusual dynamics is key to understanding the employment recovery. In particular, assessing the degree of persistence in the massive COVID-related job losses is critical in evaluating the strength of the rebound. A first step of this analysis is to look at the composition of the employment flows using more detailed labour force status information. It has been documented that the recent downturn was characterized by a dramatic increase in the stock of workers classified as "unemployed, on temporary layoffs", as compared to previous recessions (e.g. Jones, Lange, Riddell, and Warman, 2020; Canada, 2020). This suggests that many of the job losses were only temporary in nature, and have been or will be recovered. To some extent, this seems to be confirmed in the data by a decline in the number of temporary

[36] See Tables B.16 and B.17 for estimates of employment flows by occupation groups. The tables show high employment outflows and churning in all the occupation categories considered (as compares to 2019). It also reveals significant heterogeneity across occupations, with the highest job losses experienced in the 'sales and services', 'arts, culture and recreation', and 'manufacturing' sectors.

unemployed workers since the beginning of the reopening. However, by looking only at the stocks, it is difficult to tell whether such decline is indeed due to reemployment of these workers, or instead due to flows into search unemployment and inactivity.

Table B.3 shows the share of the total employment flows (i.e. flows in and out of employment) for individuals transitioning to and from temporary unemployment, search unemployment, and out of the labour force (NILF). Clearly, the temporary-unemployment share is remarkably important as compared to previous years: it represents about one-third of outflows, while it is typically around or below 5% in 2018 and 2019 (see appendix Tables B.14 and B.15). This simple fact, which is in line with the surge in the temporary unemployment stock, is a good illustration of the very unusual nature of the employment inflows and outflows in recent months. The reopening stages are also characterized by a large share of employment inflows from temporary unemployment, as suggested by the May to July numbers. This indicates that a large reemployment movement took place with the progressive easing of the COVID-19 restrictions, which is consistent with the decline in temporary unemployment over this same period. However, it appears (again) that there is an asymmetry between inflows and outflows, suggesting that a substantial fraction of the COVID-related job losses have not been recovered. Another noticeable feature of the composition of flows is the high share represented by outflows to non-participation (NILF) between February and April (hovering around 50% of the outflows). One must be careful however when interpreting this number given the fine line between labour market statuses in COVID times.[37] Although this suggests that the pandemic shock induced numerous permanent job losses and transitions out of the labour force, it is difficult to gauge the size of these flows given the lack of clear distinction between labour-market states. Our worker transition analysis of the next section will help shed more light on the nature of these movements.[38]

All in all, this analysis allows us to characterize the COVID-time dynamics of the labour market as follows: first, the net job losses of the early stages of the pandemic can be accounted for entirely by massive employment outflows; second, these outflows are persistently high, even during the reopening stages of the economy, which is indicative of significant labour market "churning"; and finally, there is an asymmetry between the size and nature of the separation flows for February to April and the hiring flows for April to July. The size of the outflows is higher, and the temporary layoffs take up a larger share of outflows than inflows. This asymmetry, taken with the persistently high separation flows, accounts for the partial employment recovery observed in July. Additionally, one can observe that the flows out of employment to non participation seem to represent an important fraction of the jobs lost between February and April. However, the murky line between non-employment states in COVID times prevents us from drawing strong conclusions about the implications of this fact

[37] See Lemieux, Milligan, Schirle, and Skuterud (2020) who emphasize the grey zone between unemployment and non participation given the fall in job search and hiring activities, and Jones, Lange, Riddell, and Warman (2020) who stress the fine line between employment and non-employment given the surge in the number of employed but absent from work.

[38] The bottom panel of Table B.3 shows the composition of flows in and out of the "employed, at work" category, but like the bottom panel of B.2, excludes movements due to vacation, parental leave, and labour conflicts. The table indicates that, since the start of the outbreak, the majority of the outflows from this stock has been towards absence from work, but that a substantial fraction (around one quarter) has been towards non-participation.

for employment dynamics, at least at this stage of the analysis.

2.6 Worker transition analysis

2.6.1 Trajectories of early-pandemic job losers

Our worker-flow analysis of the previous section suggests large employment inflows for COVID job losers as social-distancing restrictions eased. However, these inflows do not offset the large employment outflows of the early pandemic stages (February to April), given that employment is still significantly below its pre-COVID level. Moreover, as we move beyond the reopening, it appears that the share of employment inflows from temporary unemployment is on the decline. The employment dynamics of the upcoming months will critically depend on the nature of worker trajectories associated with these flows. What explains the asymmetry between the outflows of the early stages and the inflows of the reopening phase? This asymmetry and the partial employment recovery are consistent with two opposite interpretations, with completely different implications for the employment dynamics in the ensuing months. First, this could be due to a staggered but steady return to work of COVID job losers as a result of the gradual easing of social-distancing restrictions, and the adaptation of businesses to the new environment. Alternatively, this could be attributable to permanent job losses, which would cause workers to slowly reallocate across jobs, occupations, and industries, or even lose their attachment to the labour market.

Analyzing the trajectories of workers since the beginning of the pandemic could be informative for the relative importance of the two possible interpretations discussed above. Specifically, looking at the transition patterns for workers who have lost their job due to COVID factors can help us better understand the labour-market dynamics of the reopening phase. As such, accurately identifying these workers is key. To do so, we once again make use of the detailed labour market status information regarding unemployment and absence from work. Given the dramatic rise in the number of temporary layoffs and absences from work (other than vacation) during the lockdown months, we can use this labour market information to create a sample that is, arguably, a reasonable representation of the COVID job losers. Moreover, and perhaps more importantly, we can take full advantage of the longitudinal dimension of our panel to analyze four- to six-month worker histories spanning the economy's trajectory over the February to July period. The analysis below is based on these approaches

Table B.4 reports, for April to July, the monthly transition probabilities of workers from temporary unemployment (i.e. unemployed, on temporary layoff), and those classified as employed but absent from work (excluding, again, absences due to vacation, parental leave, and labour conflict). The temporary unemployed have high monthly transitions to employment, standing at around 50% from April to July. This is consistent with our worker-flow analysis, which suggests large reemployment flows after April. However, it seems that the fraction of workers flowing to states with presumably low job attachment has been increasing over time since the start of the reopening: between June and July, almost 40% of the temporary unemployed have transited to search unemployment and non-participation. The trajectories of those absent from work also display increasing flows to low job-attachment states, but to a lesser extent: the transition rate of this group to search unemployment and non-participation is around 18% over June and July.

However, this analysis provides an imperfect picture of the trajectories of COVID-19 job losers, as temporary unemployment and absence from work might be subject to seasonal variations not related to the pandemic, especially when looking at summer months. Table B.5 analyzes trajectories of the workers employed in February (i.e. the month just before the beginning of the propagation of the virus in Canada) *and* non-employed in April, when social distancing restrictions were peaking. The transitions suggest again important reemployment flows upon reopening: the share of these workers flowing back to employment after April is increasing over time, and as of July, approximately two-thirds of this group are employed again. However, with these numbers alone, it is difficult to gauge the size of these reemployment flows, given that the composition of workers with similar histories in previous years is probably very different due to the pervasiveness of the COVID shock. Hence, to get an additional sense of the magnitude of these flows, we compare the trajectories of workers employed in February 2020 (i.e. before the pandemic) with those employed in the same month in 2019 (see Table B.7). The probability of being employed in July (conditional on employment in February) is 88.4% for the 2020 group, versus 94% for 2019; the 2020 group is also slightly more likely to be non participating in July (6.2% versus 3.8%). When put in perspective — accounting for the severity of the COVID shock — these numbers can be seen as reassuring regarding the strength of the reemployment dynamics of the reopening phase.[39]

These numbers, however, might hide important heterogeneity. To address this issue, we conduct the transition analysis of Table B.5 by looking at trajectories of workers conditional on the detailed out-of-work status in April. This reveals important differences in trajectories of workers who were out of work but presumably kept a link with their former employer (i.e. those in temporary unemployment and absent from work), and the group of workers who were classified as out of the labour force. It appears that a large majority of those with an employer attachment have been reemployed, but the picture looks very different for those who were NILF in April. As of July, only 58.4% of the NILF group was reemployed, versus 75.9% for the temporary laid-off group. Given that, as emphasized in the worker-flow analysis, a large share of the employment outflows of the early pandemic stages was towards non-participation (see Table B.3), the staggered return to work of this group is presumably key in accounting for the incomplete employment recovery of July. Importantly, this also indicates that the behaviour of these workers in the next months might be an important determinant of the pace at which the labour market recovers.

All in all, the transition analysis suggests important heterogeneity in the trajectories of the COVID job losers, which is in line with the now well-established idea that the shock is being felt very differently across occupations and industries. Many appear to have regained employment, especially among those who were classified as temporarily out of work (temporary

[39]We find, moreover, that the majority of these reemployed workers went back to their previous industry and occupation. Table B.6, reports the probability of remaining in the same industry/occupation between February and the months of May, June, and July, conditional on being non-employed in April. The table shows that the majority of these workers did not experience an industry or occupation switch, even when considering the most detailed classifications (5-digit NAICS industries and 4-digit NOC occupations), which suggests that many of them went back to their previous employer. It appears, moreover, that the probability of no switch is especially high for the temporarily unemployed in April.

laid-off and absent from work). But even among this group, there are some signs of a fading link with pre-COVID employers, as suggested by increasing transitions to search unemployment and non-participation since the start of the reopening. In addition, and perhaps most concerning, the numerous workers who have transited from employment to non-participation between February and April seem to be having a hard time regaining employment. Understanding the sources of this heterogeneity might be key for drawing implications about the nature of the recovery. Moreover, this will be crucial for understanding the distributional impact of the pandemic.

2.6.2 Trajectories of non-employed individuals

Our discussion of the labour market trajectories during COVID-19 has focused on the experience of recent job losers at the expense of individuals who had been out of work before the pandemic. Table B.2 indicates that, upon the reopening, temporary unemployment represents a higher than usual share of employment inflows. In addition, the large employment probabilities of workers temporarily out of work in April (and employed before the pandemic, see Table B.5) suggest that reemployment flows account for a large share of the reopening employment gains. On still needs to evaluate the trajectories of individuals out of work at the time of the initial COVID shock if on is to better understand the employment dynamics during the pandemic. Assessing how the sizeable employment inflows have benefited individuals without a prior job is important for drawing distributional implications of the outbreak. Moreover, in a typical year, non-employed individuals classified in search unemployment or in non-participation — i.e. without any obvious attachment to a given employer, as opposed to workers classified as temporarily laid-off — represent the lion's share of employment inflows (see appendix Tables B.14 and B.15). Therefore, although the current discussions have been mainly focused on labour market outcomes of the COVID-19 job losers, the fact remains that the trajectories of individuals out of work prior to the pandemic are likely to matter a great deal for employment dynamics.

Table B.7 reports employment and unemployment probabilities for May to July, of individuals not employed in February (for 2019 and 2020). The table indicates that the employment probability of the 2020 group is quite low when compared to similar workers in 2019, and when contrasted with the reemployment dynamics of the reopening. This is true for both unemployed and non-participating individuals who report wanting a job in the LFS, i.e. for workers who can be considered as attached to the labour market. In 2020, a worker who was unemployed in February had a 42.8% probability of being employed in July; in 2019, the same employment probability was 55.8%. Strikingly, an individual not in the labour force but wanting a job in February was two times more likely to be employed in July in 2019 than 2020 (32% versus 16%).

2.6.3 Trajectories conditional on individual and job characteristics

We complement our analysis by looking at individuals' trajectories, conditioning on socio-demographic and job characteristics. The severity of the shock and the high churn experienced by the labour market suggest that the pandemic has important distributional consequences. This is in line with what has been documented in the existing literature (e.g. Lemieux, Milligan, Schirle, and Skuterud, 2020; Béland, Brodeur, Mikola, and Wright, 2020). We address the three following questions: 1) how has the initial impact of the COVID shock been distributed across workers and jobs? 2) What characteristics are correlated with reemployment in the case of the pandemic job losers? And 3) how has the trajectories of the different subgroups of the pre-COVID non-employed been affected? Given that the focus of the paper is on job mobility, our outcomes of interest are individuals' transitions across labour force states. Specifically, we look at transitions from February to July in order to follow workers along the different phases of the pandemic.

The initial impact of COVID-19 on job losses

We examine the impact of COVID-19 on gross job losses across groups of workers, by relying on mini panels for 2019 and 2020 (covering the months of February and April).[40] The econometric model takes the following form

$$y_{i,m,s} = \alpha_0 + \alpha_1 d_s + X'_{i,s}\beta + \big(d_s \times X'_{i,s}\big)\gamma + \varepsilon_{i,m,s}, \tag{2.1}$$

where $y_{i,m,s}$ is a labour-force-status dummy for individual i as of m = April in year $s \in \{2019, 2020\}$,[41] and d_s is a dummy variable taking the value one if the individual is observed in 2020. $X_{i,s}$ is a vector of indicator variables for socio-demographic and job characteristics for individual i as of February in year s. The socio-demographic variables consist of female and aboriginal indicator variables, dummy variables for age,[42] highest educational attainment,[43] young-child parenting interacted with the female indicator,[44] and province of residence. For the job characteristics variables, we have dummy variables for union status (i.e. member or covered by a union), job-tenure situation,[45] and one-digit occupation groups.[46]

The sample consists of individuals employed in February (in 2019 and 2020) and observed in the data two months later (i.e. April).[47] We consider two sets of regressions. First, we analyze transitions to non-employment (i.e. $y_{i,m,s} = 1$ if individual i is not employed in month

[40]More precisely, we rely on two panels: the two-month panel spanning the February 2019 to April 2019 period, and the two-month panel spanning the February 2020 to April 2020 period.

[41]We present the equation in a general form by using an m subscript for the month as to be consistent with the notation used in the subsequent analysis, where the month of observation is allowed to vary.

[42]Age dummies: 20 to 29, 40 to 49, and 50 to 64. Those that are 30 to 39 years of age are the reference group.

[43]Highest educational attainment dummies: dropout, college (post-secondary/trades certificate, community college, CEGEP and university certificate below bachelor's), and bachelor's degree and up. The reference group consists of those that have no more than a high school degree (and also includes those that have some post-secondary education but no certificate, diploma or degree).

[44]More precisely, a binary variable that equals one if the person is female and has a child under the age of six.

[45]Job tenure dummies: 1 to 11 months, and 12 to 35 months. Workers with 36 months or more of job tenure with the same employer are the reference group.

[46]We use the NOC's ten larger categories (i.e. one-digit groups). Health is the reference group. Including broad occupation groups instead of exploiting more detailed NOC information has the advantage of allowing us to analyze heterogeneity across occupations based on the coefficients of the associated dummies. An alternative approach, that could be followed in further research using the LFS, is to analyze transitions conditional on task-content index values (e.g. potential for remote work, public facing), as in Cheng, Carlin, Carroll, Sumedha, Lozano Rojas, Montenovo, Nguyen, Schmutte, Scrivner, Simon, Wing, and Weinberg (2020).

[47]The unit of observation is a person (with data for February and April) in a given year: each person is only

m (i.e. April) of year s, 0 otherwise), to better understand the composition of the high job loss flows of the shutdown (see Table B.2). Second, we analyze transitions to non participation (i.e. $y_{i,m,s} = 1$ if individual i transits to non-participation in month m (i.e. April) of year s, 0 otherwise). Given the large share of employment outflows to non-participation (see Table B.3) and the low reemployment probability of the NILF in April (Table B.5), we deem important to examine the characteristics of the NILF job losers. For each regression set, we add the covariates sequentially.

The results are reported in Table B.8. Column (1) focusses on the probability that an individual employed in February will be non-employed in April (of that same year), whereas Column (4) focusses on whether the individual employed in February will be in the non-participation state in April (again of the same year). In both cases, the only explanatory variable is the year dummy. These estimates, which complement the flow analysis of Section 2.5 give a good illustration of the severity of the initial shock. For instance, the coefficient estimate (of the year dummy) suggests an increase in the job loss likelihood of 14 percentage points in 2020, as compared to 2019. Given the considerable magnitude of the shock, it might be reasonable to attribute such a change to the virus. The probability of transition to non-participation is also very large compared to 2019 (i.e. an additional 7 percentage points), reflecting the important share of employment outflows towards non-participation highlighted in our flow analysis (Table B.3).

Columns (2) and (4) analyze transitions to non-employment and non-participation, but where we now add demographic covariates, on their own and interacted with the year dummy. Our estimates suggest a higher probability of job loss for women than men, as well as higher transitions to non-participation. The differential effect seems of substantial magnitude and is highly significant. This is in line with the literature documenting a more severe impact on women in the U.S. But this contrasts with the existing evidence for Canada that is mixed: for instance, Béland, Brodeur, Mikola, and Wright (2020), looking at net changes, find no distinguishable difference across genders. Looking at gross job losses offers a different picture revealing a clear negative effect for women relative to men. The estimates also reveal higher transitions to non-employment for youths (20 to 29 year olds) as compared to those 30 to 39 year of age, and to a lesser extent, for lower level of educations (as compared to college and up).

Columns (3) and (6) add job-characteristics covariates. The estimated impact is substantially higher for low-tenure jobs. Tenure below one year was associated with an additional 5 percentage points job-loss probability, as compared to workers that had been with their employer three or more years. This is consistent with the literature arguing that low-tenure jobs are the most fragile due to heterogeneity in match quality (e.g. Jung and Kuhn, 2019) and information frictions (e.g. Jovanovic, 1979; Pries and Rogerson, 2005). Obviously, industry and occupation might also play a role. Unsurprisingly, there are important differences across occupations, which tend to be amplified when looking at flows. For instance, being employed in sales and services (as compared to health) was associated with an additional 10 percentage point job-loss probability. Note that controlling for job characteristics attenuates the coefficient for the youth age group; this suggests that occupations and especially tenure might play

observed once in the sample, i.e. in 2019 or 2020. For each person one has their labour force status in April and their socio-demographic and job characteristics as of February (i.e. two months earlier).

a key role in accounting for the high impacts for youths.

Reemployment probability

We examine reemployment patterns of the early-pandemic job losers, using mini panels for 2020 only.[48] Our econometric model is as follows

$$y_{i,m} = \phi_0 + X_i'\psi + \nu_{i,m}, \tag{2.2}$$

where $y_{i,m}$ is a labour force status dummy for individual i in month $m \in$ {May, June, July}. X_i are individual socio-demographic and job characteristics as of February. We estimate equation (2) for each m separately, and as such, the sample varies with m. More precisely, for each month m, the sample consists of workers employed in February but not in April, and for whom we observe their labour force status in month m.[49]

We explore the reemployment transitions, and to further understand the trajectories of job losers we also look at the transitions out of the labour force. As previously done, we add the covariates sequentially: we first only include the socio-demographic variables, and then add the job characteristics.

The results for reemployment transitions are reported in Table B.9, whereas those for transitions out of the labour force are presented in Table B.10. The estimates indicate lower reemployment probabilities for women than men for May and June, which suggests that women have more difficulties regaining employment upon the reopening (columns (1), (3), (5)). Controlling for job characteristics attenuates the gender differential, suggesting that part of the difference could be due to occupation composition (columns (2), (4), (6)). Indeed, we observe large differences across occupations in terms of reemployment, with low reemployment probabilities for education and social services or arts and culture (as compared to health).

The results also suggest a distinct reemployment patterns for individuals 50 to 64 years of age. These individuals have remarkably low reemployment probability in July relative to other groups (column (6) of Table B.9). At the same time, they were significantly more likely to exit the labour force in June, and especially in July (column (6) of Table B.10). These results align with Coibion, Gorodnichenko, and Weber (2020), which would suggest that a large numbers of older workers are retiring early due to the virus. This possibility, combined with the typically low job-finding probabilities of these older workers, raises concerns about the pace at which employment will recover.[50] Further research should analyze the mechanisms behind these labour-market outflows.[51]

[48] More precisely, we rely on three panels: the three-month panel covering February, April and May; the three-month panel covering February, April and June; the three-month panel covering February, April and July.

[49] The unit of observation is a person. For each person i one has their labour force status in month m and their socio-demographic and job characteristics as of February.

[50] As explained by Coibion, Gorodnichenko, and Weber (2020), retired individuals typically have very low rates of entry into the labour market.

[51] Moreover, it is important to understand the motive for leaving the labour force and the profile of the leavers for assessing the implications of the crisis for the evolution of wealth and consumption distributions.

The effect of COVID-19 on the trajectory of the non-employed

We now analyze the impact of COVID-19 on the trajectories of individual who were non-employed before the pandemic. We consider again model (2.1), but focusing now on those who were non-employed (and also the sub-group of non-participants) in February and who we observe in month m of the same year, for $m \in \{\text{May, June, July}\}$. We again use data for 2019 and 2020. We analyze transitions to employment, i.e. our dependent variable, $y_{i,m,s}$ is a dummy taking the value of one if individual i (who was not employed in February) is employed in month m of the same year. Table B.11 shows regression results for the sample of workers who were unemployed in February, and Table B.12 for non-participating individuals. Again, we present separately results for the unconditional-mean 2020 effect and for after adding individual characteristics.

Columns (1), (3), and (5) of Table B.11 show the impact of the economy's reopening on the employment probabilities of those who were unemployed in February. The virus seems to have had a strong and significant negative impact on these workers, which confirms the results of Table B.7. The effect is quite persistent as it lasts even during the late phase of the reopening. This is in stark contrast with the strong (but partial) rebound of the reopening, and the large reemployment flows for certain groups of workers (Table B.5). Table B.12, which focusses on those out of the labour force in February, suggests a lower average impact for this group as compared to the unemployed. However, there is heterogeneity within this population. Moreover, the result suggest a relatively high negative effect on the non-participating youths, at least for June (see column (5)). This, taken with the high job loss rates for youths (and discussed above) can be seen as a sign of relative vulnerability.

2.7 Conclusion

This paper analyzes the dynamics of the Canadian labour market from the beginning of the propagation of COVID-19, and the ensuing lockdown, to the reopening of the economy. We take advantage of the longitudinal dimension of the confidential-use files of the Labour Force Survey to analyze the size and composition of employment inflows and outflows, and to examine worker trajectories over the course of the pandemic. Our analysis shows that the Canadian labour market has experienced high "churning", which has persisted even after the easing of social distancing restrictions. We find evidence suggesting large reemployment flows of recent job losers, especially among workers who have been temporarily laid-off during the lockdown. There is, however, important heterogeneity in worker trajectories in recent months. Even for those who seemed to have temporarily lost their jobs, we observe increasing transitions to job-search unemployment and non-participation, which suggests an erosion of the link with previous employers. Even more concerning, is the evidence suggesting reemployment difficulties of COVID job losers who are classified out of the labour force in April, and the unusually low job-finding rate of workers not employed prior to the pandemic. Examining more closely the outcomes and behaviour of these groups might be key to drawing implications about the labour market recovery moving forwards, and for distributional consequences of COVID-19.

Chapter 3

Gender Differences in Early Labour Market Trajectories and Wage Growth in Canada

3.1 Abstract

Previous studies based on the European labour market have shown that the gender pay gap is relatively low when entering the labour market, yet widens quickly in early career, over the first 5 to 10 years, which highlights the importance of examining and decomposing the differences in early career labour market trajectories and wage growth. Using the Longitudinal Workers File (LWF) linked to the 2006 and 2016 Census. This is the first study in Canada to identify differences in labour market trajectories and wage growth for young workers aged 25 to 34 at the beginning of their careers. Results from both descriptive analysis and regression analysis suggest that, although some progress has been made on the gender gap on labour market trajectories for young workers in their earlier careers, women had a greater disadvantage compared with men. More specifically, regardless of the type of job separation, changed employer or occupation, and reason for separation, women's weekly wage annual growth rate was lower compared with their male counterparts. The results from regression analysis suggest that labour market trajectories affect women's and men's weekly wage annual growth rates differently. Women's weekly wage annual growth rate was more sensitive to temporary job separation compared with men, whereas men's weekly wage annual growth rate was more sensitive to permanent job separation. The Oaxaca-Blinder decomposition shows that the total number of permanent separations and the total number of permanent separations due to parental/maternity leave each explains about one-third of the cross-sectional gender differences in weekly wage annual growth rate observed for young workers in Canada.

3.2 Introduction

In Canada, the gender gap in hourly wages among employees aged 25 to 54 was 13.3% in 2018, where more than two-thirds of this gap was still unexplained by standard controls (human capital, job attributes, occupation and industry, and demographics) (Pelletier, Patterson, and Moyser, 2019). This points to a continued need for analysis in this area in order to better understand gender-based wage disparity.

Previous studies have shown that the gender wage gap is relatively low when entering the labour market, yet widens quickly in early career, over the first 5 to 10 years (Manning and Swaffield, 2008), which highlights the importance of examining the differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason of separation. In addition, according to studies by Hirsch and Schnabel (2012) and Eryar and Tekgüç (2014) based on European labour markets, men were more likely than women to make a job-to-job move, meaning a switch in employers, while women were more likely to go from job to non-employment, typically for maternity leave. In comparison, the results from Canada remain mostly unclear due to the lack of relevant studies. Existing works (e.g. Moyser, 2019) have primarily focused on the gender wage gap itself; however, since the differences observed in the gender wage gap start low but diverge fast in later years, labour market trajectories could be an important factor to consider when tackling this difference(e.g. Topel and M. P. Ward, 1992, Bagger, Fontaine, Postel-Vinay, and Robin, 2014 and Jung and Kuhn, 2019). In addition, factors associated with a job move, such as the type of separation and post-movement labour market trajectory, remain mostly unclear and could differ substantially. Individuals with different job paths could have very different wage growth. To close these gaps, this study examines and decomposes the gender differences in labour market trajectories for young workers in Canada and their potential impact on wage growth over the first ten years of their careers. More specifically, this study attempts to address three questions. First, what are the main gender differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason for job separation for young workers? Second, how have gender differences in early career labour market trajectories impacted wage growth between 2005 and 2015? And lastly, do the gender differences in early career labour market trajectories help to explain a portion of the gender pay gap through wage growth? In this study, a temporary separation is identified as an employee returning to his or her employer during the year of the separation or in the following year. When such a return does not occur, the separation is identified as permanent separation instead. The first question is analyzed through descriptive statistics analysis. The latter two questions are examined through descriptive statistics and regression analysis, followed by an Oaxaca-Blinder decomposition.

There are two main contributions of this study. First, this is the first study in Canada to identify gender differences in labour market trajectories for young workers (aged 25 to 34) at the beginning of their careers. Second, this study attempts to decompose the gender differences in labour market trajectories by the type of job separation changed employer or occupation and reason of separation and their impact on wage growth to help to answer the unexplained portion of the gender wage gap using the Longitudinal Workers File (LWF), which is a rich longitudinal data set that can be linked with Census information, allowing the researcher to track a cohort of young workers over ten years.

Results from both descriptive statistics analysis and regression analysis suggest that although some progress has been made on the gender gap on labour market trajectories, such as the share of individuals who had a permanent job separation for young workers, women had a greater disadvantage compared with men. More specifically, regardless of the type of job separation, changed employer or occupation, and reason for separation, women's log weekly wage growth overall was lower compared with their male counterparts. The results from regression analysis suggest that labour market trajectories affect women's and men's weekly wage annual growth rates differently. For instance, women's log weekly wage growth was more sensitive to temporary job separation compared with men, whereas men's weekly wage annual growth rate was more sensitive to permanent job separation. Further, an Oaxaca-Blinder decomposition shows that the total number of permanent separations and the total number of permanent separations due to parental/maternity leave each explain the differences in the cross-sectional gender differences in weekly wage annual growth rate observed in Canada by 33%, conditional on other observables.

This study contains seven sections in total. Section 3.3 provides a literature review. Section 3.4 discusses the framework of the study. Section 3.5 provides the background of the dataset and methodology used. Section 3.6 presents the empirical results in detail. Section 3.7 discusses the limitations of this study, and Section 3.8 presents the conclusions.

3.3 Literature Review

Several studies have examined how gender influences job mobility and labour market trajectories. Eryar and Tekgüç (2014) studied job mobility and labour market trajectories from a gender perspective in Turkey, where they found women are more likely than men to move from employment to non-employment, and at the same time, women are less likely to have a job-to-job move compared to men. Similarly, Hirsch and Schnabel (2012) used independent competing risk models, route-specific separation rates modelled as stratified Cox models to study the gender differences in job separations. The study finds that the separation rate for women is lower than for men, and women are less inclined to make "wages increasing" voluntary job-to-job moves. Also, they found that older workers are less likely than younger workers to have job transitions, while collective agreement declines the likelihood. Their regression analysis results conclude that job separations for women are less likely to be affected by a collective union and poor economic performance when compared with men. From the perspective of family composition, Frederiksen (2008) looked at the gender differences in job separation rates and employment stability based on data from the Integrated Database for Labour Market Research (IDA) in Denmark, which include over 3 million individual records. The author concludes that marriage is associated with more job mobility stability for women, while children have an opposite effect, meaning women are more likely to separate, while men are more likely to stay infirm. Wage level and workplace size have similar effects on job separation effect by gender.

Some recent studies have examined the conventional socioeconomic dimensions from various aspects that are associated with the gender wage gap. In a recent study by Joshi, Bryson, Wilkinson, and K. Ward (2019), they track those born in a single week in Great Britain in 1958 through to their mid-50s based on repeated cross-sectional surveys to study the change in the gender wage gap over time. The data used in their study shares great similarities and properties to the LWF-Census linked file, which is used in this study. For instance, the LWF-Census file allows the researcher to track the same individual from different cohorts between 1989 and 2017 and to study the changes in their labour market outcomes. According to Joshi, Bryson, Wilkinson, and K. Ward (2019), there is an inverse U-shaped gender wage gap over their life-course: an initial gap in early adulthood widened substantially during childrearing years, affecting earnings in full-time and part-time jobs. In their descriptive approach, they found that education-related differences are minor. On the other hand, they found that gender differences in work experience are the most significant contributor to that part of the gender wage gap, which we can explain in their models. In addition, they conclude that family formation primarily affects the gender wage gap through its impact on work experience. As a final remark, the authors mention that not all of the gender wage gap is linked to family formation. For instance, there was a sizeable gender wage gap on labour market entry, and there are some otherwise unexplained gaps between the pay of men and women who do not become parents. Unlike other traditional gender-based wage inequality studies, Card, Cardoso, and Kline (2015) examined the firm-specific pay premiums as a source of gender wage inequality from the labour demand side. Using data collected from the firm's side, they found that the wages of both men and women contain firm-specific premiums that are strongly correlated with simple measures of the potential bargaining surplus at each firm. They have shown that the impact of these firm-specific pay differentials on the gender wage gap could be decomposed into a combination of sorting and bargaining effects. In conclusion, the authors found that women are less likely to work at firms that pay higher premiums to either gender, with sorting effects being most important for low and middle-skilled workers. Moreover, they found that women receive only 90% of the firm-specific pay premiums earned by men.

Del Bono and Vuri (2011) studied the labour market trajectories and the gender wage gap in Italy and found that the difference in average men's and women's earnings has been relatively stable since the mid-90s. They found that the gender gap in earnings is low in early career and then grows quickly afterwards; most importantly, authors decomposed log wage growth for job mobility and labour market trajectories and found that about 30% of log wage growth for men over the first ten years of experience is accounted for by transition. Wage growth was much higher for men who moved than those who did not in the first ten years, whereas it's the opposite for women. Mobility only contributes to 8.3% of log wage growth for women. In addition, 53% of men in the sample move to another employer, compared with 47.7% of women, but the gender differences in returns to mobility are what stand out. Differences in returns to between-firm mobility account for half of the observed gender pay gap. Hahn, H. R. Hyatt, Janicki, and Tibbets (2017) conducted a decomposition of labour market trajectories to explain earnings growth at the macro-level in the U.S. Although their methodology is not relevant to this study, they found that consistent with previous results, job stayers are responsible for the bulk of fluctuations in earnings. Moves into non-employment have a downward pressure, while job-to-job moves have a modest positive effect on earnings.

Focusing on the interactions between family, labour market trajectories and the gender wage gap, Fuller (2008) found that women with children or husbands in the U.S. are less able to capitalize on employer change, especially when their mobility is above the norm. Workers with high mobility not able to benefit from higher tenure in the firm, especially women, are penalized in terms of mobility. Essentially, workers who are already doing well in the labour market have less to gain in wages from mobility, though there is considerable variation in mobility and wages. In a similar context, Bertrand, Goldin, and Katz (2010) studied the post-MBA labour market outcomes from Chicago Booth in the U.S. They found that women and men from the program have identical labour incomes and hours of work initially, but the gender gap widens considerably over their career as measured at five years and ten to sixteen years after MBA completion. For instance, this study finds major gender differences in career interruptions associated with loss of earnings during career interruptions and growing differences in weekly hours worked. The presence of children is responsible for less job experience, more interruptions, and less likely to return to the labour market if the spouse has a high income, shorter work hours and lower earnings. In contrast, the study did not find any impact on men, and fathers seem to do better after having children. In a similar fashion, Kleven, Landais, Posch, Steinhauer, and Zweimülle (2020), Kleven, Landais, Posch, Steinhauer, and Zweimüller (2019), Kleven, Landais, and Søgaard (2019) and X. Zhang (2010), looked at the family gap, and the difference in earning trajectories for people with and without children. They find that the large penalty for mothers explains a large fraction of the gender gap in earnings.

From the perspective of job quality and gender, Manning and Swaffield (2008) examined the gender gap in early-career wage growth based on the British Household Panel Study from 1991 to 2002; where they used reduced-form estimates of wage growth and found that the gender pay gap emerges later in life, after five years and then ten years of experience; and is not evident at the beginning of the career. Further, they found that there is a penalty for taking time out from the labour market of about 4.4% per year for men and 4.7% for women. Most importantly, they found that occupational differences cannot explain much of the gender gap in early-career wage growth. Women who work continuously full-time for ten years can also expect to be 12% behind men with the same characteristics. Men are more likely to move to a better job than women are. Essentially, this study shows that researchers need to pay more attention to the emergence of the gender pay gap in early career, as women who entered with the same pay were 25 log points behind ten years later. Job-shopping theories explain about 1.5 log points, while psychological differences explain up to 4.5 log points. Lima (2004), based on the Portuguese employer-employee dataset from 1991 to 1999, explores the relationship between labour market trajectories and wages. They found that a high incidence of long-term employment relationships and new jobs tend to be short-lived. In addition, part-time workers are more likely to leave, and workers in large firms are more likely to stay. Lima (2004) also finds that workers who exit the firm tend to earn higher wages after a job separation. Abbott and Beach (1994) use the Labour Market Activity Survey from 1986-87 (about 2,100 women) to examine how wage rates of working women in Canada evolve when they change jobs. Their main findings suggest that Canadian women who changed jobs in 1986 realized short-run wage gains of 8-9% and that women who quit for non-personal or job-related reasons had substantially greater gains than those who quit for personal reasons or were laid-off.

In summary, most existing gender-based literature on labour market trajectories and wage inequality focus on examining the factors associated with the gap with respect to the conventional socioeconomic dimensions of education, family composition and job quality. However, factors associated with a job move, such as the type of separation and post-movement labour market trajectory, remain mostly unclear and could have long-term implications on wage growth that could ultimately contribute to the gender pay gap. This points to a continued need for analysis in this area, including the unexplored factors under the channel of labour market trajectories that could help to explain a portion of the unexplained gender pay gap through wage growth.

3.4 Framework

According to the existing literature discussed above, there is still a gap in the literature of gender differences on labour market trajectories with respect to wage growth from a gender perspective, especially for young workers in Canada at the beginning of their careers (e.g. Moyser, 2019). Joshi, Bryson, Wilkinson, and K. Ward (2019) concluded in their study that work experience is the biggest contributor to that part of the gender wage gap in their model. Combined with Manning and Swaffield (2008)'s findings on penalties of labour market interruption, it is reasonable to think that gender difference in labour market trajectories could be an important channel that contributes to the gender wage gap through wage growth. Gender differences in early career labour market trajectories depend on the type of job separation, and factors associated with job move could affect wage growth very differently. These differences could ultimately contribute to the gender wage gap that is largely remained unexplained in the existing literature. Therefore, in this study, the concept of labour market trajectories is identified separately into three major definitions. First, the type of separation, including both permanent and temporary separation. Second, changed employer or occupation, which indicates whether individuals have switched employer or changed occupation between 2005 and 2015. And lastly, the reason for job separation is recorded on the Record of Employment (ROE).[1] As Lima (2004) mentioned in his study, while the traditional concept of job separation is an important factor in determining the the gender wage gap in Portugal, it depends on the underlying reason for such separation, and it could lead to very different long-term outcomes. Although there are important cross-country differences in labour market policies such as maternity and paternity leave between Portugal and Canada. And as such, one would not expect the finding for Portugal to fully translate to Canada since the labour market policies are dramatically different. However, given the high relevance of the subject matter between Lima (2004) and this paper, citetLima2004 's results can still at least shed some light on the framework of this study.

Given the high degrees of relevance to the data structure and subject matter, this study based from Del Bono and Vuri (2011) and Joshi, Bryson, Wilkinson, and K. Ward (2019)'s approach and attempts to further explore and decompose the impact of labour market trajectories by type of separation, changed employer or occupation and reason for job separation on the unexplained portion of the gender wage gap through wage growth. In other words, this study uses similar groups of regressor or blocks from Joshi, Bryson, Wilkinson, and K. Ward (2019) with additional controls identified by Del Bono and Vuri (2011) to explore the impact of labour market trajectories on the gender wage gap. In Joshi, Bryson, Wilkinson, and K. Ward (2019)'s study, the regression model contains three main blocks: "Education (ED)" block, which includes individuals' highest educational qualification, academic or vocational equivalent achieved by the time of each survey; the set of work experience variables "Experience (EXP)" which includes the number of months that cohort members had worked full-time or part-time up to each survey; and lastly the "Family (FAM)" block which includes the current presence of a partner in the household, and of children at various ages; and whether the individual had ever reported a child in the household. Most of these variables

[1]As described later in the data section, in this study, a temporary separation is identified as an employee returning to her or his employer during the year of the separation or in the following year. When such a return does not occur, the separation is identified as a permanent separation.

are identifiable in the LWF-Census file. For instance, in this study, each variable in these three blocks in the model is replaced with a similar variable or a proxy variable with a similar concept from the LWF-Census file. For example, for the "Experience (EXP)" block, instead of the number of months that cohort members worked, total years of work experiences are used instead. For the "Family (FAM)" block, marital status is used instead, which contains several binary variables rather than one binary variable on the current presence of a partner. Additional control variables identified by Del Bono and Vuri (2011), such as occupation, are also included. Ultimately, the objective of this study is to investigate the composition of the gender wage gap through wage growth, in terms of labour market trajectories based on the structure model of Joshi, Bryson, Wilkinson, and K. Ward (2019) and Del Bono and Vuri (2011).

3.5 Data & Methodology

3.5.1 Data source

This study has three main data sources: The Longitudinal Workers File (LWF), the 2006 and 2016 Census long-form, along with the special LWF-Census linkage weight file.

The LWF is constructed by integrating data from four sources: the T1 and T4 file from Canada Revenue Agency (CRA), the Record of Employment (ROE) files from Employment and Social Development Canada (ESDC) and the Longitudinal Employment Analysis Program (LEAP) from Statistics Canada (StatCan).[2] The LWF contains rich information on job-level information, which is drawn from T4 records. This includes province of employment, employment income, employment insurance premiums, employment insurance income (from T4E file), pension adjustment amount, and union dues. In addition, the LWF includes ROE information on the reason for job termination, which is essential for identifying labour market trajectories. Information on pension adjustment amount and union dues are also available in the LWF drawn from T1 personal tax files, which will be used as a block in the regression analysis.

The LWF is a key data source that can identify labour market trajectories for the same individual between 1989 and 2016. For instance, the LWF distinguishes between various types of employee separations, allowing distinctions to be made between layoffs, quits and other separations. This is done through the ROE, which specifies the reason for the work interruption or separation. The LWF also allows a distinction between temporary and permanent separations. In the LWF, a temporary separation is identified as an employee returning to his or her employer during the year of the separation or in the following year. When such a return does not occur, the separation is identified as permanent separation instead. Note that according to Chan, Davidson, T. Qiu, and Morissette (2019), job end occurring in year t followed by reemployment in the same year is not counted as a separation. When a job ends (i.e. when an employer-employee pairing observed in year t is no longer observed in year t+1), employers are generally required to issue a Record of Employment (ROE) (Chan, Davidson, T. Qiu, and Morissette, 2019). This requirement holds whether or not the employees leaving

[2]Please refer to the Data Flow Chart in the appendix (D.1) for details regarding the visual representation of the construction of the analytical data file.

the firm intend to claim EI benefits. Hence, one would expect the majority of job endings to be associated with an ROE and thus a reason for the separation. In this study, for consistency purposes, the same definition of job separation is adapted from the LWF. It is worth noting that individuals could have a separation without ROE which the LWF would not capture.

There are certain limitations to the LWF file. First, none of the data sources used to construct the LWF contain information on workers' educational attainment or occupation, and hence these variables are not available on the file. Second, information is also not available on the number of hours individuals work each week nor on the number of weeks worked each year. As a result, weekly earnings or hourly earnings cannot be computed. Only annual income can be measured using workers' annual employment income (from T4 records) and workers' net income from self-employment (from T1 records).

To overcome these limitations in the LWF, the 2006 and 2016 Census long-form are merged with the LWF file in this study for the years 2005 and 2015.[3] The Census long-form provides a snapshot of key information such as educational attainment, weeks worked, and occupations that are not available on the LWF. The Census is a nationwide survey that provides a statistical portrait of the country and its people every five years. In Canada, it is mandatory for all residents to participate in the Census. The Census long-form contains additional questions besides the basic demographic questions asked in the short-form questionnaire, which are crucial for this study. In 2006, 20% of Canadian households received a long-form questionnaire; in 2016, this number increased to 25%.

Finally, it is worth mentioning that not all Census respondents are observed/matched in the LWF, and since the original Census long-form weight does not account for the probability of individuals being observed in the LWF, it cannot be used directly as the weight for estimation in the LWF-Census linked file. For this purpose, in February 2020, Statistics Canada created the very first version of the 2006 LWF-Census linkage weight file, which takes into account the probability of individuals being observed in 2006 Census long-form and the LWF at the same time. Although it does not account for the probability that individuals are observed at the same time in the LWF and both 2006 and 2016 Census long-form, it is the best alternative at the time of this study. A revision using the new version of the LWF-Census linkage weight is possible in the near future.

3.5.2 Sample selection

The population of interest for this study consists of employed individuals aged 25 to 34 (young workers) in 2005, as individuals within this age group are more likely to be out of school and working full-time. In addition, they have to be observable on both the 2006 and 2016 Census long-form and also identifiable on the LWF file at the same time. The main reason for this rather strict restriction is to allow the tracking of one's weekly wage, occupations, and other key control variables such as level of education from the Census, which are not available on the LWF file alone. It is worth mentioning that according to (Robson, 2017), there was no significant change to the paternity leave policies during the sample period between 2005 and

[3]The reference period of the Census long-form is the year prior to the day where the survey is conducted. There was no long-form census in 2011, and it was temporarily replaced by National Household Survey (NHS). In this study, NHS is not used due to the fact it was not a mandatory survey.

2015, which helped to it avoid a potential confounding factor that could potentially introduce bias to the results.

To construct the analytical dataset, first, the Census is linked to the LWF key files separately for the years 2006 and 2016 to retrieve the longitudinal person identifier. Note that although, in theory, 100% of the Census respondents should be observed in the LWF, however during the data processing and matching procedures, some individuals in the Census cannot be accurately identified longitudinally in the LWF and, therefore, were dropped. 55% of Census long-form respondents are matched to the LWF in 2006 and 78% in 2016. Next, the Census 2006 is linked to the Census 2016 using the LWF's longitudinal person identifier "casenum2019", which is a unique id that keeps track of the same individual across Census and different years among the LWF file. Finally, the LWF job level (T4 and T4E) files and person-level (T1) files, along with the LWF-Census weight file, are merged together with the linked Census file to construct the final analytical dataset.

Note on the LWF file; individuals could hold multiple jobs in a given year recorded by T4. Identifying labour market trajectories for the same individual for all jobs in a given year across the ten-year period is exceedingly difficult. Therefore, for the purpose of simplicity, only the main job, which is defined as the job where individuals had the highest employment income from the LWF file in a given year, is kept.[4] In addition, only individuals with a complete record of T1 and T4 or T4E are kept in the sample to ensure that certain key control variables, such as pension contribution and union dues, are in place. In the end, the final sample in this study contains 110,430 unique individuals, which are about 22.37% of the young workers (weighted) in the 2006 Census long form. The descriptive statistics analysis is conducted to focus on this sample. For the regression analysis, only individuals with valid values for control variables in the regression analysis are included. The final sample size for the regression analysis is 6,002 unique individuals (2,710 female and 3,292 male), which is about 1.12% of the young workers (weighted) in the 2006 Census long form.[5] For the complete summary statistics of the selected sample, please refer to Table D.1 in Appendix D.

3.5.3 Summary Statistics

The summary statistics of the analytical dataset are presented in Table D.1 in Appendix D. The control variables in the Education (ED), Experience (EXP), Family (Fam) and Occupation (OCCU) blocks are all derived from the 2006 and 2016 Census long-form. Among the selected sample population in 2016, 19.72% of the individuals have a high school equivalency certificate without further schooling. Among women, 22.17% have at least a high school equivalency certificate compared to 22.72% of men, which is quite similar. As for work experience, women had an average of 11.27 years of work experience in 2015 since the first day they started working, which was slightly higher compared to men with 10.92 years of total work experience. In terms of family status, overall, 55.66% of individuals are married, and 70.60% of all individuals are the parent of another household member. 54.79% of women were married, slightly lower than 56.50% of men. Regarding the presence of children, 74.64% of

[4] On average, there is no significant difference between women and men on the likelihood of holding multiple jobs in a given year for the selected sample.

[5] The significant reduction in sample size is due to the inconsistent or missing values on the 2006 and 2016 Census in education, work experiences and occupation.

women observed in 2015 had a least one child present in the household, which is much higher than 66.70% of men. Women were more likely than men to be working in the business, finance and administration occupation (23.28% versus 9.63%, respectively). Similarly, women were twice as likely to work in education, law and social, community and government services than men (20.64% versus 9.18%, respectively). On the other hand, women were much less likely to work in trades, transport and equipment operators and related occupations than men (1.55% versus 23.89%). These drastic occupational differences between women and men could be one of the reasons behind the fact that 54.20% of women were unionized compared with 48.45% of men. In addition, women were also more likely to have a job with a pension plan than men (63.37% versus 59.23%), perhaps because women were more likely to work in government services. Table D.2 contains the summary statistics of the analytical dataset after applying the sample restriction for the regression analysis. Overall, the results are quite similar to Table D.1. with minor differences.

3.5.4 Methodology

Regression analysis

The basic model specification is adapted from Joshi, Bryson, Wilkinson, and K. Ward (2019) with minor modifications suggested by Del Bono and Vuri (2011), which are not available in their study.[6] Similar to Joshi, Bryson, Wilkinson, and K. Ward (2019), the coefficient estimates in this equation are estimated separately for each gender. The complete list of variables is listed in the Technical Annex in Appendix A.2. The purpose of equation (3.1) is to investigate how gender differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason of separation, could impact wage growth, given the base model specification from Joshi, Bryson, Wilkinson, and K. Ward (2019) and Del Bono and Vuri (2011). Note that unlike Joshi, Bryson, Wilkinson, and K. Ward (2019) hours worked is not observed in the LWF and the Census long-form and therefore are omitted. In this case, since weeks worked are highly correlated with the instance of separations and therefore are viewed as outcomes in the model. As a result, I decided not to include them as control variables (bad control). For the complete summary statistics of the selected sample for the regression analysis, please refer to Table D.2 in Appendix D.

The specification of the regression analysis of equation (3.1) is given by:

$$W_i = \alpha_0 + ED'_i\beta + EXP'_i\gamma + FAM'_i\delta + OCCU'_i\epsilon + Other'_i F + TP'_i\zeta + CEO'_i\eta + RS'_i\theta + u_i \quad (3.1)$$

W_i here is the main dependent outcome variable. It is the weekly wage annual growth rate between 2005 and 2015, which is defined as (ln(weekly wage 2015)-ln(weekly wage 2005)/10). This would give us ln (1+r (real weekly wage growth)), which is approximately r (when r is close

[6] As mentioned in the framework section, certain variables used in Joshi, Bryson, Wilkinson, and K. Ward (2019) are substituted with close proxy or variables with similar concepts in the Census. Please refer to the Technical Annex for detail.

to zero).[7] This definition is in line with The Ontario Inclusive Innovation (i2) Action Strategy.[8] $Education(ED)_i$ is a vector that contains information on individuals' highest level of education achieved by 2015, which is derived from the 2016 Census long-form. $Experience(EXP)_i$ is a vector that contains information on individuals' total length of work experiences by 2015, derived from the first year where individuals started working according to the LWF. Note here the work experiences are not job-specific but rather a total length calculated from the first day individuals started working. Inclusion of the total length of work experiences could reveal returns to accumulated human capital or job-specific skills. $Family(FAM)_i$ is a vector that contains various information on individuals' family status in 2005 and in 2015 derived from the 2006 and 2016 Census long-form. The variables in the vector include marital status, a binary measurement of parent flag in 2015 and if there were the presence of children by age category in 2005.[9] $Occupation(OCCU)_i$ is the occupational fixed effect of individuals in 2015, which includes the occupational dummies, binary measurement on ever paid union fees, and binary measurement of if ever had a contribution to the registered pension plan or deferred profit-sharing plan between 2005 and 2015.[10] The purpose of adding this extra block is to capture the impact of the quality of jobs on wage growth, which is not available in Joshi, Bryson, Wilkinson, and K. Ward (2019)'s original data source. Note here that firm size is excluded from the model due to a large number of missing values, which would reduce the sample size by more than half. $Other_i$ is a vector that contains other control variables that did not below any of the blocks above but are considered to be important either by Joshi, Bryson, Wilkinson, and K. Ward (2019) or the context in this study. The vector includes a continuous age variable, age square, number of times each individual appears in the LWF file over ten years and lastly, the province of residence of individuals in 2015 according to the Census long-form. Note in Joshi, Bryson, Wilkinson, and K. Ward (2019), the regression analysis is carried out for different age groups and thus, age is omitted from their model. However, in this study, since the age focus is rather narrow (aged 24-35 in 2005), age is included as a control variable in the regression analysis instead.

The main blocks of interest in equation (3.1) are $Type\ of\ separation(TP)_i$, $Changed\ employer\ or\ occupation(CEO)_i$, and $Reason\ for\ separation(RS)_i$, which in turn refers to the block of type of separation, changed employer or occupation and reason for job separation. All variables in these blocks are derived from the LWF between 2005 and 2015 for each individual. The type of separation blocks includes the total number of separations, the total number of permanent separations and the total number of temporary separations. The changed employer or occupation block includes a binary measure on whether individu-

[7]For each individual, the weekly wage is computed using the total annual employment income divided by the number of weeks worked in 2005 and 2015. Ideally, an hourly wage would be preferred; however, due to data limitation on hours worked, only weekly wage can be computed using weeks worked in the reference month from the Census.

[8]The Ontario Inclusive Innovation (i2) Action Strategy is a joint initiative between the Telfer Centre for Executive Leadership and the Ryerson Diversity Institute to support research and projects in the field of Gender-based analysis (GBA+) and diversity.

[9]The variable for the presence of children by age category is not available on the 2016 Census long-form.

[10]Similar to weekly wage, occupation information is only available on the 2006 and 2016 Census. While union status and pension are available on a yearly basis from the LWF. For the purpose of regression analysis, a single binary variable for union status and pension status is created from the original LWF variables, which are longitudinal.

als switched employer between 2005 and 2015 and a binary measure on whether individuals switched occupation between 2005 and 2015. The reasons for separation block include the total number of permanent separations due to layoff, the total number of temporary separations due to layoff, the total number of permanent separations due to quit, total number of temporary separations due to quit, the total number of permanent separations due to parental/maternity leave and lastly, total number of temporary separations due to parental/maternity leave.

For equation (3.1), the estimation is carried out gradually in three different specifications separately for women and men, starting with the base model (including control variables) with only the type of separation block. Then other blocks on changed employer or occupation and reasons for job separation are added once at a time to see how these coefficient estimates change.

In addition to equation (3.1), a separate regression analysis based on equation (3.2) is also estimated given by:

$$W_i = \iota_0 + \iota_1 Female_i + ED'_i\kappa + EXP'_i\lambda + FAM'_i\mu + OCCU'_i\nu + Other'_i\xi + TP'_i\emptyset + CEO'_i\pi + RS'_i\rho + e_i \tag{3.2}$$

The purpose of equation (3.2) is to investigate by adding blocks on early career labour market trajectories, including the type of separation, changed employer or occupation and reason of separation into the model, what would happen to the coefficient estimates of the female dummy. For instance, does the coefficient become smaller and less statistically significant? Here the binary variable "Female" is the main variable of interest. "Female" is defined as a binary variable equal to 1 if the respondent's gender is female and 0 otherwise, derived from the 2005 Census long-form.

Like equation (3.1), equation (3.2) is also carried out gradually, but in twelve different specifications, starting with the basic point estimation just with the female dummy variable. Then each block is added once at a time to see how the coefficient estimates of the female dummy change. Following the regression analysis, the gap in weekly wage annual growth rate observed in the raw data is decomposed between men and women using Oaxaca-Blinder decomposition based on equation (3.2). As explained in the conceptual and methodological overview on the gender pay gap by Moyser (2019), the Oaxaca-Blinder decomposition divides the gender gap in weekly wage annual growth rate into the part associated with individual attributes and the part associated with unequal coefficients and any interaction between them. In other words, allowing the calculation of "how much of the gender gap is accounted for?" by each variable in the regression model, including the blocks on type of separation changed employer or occupation and reason for job separation.

Similar to Joshi, Bryson, Wilkinson, and K. Ward (2019) and Goldthorpe (2001), the purpose of the regression analysis is not to recover causal estimates of the influence of labour market trajectories on wage growth; rather, the estimates offer an accounting exercise to map the correlates of unequal pay growth over the early career period and to establish empirical regularities. Therefore, the possibility of omitted factors should be borne in mind when interpreting the results. All analyses presented in the following sections are weighted using the 2006 LWF-Census linkage weight.

3.6 Results

3.6.1 Descriptive statistics

The main purpose of the descriptive statistics analysis is to investigate the gender differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason for job separation for young workers. Interaction between these dimensions and weekly wage annual growth rate is also analyzed in this section.[11]

Type of separation

This section provides descriptive statistics analysis of the gender differences in early career labour market trajectories by type of separation and its interaction with the weekly wage annual growth rate between women and men.

According to Figure C.1, overall, women were less likely than men to experience a job separation, though the difference was small and not statistically significant. A job separation is defined as a separation from the employer, regardless of whether the employee returns to the employer during the year of the separation or in the following year. In general, there is a decreasing trend in job separation over time for both women and men. This is consistent with previous studies based on the American labour market that show that older workers are less likely than younger workers to experience job transitions (Haltiwanger, H. Hyatt, and McEntarfer, 2018).

While the differences in overall job separations between women and men were minor among workers who had a job separation between 2005 and 2015, conditional on having a separation, women were less likely than men to experience a permanent job separation (i.e., a separation where employees do not return to their employer during the year of the separation or in the following year), where the difference are statistically significant (see Figure C.2). Most of the existing literature does not differentiate between permanent and temporary job separations but instead considers moves to different jobs or moves from employment to non-employment. Therefore, a direct comparison with other studies is difficult. Generally speaking, the existing literature based on the European labour market indicates that women are less likely to make job-to-job moves (Hirsch and Schnabel, 2012; Eryar and Tekgüç, 2014), which is consistent with the trend observed here. More specifically, women were more likely than men to stay with the same employer, even after accounting for the number of workers who moved from employment to non-employment. There may be many reasons behind this phenomenon—including gender-based factors such as family responsibilities, including housework and childcare (Fuller, 2008; Bertrand, Goldin, and Katz, 2010) or other job-specific characteristics, such as unionization (Hirsch and Schnabel, 2012).

According to Figure C.3, for women, as the total number (instance) of permanent job separations increases, the weekly wage annual growth rate decreases, while the same is not observed for men. For young workers who worked continuously with no permanent job separation

[11] The main focus of the paper is to look at the gender wage gap through wage growth; therefore, the variables on labour market trajectories, including the type of separation, changed employer or occupation and reasons for separation, are not included as dependent/outcome variables in the regression analysis.

between 2005 and 2015 (point of "0" on the x-axis), women and men's weekly wage annual growth rate was almost at the same level with a minor difference. On the other hand, on average, men who had a permanent job separation have a higher weekly wage annual growth rate than those who do not (points between "1" and "7" on the x-axis). At the same time, for those who had labour market interruptions (points between "1" and "7" on the x-axis), women's weekly wage annual growth rate was lower than men's at all levels of the total number (instance) of permanent separations where the differences are statistically significant at the 5% level beyond the total of two instances. These findings are consistent with existing literature such as Del Bono and Vuri (2011) that wage growth is much lower for women who move than those who do not in the first ten years. It is worth mentioning that unlike the results shown here for men, in the regression results presented in Table C.15, where in the third model specification with all the control variables in place, the sign of the coefficient estimates of the total number of permanent separations are negative for both women and men, which suggest a negative correlation between permanent separation and wage growth for men.

Previous studies based on the European labour market have found that women faced a higher penalty for taking time out of the labour market compared with men (Manning and Swaffield, 2008). The results of this study support this finding. According to Table C.1, among women (comparison by column), on average, the weekly wage annual growth rate was 0.84 percentage points higher for those who had no permanent job separation than those who moved jobs. The opposite applies to men, with a difference of 0.47 percentage points. Comparing across gender (comparison by row), however, for those who had no permanent job separation, women's weekly wage annual growth rate was higher than men by 0.70 percentage points, while this was the opposite for individuals who had permanent job separations. In addition, for those who had a permanent job separation, women had a lower weekly wage annual growth rate than men by 0.61 percentage points. These results suggest that compared with men, women were penalized for having a permanent job separation. One possible explanation is that women are less inclined to make "wages increasing" voluntary job-to-job moves compared with men, such as job-hopping (Hirsch and Schnabel, 2012).

In summary, according to the results presented above, although women in their early career were less likely to experience job separation and permanent job separation compared with men, their weekly wage annual growth rate was at a disadvantage. More specifically, women had a lower weekly wage annual growth rate compared with men by the total number of permanent job separations and status of permanent job separation. This means that at the beginning of their career, women had a greater disadvantage for taking time out of the labour market due to permanent job separation compared with men. This evidence points out the possible existence of the gender gap in wage growth caused by permanent job separation, which could be a potential channel that ultimately contributed to the unexplained wage gap between women and men.

In contrast to permanent job separation, according to Figure C.4, among individuals that had a job separation between 2005 and 2015, women were more likely than men to experience a temporary job separation (i.e., a separation where an employee returning to his or her employer during the year of the separation or in the following year) with statistically significant differences. More specifically, compared with men, women were more likely to leave their employer temporarily. Again, since most of the existing literature does not differentiate job separation into permanent and temporary, a direct comparison with other studies is difficult. Especially in this case, the definition of temporary job separation derived from the LWF is unique, as it captures the employer-level mobility change (return to the previous employer or not) instead of voluntary job-to-job moves that could potentially be beneficial to wage growth (Hirsch and Schnabel, 2012).

Unlike the results discussed above on permanent job separation, as the total number of temporary job separation increases, the weekly wage annual growth rate increase for both women and men (see Figure C.5). For young workers who worked continuously with no temporary job separation between 2005 and 2015 (point"0" on the x-axis), women had a lower weekly wage annual growth rate than men. Although the differences were minor and not statistically significant. On average, both women and men who had a temporary job separation (points between "1" and "8" on the x-axis) have higher weekly wage annual growth rates than those who do not, though the differences were minor and not statistically significant. These results are consistent with previous findings by Abbott and Beach (1994) and Hahn, H. R. Hyatt, Janicki, and Tibbets (2017), where Canadian women who changed jobs in 1986 had short-run wage gains, and in general, job-to-job moves have a modest positive effect on earnings. Compared with permanent job separation, the difference in weekly wage annual growth rate by the total number of temporary separations was rather stable and minor across gender.

Similarly, different than the results discussed above on permanent job separation and wage growth, women who had no temporary job separation between 2005 and 2015 had a lower wage growth than those who did. More specifically, according to Table C.2, among women (comparison by column), on average, the weekly wage annual growth rate was 0.74 percentage points lower for those who had no temporary job separation compared with those who had a separation. The same applies to men with a smaller difference of 0.10 percentage points, although the differences are not statistically significant. Comparing across gender (comparison by row), for those who had no temporary job separation, women had a lower weekly wage annual growth rate than men by 0.76 percentage points, while the same applies for individuals who had a temporary job separation. Among those who had a temporary job separation, women's weekly wage annual growth rate was lower than men by 0.12 percentage points. Unlike the case of permanent job separation, the differences in wage growth by gender are not statistically significant; however, the point estimates are not small, which is likely due to the fact that a smaller number of individuals experience temporary separations. In sum, both women and men were slightly better off as a result of temporary job separations, although women benefited slightly less compared to men.

In summary, according to the results presented above, women in their early careers were more likely to experience a temporary job separation compared with men, which is the opposite of the case for permanent job separation. However, their weekly wage annual growth rate was still at a disadvantage. More specifically, women had a slower growth rate on log weekly wage by the total number of temporary job separations compared with men and a lower level of weekly wage annual growth rate by the status of permanent job separation. These results reinforce the discussion from the permanent job separation listed above, where a gender gap in wage growth could exist that is caused by temporary job separation, which could also be a potential channel in addition to permanent job separation that ultimately contributed to the unexplained wage gap between women and men.

Changed employer or occupation

This section provides a descriptive statistical analysis of the gender differences in the early career labour market trajectories, including change in employer and occupation and their interaction with weekly wage growth between women and men.

During the period between 2005 and 2015, women in their early careers were less likely than men to switch employers. More specifically, according to Table C.3, women were less likely to switch employers compared with their male counterparts overall for the ten years, but also the first five years between 2005 and 2010 and the last five years between 2010 and 2015. For both women and men, there was a decreasing trend in the number of young workers who switched employers by age, which is consistent with previous findings by Hirsch and Schnabel (2012). Another interesting highlight is that the differences between women and men were minor at the beginning of their careers but gradually increased as age increased. This suggests that, unlike men, women tend to be more likely to stay with the same employer, and the likelihood of switching employers decreases at a faster rate than their male counterparts as age increases.

Table C.4 contains the weekly wage annual growth rate of young workers aged 24 to 35 in 2005 by whether or not individuals changed employer and gender between 2005 and 2015. When comparing among gender (comparison by column), the weekly wage annual growth rate was 0.73 percentage points lower for those who stayed with the same employer than those who switched employers between 2005 and 2015. The same applies to the men with a larger difference of 0.83 percentage points. Comparing across gender (comparison by row), for those who stayed with the same employer, women had a lower weekly wage annual growth rate than men by 0.08 percentage points, while the same applies for individuals who switched employers. For instance, among those who switched employers, women's weekly wage annual growth rate was lower than men by 0.18 percentage points. Both women and men were better off as a result of switching employer; however, women still had lower weekly wage growth compared with men and women who switched employer was still worse off compared with their male counterparts.

According to Table C.5, between 2005 and 2015, women in their early careers were slightly more likely than men to switch occupations, though the differences were minor. Note that like weekly wage, occupation information is only available on the 2006 and 2016 Census. The lack of variation here on the gender difference in change of occupation is in line with Manning and Swaffield (2008), where they find that occupational differences are minor and cannot explain much of the gender gap in early-career wage growth.

Table C.6 contains the weekly wage growth of young workers aged 24 to 35 in 2005 by the status of occupation and gender between 2005 and 2015. When comparing among gender (comparison by column), the weekly wage annual growth rate for those who stayed within the same occupation was 0.51 percentage points higher than those who changed occupations between 2005 and 2015. The same applies to men, with a slightly lower difference of 0.48 percentage points. Comparing across gender (comparison by row), for those who stayed within the same occupation, women had a lower weekly wage annual growth rate than men by 0.11

percentage points. Similarly, among those who changed occupation, women's weekly wage annual growth rate was lower than men by 0.14 percentage points. Both women and men were better off if they stayed within the same occupation; however, women had a lower increase in weekly wage growth compared with men and women who stayed within the same occupation group since 2005 were worse off compared with men.

In summary, according to the results presented above, women in their early careers were less likely to switch employers compared with men but were slightly more likely to change occupations. At the same time, women's weekly wage annual growth rate was at a lower level compared with men regardless of changes in employer or occupation. For instance, among those who stayed with the same employer, women's weekly wage annual growth rate was lower than men's. Similarly, for those who switched employers, women's weekly wage annual growth rate was still slightly less compared with men. The same applies to the occupation, which for both groups of individuals who either stayed or switched occupations, women's weekly wage annual growth rate was lower than men.

As part of the questions, the following section provides a more detailed descriptive statistical analysis of the gender differences in early career labour market trajectories, including change in occupation for those who stayed with the same employer and those who switched employer, and their interaction with weekly wage annual growth rate between women and men.

According to Table C.7, among those who stayed with the same employer between 2005 and 2015, women were more likely than men to switch occupations. This result is consistent with the overall gender difference in the likelihood of changing occupations discussed in Table C.6. Table C.8 contains the weekly wage growth of young workers aged 24 to 35 in 2005 who stayed with the same employer by the status of occupation and gender between 2005 and 2015. When comparing among gender (comparison by column), among women, for those who stayed with the same employer and the same occupation, their weekly wage annual growth rate was 0.31 percentage points higher than those who stayed but changed occupation between 2005 and 2015. The same applies to men with a slightly large difference of 0.32 percentage points. Comparing across gender (comparison by row), for those who stayed with the same employer and the same occupation, women had a lower weekly wage annual growth rate than men by 0.05 percentage points, while the same applies for individuals who stayed but changed occupation. For instance, among those who stayed with the same employer but changed occupation, women's log weekly wage growth was lower than men by 0.04 percentage points. These results are consistent with the overall gender difference in weekly wage annual growth rate by occupation status discussed in Table C.6. Again, among individuals who stayed with the same employer, both women and men were better off if they stayed within the same occupation. However, women benefited slightly less compared with men and women who stayed with the same employer that also stayed within the same occupation group since 2005 were worse off compared with their male counterparts.

According to Table C.9, among those who switched employers between 2005 and 2015, women were less likely than men to switch occupations. This result is again consistent with the overall gender difference in the likelihood of changing occupation discussed in Table C.5 and also with those who stayed with the same employer that changed occupation discussed in

Table C.7. Compared with results shown above in both Table C.5 and Table C.7, the variation here between women and men were larger, which suggest that compared with their male counterparts, women were even more likely to switch occupation when they switch employer.

Table C.10 contains the weekly wage annual growth rate of young workers aged 24 to 35 in 2005 who switched employer by the status of occupation and gender. When comparing among gender (comparison by column), for women, those who switched employer but stayed within the same occupation, their weekly wage annual growth rate was 0.42 percentage points higher than those who switched employer and also changed occupation between 2005 and 2015. The same applies to men, with a slightly larger difference of 0.44 percentage points. Comparing across gender (comparison by row), for those who switched employer but stayed within the same occupation, women had a lower weekly wage annual growth rate than men by 0.19 percentage points. Similarly, among those who switched employers and also changed occupation, women's weekly wage annual growth rate was lower than men by 0.17 percentage points. These results are consistent with the overall gender difference in weekly wage annual growth rate by occupation status discussed in Table C.6 and with those who stayed with the same employer discussed in Table C.8. Again, for those who switched employers, both women and men were better off if they stayed within the same occupation. However, women benefited slightly less compared with men and women who switched employers but stayed within the same occupation group since 2005 were worse off when compared with men.

In summary, according to the results presented above, women in their early careers who stayed with the same employer were slightly more likely to change occupation compared with men. On the other hand, among individuals who switched employers, women were less likely than men to switch occupations. At the same time, similar to the overall gender gap discussed in the previous section, women's weekly wage annual growth rate was at a lower level compared with men regardless of changes in employer or occupation. For instance, among those who stayed with the same employer and also the same occupation, women's weekly wage annual growth rate was lower than men's. Similarly, for those who stayed with the same employer but changed occupation, women's weekly wage annual growth rate was still less compared with men. The same applies to those who switched employers, regardless of changing occupation or not.

Reason for job separation

This section provides descriptive statistics analysis on the gender differences of the reason for job separation and their interaction with weekly wage growth between women and men.

In addition to the type of job separation, the reason for separation is also an important factor that could affect the gender wage gap that still remains mostly unexplored (Manning and Swaffield, 2008). Table C.11 contains the total number (instances) of permanent job separation of young workers aged 24 to 35 in 2005 by reasons of separation and gender between 2005 and 2015. In Canada, between 2005 and 2015, among individuals who had a permanent job separation, "quit" and "layoff" was the leading reason for job separation for both women and men. However, women were less likely than men to experience permanent job separation due to layoffs and quitting. On the other hand, women were more likely than men to experience permanent job separation due to parental or maternity leave.

Previous studies based on the Canadian Labour Market Activity Survey 1986-87 have found that overall, both women and men who changed jobs in 1986 realized short-run wage gains; however, women who were laid-off or quit for personal reasons had substantially greater wage losses than men (Abbott and Beach, 1994). The results of this study support this finding with more recent data. According to Table C.12, among all the reasons for permanent job separation listed above recorded in the Record of Employment (ROE), women's weekly wage annual growth rate was lower compared with men. For instance, for layoff and quit, women's weekly wage annual growth rate was 0.49 percentage points and 0.80 percentage points lower than men, respectively. Similarly, women who had a permanent job separation due to parental or maternity leave had a lower weekly wage annual growth rate compared with men by 0.42 percentage points. Refer to back to results in Table C.11, although women were less likely than men to experience a permanent job separation due to layoff and quit, their weekly wage annual growth rate was still at a lower level compared with their male counterparts. This means that for young workers in Canada, women had a greater disadvantage for taking time out of the labour market compared with their male counterparts, regardless of the reason for separation.

In summary, according to the results presented above, although the leading reasons for permanent job separation were the same for both women and men. Women's weekly wage annual growth rate was less compared with men by all reasons for permanent separation listed above. This means that during the period of early career, women had a greater disadvantage for taking time out of the labour market compared with their male counterparts due to permanent job separation, regardless of the reason for job separation. These results reinforce the evidence from the type of separation session discussed above where a gender gap in wage growth is observed in permanent job separation by all reasons of separation, which could be a potential channel that ultimately contributed to the unexplained wage gap between women and men.

Table C.13 contains the total number (instances) of temporary job separation of young workers aged 24 to 35 in 2005 by reasons and gender between 2005 and 2015. For those who had a temporary job separation, "layoff" and parental or maternity leave were the leading reasons for job separation for women. On the other hand, "layoff" and other reasons were the leading reasons for men. Compared with permanent job separations, the story for temporary job separations is quite different. In particular, instead of "quit" being one of the top reasons for job separation, parental or maternity leave was listed instead. Also, for men, instead of "quit," other reasons were listed as one of the top reasons for job separation. Comparing across gender, women were less likely than men to experience a temporary job separation due to "layoff" and other reasons. However, women were much more likely than men to experience a temporary job separation due to parental or maternity leave.

According to Table C.14, unlike the case of permanent job separation, compared with men, women's weekly wage annual growth rate was lower only for the reason of separation by "layoff" and "quit." On the other hand, for the reason of parental or maternity leave and other, women's weekly wage annual growth rate was higher than men's. For instance, for parental or maternity leave and other reasons, women's weekly wage annual growth rate was 0.75 percentage points and 0.02 percentage points higher than men, respectively. Compared with the case of permanent job separation, women who experienced temporary job separation due to parental or maternity leave and other reasons had null and positive effects on their weekly wage annual growth rate compared with men. Referring back to results discussed above in Table C.13, although women were more likely to have a temporary job separation due to parental or maternity leave and other reasons compared with men, they were not penalized by taking time out of the labour market, which is the opposite of the case of permanent job separation. A possibility is that permanent separations due to parental leave are more likely to be accompanied by job changes leading to lower wage growth, as opposed to women having a temporary separation. Future work could look at wage growth penalties associated with maternal leave by occupation, accounting for the distinction between permanent and temporary separations. Lastly, for the reason of "layoff" and "quit," women still had a lower level of wage growth compared with their male counterparts, which is similar to the case of permanent job separation.

In summary, according to the results presented above, layoff and parental or maternity leave were the leading reason for temporary job separation for women. On the other hand, layoff and others were the leading reasons for men. However, women who had temporary job separation due to layoff and quit were still worse off compared with men but were in a slightly better situation if the separation was related to parental or maternity leave and other reasons. Again, these results reinforce the evidence from the type of separation session discussed above where a gender gap in wage growth is observed in temporary job separation by reasons of separation, which could also be a potential channel that ultimately contributed to the unexplained wage gap between women and men.

3.6.2 Regression

As discussed in the framework and methodology section, the model specification for the regression analysis is adapted from Joshi, Bryson, Wilkinson, and K. Ward (2019) and Del Bono and Vuri (2011), with new blocks added on type of separation, labour market trajectories and reason for job separation.

Recall that the purpose of equation (3.1) is to investigate how gender differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason for job separation, could impact wage growth, given the base the gender wage gap model specification from Joshi, Bryson, Wilkinson, and K. Ward (2019). Table C.15 contains the regression results estimated using euation (3.1). As shown in Table C.15, equation (3.1) is estimated separately for women and men in three different model specifications. In the first model specification, only the type of separation variables is included in addition to the base model. For the next two specifications, variables on labour market trajectories and reasons for separation are gradually added to the base model.

For women, among all three specifications, there is a negative correlation between the total number of temporary separations and the weekly wage annual growth rate. For instance, according to the third model specification, one additional instance of temporary job separation is associated with a decrease of the weekly wage annual growth rate for women by 0.2 percentage points. At the same time, compared with those who did not switch employers, those who did have a higher weekly wage annual growth rate by 1.1 percentage points. Both of these coefficient estimates are statistically significant at the 10% confidence interval level.

Men, on the other hand, were more affected by permanent separations instead of temporary separations. For instance, among all three specifications, there is a negative correlation between the total number of permanent separations and the weekly wage annual growth rate. According to the results in the third specification, one additional instance of permanent job separations decreases the weekly wage annual growth rate for men by 0.6 percentage points. Similar to women, men who switched employers compared with those who did not have a higher weekly wage annual growth rate by 0.7 percentage points. In addition, men who changed occupations have a lower weekly wage annual growth rate by 0.7 percentage points compared with those who didn't, where the same is not observed for women. These coefficient estimates are statistically significant at 5%, 10% and 10%, respectively.

The regression results here in equation (3.1) suggest that women's weekly wage annual growth rate is more sensitive to temporary job separation compared with men's, where men's weekly wage annual growth rate is more sensitive to permanent job separation. These results are in line with the descriptive statistics analysis in the previous sections, where women were more likely to experience a temporary job separation, and men were more likely to experience a permanent job separation.

Table C.16 contains the regression results based on equation (3.2). Recall that the purpose of equation (3.2) is to investigate the impact of early-career labour market trajectories, including the type of separation, changed employer or occupation and reason of separation on the female dummy based on the model specification from Joshi, Bryson, Wilkinson, and K. Ward (2019). As shown in Table C.16, part I and part II, equation (3.2) is estimated using twelve different model specifications. In the first six model specifications from Table C.16 part I, each block of control variables is added gradually into the model. The purpose of this is to attempt to reproduce the results from Joshi, Bryson, Wilkinson, and K. Ward (2019)'s original model. From the seventh column to the eleventh column in Table C.16b part II, blocks on type of separation, including the type of separation, changed employer or occupation and reason for job separation, are gradually added into the model. The last model specification in column twelve includes only parental/maternity from the reason for the job separation block. The purpose of this last model specification is to test out the sole impact of the reason for job separation due to parental/maternity from all the other reasons and blocks on the female dummy in respect to log wage growth.

According to the first six model specifications, there is a negative correlation between being women and the weekly wage annual growth rate compared with men. For instance, according to the sixth model specification, women's log wage growth is 0.9 percentage points lower compared with their male counterparts, which is statistically significant at the 5% level. These results are consistent with Joshi, Bryson, Wilkinson, and K. Ward (2019), where the coefficient estimates of the female dummy remain statistically significant conditional on all the blocks of control variable.[12]

In the next four model specifications, when the blocks on labour market trajectories, including the type of separation, changed employer or occupation, are added to the model, the coefficient estimates of the female dummy remain statistically significant with a similar mag-

[12]Note that this is true even after controlling for occupation, union and pension status which is absent in Joshi, Bryson, Wilkinson, and K. Ward (2019)'s model.

nitude of 0.8%. However, in the eleventh column, where the last reason for job separation on parental/maternity leave is added to the model, the coefficient estimate of the female dummy decreases to 0.5% and is no longer statistically significant. In other words, In specification 12, where variables on type of separation, labour market trajectories and all other reasons for separation are absent from the model except for a reason for job separation due to parental leave, the coefficient estimate of the female dummy is exactly the same to specification 11 where all variables are in place. Such a finding suggests that parental leave alone is sufficient to explain the significant differences observed in weekly wage growth by gender (the female dummy). This means that conditional on observables from all the blocks and reasons above, differences in weekly wage annual growth rate associated with the reason for job separation on parental/maternity accounted for the majority of the differences associated with gender. To further ensure this is indeed the case, the model is re-estimated in the last specification with the absence of other reasons and other labour market trajectories blocks except for the reason on parental/maternity leave. As suspected, the coefficient estimate of the female dummy remains at 0.5% and is not statistically significant.

The regression results here in equation (3.2) suggest that differences in weekly wage annual growth rate associated with gender can largely be explained by the reason for job separation on parental/maternity leave.

Next, to get a more precise measure of how much each of the variables in these labour market trajectories blocks can explain the gender gap in weekly wage annual growth rate observed in the raw data, the Oaxaca-Blinder decomposition of weekly wage annual growth rate by gender is estimated based on the last two model specifications of equation (3.2).

Table C.17 contains the Oaxaca-Blinder decomposition of gender difference in weekly wage growth estimated based on equation (3.2). The results of the Oaxaca-Blinder decomposition show that based on the full model specification in Table C.16b part II column eleven, the most important factors that can explain the gender gap in weekly wage annual growth rate observed in the raw data are the total number of permanent job separation from the type of separation block; switched employer between 2005 and 2015 from the changed employer or occupation block and both permanent & temporary separation due to parental/maternity leave from the reason for job separation block. Recall in the first specification in Table C.16 part I, the gender differences in weekly wage annual growth rate in the raw data are equal to 0.6 percentage points. This means that according to the decomposition results, both the total number of permanent separations and the total number of permanent separations due to parental/maternity leave each explain the gender differences in weekly wage annual growth rate in the raw data by -0.002/-0.006=33%, conditional on other observables. Similarly, the model is re-estimated in the last specification with the absence of all other reasons and labour market trajectories blocks except for the reason for parental/maternity leave. According to the results, the total number of permanent separations due to parental/maternity by itself explains the gender differences in weekly wage annual growth rate in the raw data by -0.003/-0.006=50%. These results are consistent with the regression results presented in Table C.16 and also in line with the results from the descriptive statistics analysis. Women in their early careers were penalized for taking time out of the labour market compared with men due to permanent job separation and permanent separation due to parental/maternity leave.

3.7 Limitations

There is a major endogeneity problem when trying to identify the impact of labour market trajectories on wage growth due to the issue of inverse causality. For example, when workers are offered a higher wage in a different employer, they are more likely to switch jobs, and high-performing workers are more likely to be offered opportunities. Since the LWF-Census linked file does not separate voluntary and involuntary job moves, and it is difficult to carry to test this hypothesis. Moreover, the analysis in this study does not distinguish between permanent separations due to job-to-job moves and those due to moves out of employment. As a result, the difference in the wage trajectories conditional on permanent separations could be due to the fact that these tend to be accompanied by job-to-job to reallocation for men, whereas these are more likely to transition out of employment for women. These two types of permanent separations have entirely different implications for earnings dynamics which is not captured in this study; as a result, the correlation of permanent and temporary separations estimated in this study on weekly wage growth could be biased. In addition, the regression model could suffer from the issue of endogeneity, which is caused by omitted variable bias. Some of the existing literature offers a number of suggestions in terms of the model structure, like the case here by Joshi, Bryson, Wilkinson, and K. Ward (2019) and Del Bono and Vuri (2011). However, these options are not feasible in this paper due to the data limitations with missing key variables in the Census and LWF, such as the duration of job separation. In addition, there are no solid theories explaining which variables would affect wage growth the most. Thus, in this study, the regression models are mostly built based on common sense and selected existing literature on the gender wage gap that is not directly linked to the topic of this study. This could lead to the absence of other important factors that should be considered if this study is to be taken further. Potential solutions do exist for some of the issues mentioned above, such as using instrumental variables and two-stage least square estimation to eliminate endogeneity bias and identify causal relationships.

3.8 Conclusion

Previous studies based on the European labour market highlights the importance of examining the differences in early career labour market trajectories, including the type of separation, changed employer or occupation and reason for separation. In addition, results on gender-based differences in early career labour market trajectories of young workers in Canada remain mostly unclear due to the lack of relevant studies. To fulfill these gaps, this study examines and decomposes the gender differences in labour market trajectories for young workers in Canada and their potential impact on wage growth over the first ten years of their career based on the LWF-Census linked data file.

In summary, according to the results from descriptive statistics analysis, among the three blocks of labour market trajectories on type of job separation, changed employer or occupation and reason for job separation, women had a greater disadvantage taking time out of the labour market compared with men. More specifically, women were less likely to experience a permanent job separations and more likely to have temporary job separations over the first ten years of their career, which is consistent with the findings from existing literature based on the

European and U.S. labour market. However, women were penalized for having permanent job separations while men were rewarded instead. For instance, for those who had no permanent job separation, women's weekly wage growth was higher than men's. On the other hand, among those who had a permanent job separation, women had lower weekly wage growth compared with men. At the same time, both women and men were slightly better off as a result of temporary job separations; however, women still had a lower increase in weekly wage growth compared with men

However, regardless of the type of job separation, women's weekly wage annual growth rate was lower compared with their male counterparts. At the same time, women were less likely to switch employers and more likely to change occupations. A more detailed analysis revealed that women in their early careers who stayed with the same employer were slightly more likely to change occupation compared with men. On the other hand, among individuals who switched employers, women were less likely than men to switch occupations. However, from the perspective of the weekly wage annual growth rate, women were worse off compared with men regardless of switching of an employer or changing in occupation.

Lastly, while the leading reasons for job separation between women and men were in large similarly, women's weekly wage growth was lower than men among the majority of the reasons investigated, except for temporary job separation due to parental or maternity leave and other reasons. These results suggest that the gender-based gap in wage growth as the result of differences in early labour market trajectories exists, and it could be a potential channel that ultimately contributed to the unexplained gender-based wage disparity observed in Canada.

The results from regression analysis suggest that labour market trajectories affect women's and men's weekly wage annual growth rates differently. For instance, women's weekly wage annual growth rate is more sensitive to temporary job separation compared with men, whereas men's weekly wage annual growth rate is more sensitive to permanent job separation. Further, an Oaxaca-Blinder decomposition shows that the most important factors that can explain the gender gap in weekly wage annual growth rate observed in the raw data are the total number of permanent job separation from the type of separation block; switched employer between 2005 and 2015 from the changed employer or occupation block and both permanent and temporary separation due to parental/maternity leave from the reason for job separation block. More specifically, the total number of permanent separations and the total number of permanent separations due to parental/maternity leave each explain the gender differences in weekly wage annual growth rate in the raw data by 33%, conditional on other observables. Together with the results from the descriptive statistics, these results suggest that women in their early career had a greater disadvantage for taking time out of the labour market due to permanent job separation and permanent separation due to parental/maternity leave compared with their male counterparts in Canada.

As a final remark, this study recognizes the endogeneity problem when trying to identify the impact of labour market trajectories on wage growth. However, the purpose of the regression analysis in this study is not to recover causal estimates of the influence of labour market trajectories; rather, the estimates offer an accounting exercise to map the correlates of unequal pay growth over the early career period and to establish empirical regularities.

As for future studies, much of the existing literature, including this study, focused on investigating and expanding the factors that contributed to the gender wage gap from the supply side; therefore, a demand-side study from the employer's perspective is desired. Perhaps using the Canadian Employer-Employee Dynamics Database (CEEDD) linked to the Census, which already contains many of the same variables from the LWF but also with the firm side information to expand research in this area.

Bibliography

[1] M. G. Abbott and C. M. Beach, "Wage changes and job changes of canadian women: Evidence from the 1986-87 labour market activity survey," *Journal of Human Resources*, vol. 29, no. 2, pp. 429–60, 1994.

[2] J. Bagger, F. Fontaine, F. Postel-Vinay, and J.-M. Robin, "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics," *American Economic Review*, vol. 104, no. 6, pp. 1551–96, 2014.

[3] G. Barrett, P. Levell, and K. Milligan, "A comparison of micro and macro expenditure measures across countries using differing survey methods," in *Improving the Measurement of Consumer Expenditures*, ser. NBER Chapters, National Bureau of Economic Research, 2013.

[4] G. S. Becker, *Human Capital and the Personal Distribution of Income: An Analytical Approach*, ser. Woytinsky lecture. Institute of Public Administration, 1967, Bibliographical footnotes.

[5] J. R. Behrman, O. S. Mitchell, C. K. Soo, and D. Bravo, "How financial literacy affects household wealth accumulation," *American Economic Review*, vol. 102, no. 3, pp. 300–304, 2012.

[6] L.-P. Béland, A. Brodeur, D. Mikola, and T. Wright, "The short-term economic consequences of COVID-19: Occupation tasks and mental health in Canada," Institute of Labor Economics, Working Paper 13254, 2020.

[7] M. Bertrand, C. Goldin, and L. F. Katz, "Dynamics of the gender gap for young professionals in the financial and corporate sectors," *American Economic Journal: Applied Economics*, vol. 2, no. 3, pp. 228–255, 2010.

[8] D. E. Bloom, D. Canning, and M. Moore, "Optimal retirement with increasing longevity," *The Scandinavian Journal of Economics*, vol. 116, no. 3, pp. 838–858, 2014.

[9] R. Branch-Allen and J. Jayachandran, "Determinants of life satisfaction in Canada: A causal modeling approach," *SHS Web of Conferences*, vol. 26, H. Çalişkan, I. Önder, E. Masal, and Ş. Beşoluk, Eds., p. 01 073, 2016.

[10] P. Brochu, "A researcher's guide to the labour force survey: Its evolution and the choice of public use versus master files," *Canadian Public Policy*, vol. 47, no. 3, pp. 335–357, 2021.

[11] P. Brochu and J. Créchet, "Survey non-response in Covid-19 times: the case of the labour force survey," *Canadian Labour Economics Forum Working Paper Series*, vol. 38, 2021.

[12] P. Brochu and D. A. Green, "The impact of minimum wages on labour market transitions," *The Economic Journal*, vol. 123, no. 573, pp. 1203–1235, 2013, ISSN: 0013-0133.

[13] G. Bronshtein, J. Scott, J. B. Shoven, and S. N. Slavov, "The power of working longer," National Bureau of Economic Research, Working Paper 24226, 2018.

[14] S. Burgess, J. Lane, and D. Stevens, "Job flows, worker flows, and churning," *Journal of Labor Economics*, vol. 18, no. 3, pp. 473–502, 2000.

[15] Z. Cai and S.-H. P. Lau, "Impact of mortality reductions on years of schooling and expected lifetime labor supply," *Journal of Mathematical Economics*, vol. 72, pp. 134–144, 2017, ISSN: 0304-4068.

[16] S. Canada, "Guide to the Labour Force Survey, 2018," *Surveys and statistical programs — Documentation:*, vol. 71-543-G, 2018, ISSN: 1704-8842.

[17] ——, "Labour force survey, july 2020," The Daily, Tech. Rep., 2020.

[18] D. Card, "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, vol. 69, no. 5, pp. 1127–1160, 2001.

[19] D. Card, A. R. Cardoso, and P. Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, vol. 131, no. 2, pp. 633–686, 2015.

[20] M. Cervellati and U. Sunde, "Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited," *Econometrica*, vol. 81, no. 5, pp. 2055–2086, 2013.

[21] P. C. W. Chan, J. Davidson, T. Qiu, and R. Morissette, "Longitudinal worker file (LWF) user guide and data dictionary.," *Statistics Canada*, 2019.

[22] P. C. W. Chang, R. Morissette, and H. Qiu, *COVID-19 and job displacement: Thinking long term*, `https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00030-eng.htm`, Statcan COVID-19: Data to Insights for a Better Canada, Statistics Canada, Cat. No. 45280001. Online; Accessed 25 August 2020, 2020.

[23] W. Cheng, P. Carlin, J. Carroll, Sumedha, G. F. Lozano Rojas, L. Montenovo, T. D. Nguyen, I. M. Schmutte, O. Scrivner, K. I. Simon, C. Wing, and B. Weinberg, "Back to business and (re)employing workers? labor market activity during state COVID-19 reopenings," National Bureau of Economic Research, Inc, NBER Working Paper 27419, 2020.

[24] O. Coibion, Y. Gorodnichenko, and M. Weber, "Labor markets during the COVID-19 crisis: A preliminary view," National Bureau of Economic Research, Inc, NBER Working Paper 27017, 2020.

[25] B. W. Cowan, "Short-run effects of COVID-19 on U.S. worker transitions," National Bureau of Economic Research, Inc, NBER Working Paper 27315, 2020.

[26] H. d'Albis, S.-H. P. Lau, and M. Sánchez-Romero, "Mortality transition and differential incentives for early retirement," *Journal of Economic Theory*, vol. 147, no. 1, pp. 261–283, 2012, ISSN: 0022-0531.

[27] S. J. Davis and J. Haltiwanger, "Gross job creation, gross job destruction, and employment reallocation," *Quarterly Journal of Economics*, vol. 107, no. 3, pp. 816–863, 1992, ISSN: 0033-5533.

[28] E. Del Bono and D. Vuri, "Job mobility and the gender wage gap in italy," *Labour Economics*, vol. 18, no. 1, pp. 130–142, 2011.

[29] D. Eryar and H. Tekgüç, "Gender Effect in Explaining Mobility Patterns in the Labor Market: A Case Study of Turkey," *The Developing Economies*, vol. 52, no. 4, pp. 322–350, 2014.

[30] E. Forsythe, "Labor market flows and the COVID-19 economy," *SSRN*, 2020.

[31] A. Frederiksen, "Gender differences in job separation rates and employment stability: New evidence from employer-employee data," *Labour Economics*, vol. 15, no. 5, pp. 915–937, 2008, ISSN: 0927-5371.

[32] S. Fuller, "Job mobility and wage trajectories for men and women in the united states," *American Sociological Review*, vol. 73, no. 1, pp. 158–183, 2008.

[33] G. Gallacher and I. Hossain, "Remote work and employment dynamics under COVID-19: Evidence from Canada," *Canadian Public Policy*, vol. 46, no. S1, S44–S54, 2020, ISSN: 0317-0861.

[34] J. H. Goldthorpe, "Causation, statistics, and sociology," *European Sociological Review*, vol. 17, no. 1, pp. 1–20, 2001.

[35] D. A. Green and K. Milligan, "The importance of the long form census in Canada," *Canadian Public Policy*, vol. 36, no. 3, pp. 383–388, 2010, ISSN: 0317-0861.

[36] J. K. Hahn, H. R. Hyatt, H. P. Janicki, and S. R. Tibbets, "Job-to-job flows and earnings growth," *American Economic Review*, vol. 107, no. 5, pp. 358–363, 2017.

[37] J. Haltiwanger, H. Hyatt, and E. McEntarfer, "Who moves up the job ladder?" *Journal of Labor Economics*, vol. 36, no. S1, S301–S336, 2018.

[38] M. Hazan, "Longevity and lifetime labor supply: Evidence and implications," *Econometrica*, vol. 77, no. 6, pp. 1829–1863, 2009.

[39] B. J. Heijdra and W. E. Romp, "Human capital formation and macroeconomic performance in an ageing small open economy," *Journal of Economic Dynamics and Control*, vol. 33, no. 3, pp. 725–744, 2009, ISSN: 0165-1889.

[40] B. Hirsch and C. Schnabel, "Women Move Differently: Job Separations and Gender," *Journal of Labor Research*, vol. 33, no. 4, pp. 417–442, 2012.

[41] F. Hou, G. Picot, and J. Zhang, "Transitions into and out of employment by immigrants during the COVID-19 lockdown and recovery," *StatCan COVID-19: Data to Insights for a Better Canada*, vol. 45280001, 2020.

[42] S. R. Jones, F. Lange, W. C. Riddell, and C. Warman, "Waiting for recovery: The canadian labour market in june 2020," *Canadian Public Policy*, vol. 46, no. S2, S102–S118, 2020.

[43] H. Joshi, A. Bryson, D. Wilkinson, and K. Ward, "The gender gap in wages over the life course: Evidence from a british cohort born in 1958," Institute of Labor Economics (IZA), Bonn, IZA Discussion Papers 12725, 2019.

[44] B. Jovanovic, "Job matching and the theory of turnover," *Journal of Political Economy*, vol. 87, no. 5, pp. 972–990, 1979.

[45] P. Jung and M. Kuhn, "Correction to: Earnings losses and labor mobility over the life cycle," *Journal of the European Economic Association*, vol. 17, no. 3, pp. 990–990, 2019.

[46] S. Kalemli-Ozcan and D. N. Weil, "Mortality change, the uncertainty effect, and retirement," *Journal of Economic Growth*, vol. 15, no. 1, pp. 65–91, 2010.

[47] H. Kleven, C. Landais, J. Posch, A. Steinhauer, and J. Zweimülle, "Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation," Working Paper w28082, 2020.

[48] H. Kleven, C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller, "Child penalties across countries: Evidence and explanations," *In AEA Papers and Proceedings*, vol. 109, pp. 122–26, 2019.

[49] H. Kleven, C. Landais, and J. E. Søgaard, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, vol. 11, no. 4, pp. 181–209, 2019.

[50] K. Koebel and D. Pohler, "Labor markets in crisis: The causal impact of Canada's COVID19 economic shutdown on hours worked for workers across the earnings distribution," University of Waterloo, Canadian Labour Economics Forum (CLEF), Waterloo, CLEF Working Paper 25, 2020.

[51] T. Lemieux, K. Milligan, T. Schirle, and M. Skuterud, "Initial impacts of the COVID-19 pandemic on the canadian labour market," *Canadian Public Policy*, vol. 46, no. S1, S55–S65, 2020.

[52] F. Lima, "How much mobility? careers, promotions, and wages," *CEG-IST*, vol. 3, 2004.

[53] B. C. Madrian and L. J. Lefgren, "An approach to longitudinally matching current population survey (CPS) respondents," *Journal of Economic and Social Measurement*, vol. 26, pp. 31–62, 2000.

[54] A. Manning and J. Swaffield, "The gender gap in early-career wage growth," *The Economic Journal*, vol. 118, no. 530, pp. 983–1024, 2008, ISSN: 0013-0133.

[55] K. S. Markides and H. W. Martin, "A causal model of life satisfaction among the elderly," *Journal of Gerontology*, vol. 34, no. 1, pp. 86–93, 1979, ISSN: 0022-1422.

[56] S. Meeks and S. A. Murrell, "Contribution of education to health and life satisfaction in older adults mediated by negative affect," *Journal of Aging and Health*, vol. 13, no. 1, pp. 92–119, 2001.

[57] D. Messacar, "The effects of education on Canadians' retirement savings behaviour," *Analysis*, Research paper series, vol. 11F0019M2017391, 2017, ISSN: 1205-9153.

[58] V. Møller, "Quality of life in retirement: A case study of zulu return migrants," *Social Indicators Research*, vol. 20, no. 6, pp. 621–658, 1988.

[59] M. Moyser, *Measuring and analyzing the gender pay gap : a conceptual and methodological overview.* Ottawa: Studies on Gender and Intersecting Identities, Statistics Canada, 2019, Catalogue no. 45-20-0002, ISBN: 978-0-660-31741-0.

[60] P. Oreopoulos, "Estimating average and local average treatment effects of education when compulsory schooling laws really matter," *American Economic Review*, vol. 96, no. 1, pp. 152–175, 2006.

[61] P. Oreopoulos and K. G. Salvanes, "Priceless: The nonpecuniary benefits of schooling," *Journal of Economic Perspectives*, vol. 25, no. 1, pp. 159–184, 2011.

[62] R. Pelletier, M. Patterson, and M. Moyser, *The gender wage gap in Canada : 1998 to 2018.* Ottawa: Government of Canada Publications, 2019, Catalogue no. 75-004-M2019004, ISBN: 978-0-660-32214-8.

[63] N. Powdthavee, W. N. Lekfuangfu, and M. Wooden, "What's the good of education on our overall quality of life? a simultaneous equation model of education and life satisfaction for Australia," *Journal of Behavioral and Experimental Economics*, vol. 54, no. 1, pp. 10–21, 2015, ISSN: 2214-8043.

[64] M. Pries and R. Rogerson, "Hiring policies, labor market institutions, and labor market flows," *Journal of Political Economy*, vol. 113, no. 4, pp. 811–839, 2005.

[65] Y. Qian and S. Fuller, "COVID-19 and the gender employment gap among parents of young children," *Canadian Public Policy*, vol. 46, no. S2, S89–S101, 2020.

[66] D. Restuccia and G. Vandenbroucke, "A century of human capital and hours," *Economic Inquiry*, vol. 51, no. 3, pp. 1849–1866, 2012.

[67] J. A. Rivera Drew, S. Flood, and J. R. Warren, "Making full use of the longitudinal design of the current population survey: Methods for linking records across 16 months," *Journal of Economic and Social Measurement*, vol. 39, no. 3, pp. 121–144, 2014, ISSN: 18758932, 07479662.

[68] J. Robson, "Parental Benefits in Canada: Which Way Forward?" *Montreal: Institute for Research on Public Policy*, 2017.

[69] M. C. J. van Rooij, A. Lusardi, and R. J. M. Alessie, "Financial literacy, retirement planning and household wealth," *The Economic Journal*, vol. 122, no. 560, pp. 449–478, 2012, ISSN: 0013-0133.

[70] A. Şahin, M. Tasci, and J. Yan, "The unemployment cost of COVID-19: How high and how long?" *Economic Commentary (Federal Reserve Bank of Cleveland)*, vol. 202009, pp. 1–7, 2020, ISSN: 2163-3738.

[71] Statistics Canada, *Methodology of the Canadian Labour Force Survey*, ser. Surveys and statistical programs — Documentation 71-526-X. Ottawa: Statistics Canada, 2017, ISBN: 978-0-660-24068-8.

[72] R. H. Topel and M. P. Ward, "Job mobility and the careers of young men," *The Quarterly Journal of Economics*, vol. 107, no. 2, pp. 439–79, 1992.

[73] U.S. Bureau of Labor Statistics, *Frequently asked questions: The impact of the coronavirus (COVID-19) pandemic on the employment situation for may 2020*, `https://www.bls.gov/cps/employment-situation-covid19-faq-may-2020.pdf`, Online; Accessed 16 April 2021, 2020.

[74] ——, *The impact of the coronavirus (COVID-19) pandemic on the employment situation for June 2020*, `https://www.bls.gov/cps/employment-situation-covid19-faq-june-2020.pdf`, Online; Accessed 21 August 2021, 2020.

[75] X. Zhang, "Can motherhood earnings losses be ever regained? Evidence from Canada," *Journal of Family Issues*, vol. 31, no. 12, pp. 1671–1688, 2010.

Appendix A

Table A.1: Descriptive Statistics for Native Born Respondents aged 14 between 1931 and 1975 in the 2013 to 2016 General Social Survey

	Survey Weights		Equal Province / Cohort Weights	
	Mean	Standard Deviation	Mean	Standard Deviation
Years of schooling	12.923	(3.262)	12.893	(2.976)
Retirement status (if retired)	0.678	(0.467)	0.677	(0.258)
Retirement age	56.728	(9.837)	55.558	(3.360)
Exposed to school leaving age ≤ 13	0.003	(0.053)	0.005	(0.074)
Exposed to school leaving age = 14	0.181	(0.385)	0.158	(0.365)
Exposed to school leaving age = 15	0.361	(0.480)	0.378	(0.485)
Exposed to school leaving age = 16	0.455	(0.500)	0.459	(0.499)
Exposed to school entry age = 8	0.062	(0.241)	0.041	(0.199)
Exposed to school entry age = 7	0.333	(0.471)	0.450	(0.498)
Exposed to school entry age = 6	0.605	(0.489)	0.509	(0.501)
Have a high school certificate	0.757	(0.429)	0.755	(0.126)
Highest educ. attainment = College	0.176	(0.381)	0.182	(0.065)
Have some post secondary schooling	0.493	(0.500)	0.492	(0.107)
Have a university degree	0.181	(0.385)	0.177	(0.059)
Log income	9.540	(0.873)	9.528	(0.164)
Log income males	9.781	(0.818)	9.482	(0.473)
Log household income	10.184	(0.755)	10.053	(0.240)
N	28,884	28,884	2,703	2,703
Change in enjoyment of life before and after retirement (if improved)	0.470	(0.499)	N/A	N/A
Change in financial standard of living after retirement (if improved)	0.192	(0.394)	N/A	N/A
N	3,887	3.887	N/A	N/A

Income is adjusted using the 2002 Canadian Consumer Price index (CPI).

Table A.2: Descriptive Statistics for Native Born Respondents and Non-Native Born Respondents aged 14 between 1931 and 1975 in the 2013 to 2016 General Social Survey

	Survey Weights		Equal Province / Cohort Weights	
	Mean	Standard Deviation	Mean	Standard Deviation
Years of schooling	12.985	(3.267)	12.942	(2.979)
Retirement status (if retired)	0.675	(0.469)	0.674	(0.257)
Retirement age	56.738	(9.789)	55.609	(3.309)
Exposed to school leaving age ≤ 13	0.003	(0.057)	0.005	(0.072)
Exposed to school leaving age $= 14$	0.173	(0.378)	0.153	(0.361)
Exposed to school leaving age $= 15$	0.360	(0.480)	0.376	(0.485)
Exposed to school leaving age $= 16$	0.464	(0.499)	0.466	(0.499)
Exposed to school entry age $= 8$	0.062	(0.241)	0.042	(0.199)
Exposed to school entry age $= 7$	0.332	(0.471)	0.447	(0.498)
Exposed to school entry age $= 6$	0.606	(0.489)	0.512	(0.501)
Have a high school certificate	0.763	(0.425)	0.759	(0.126)
Highest educ. attainment = College	0.176	(0.381)	0.183	(0.064)
Have some post secondary schooling	0.498	(0.500)	0.496	(0.107)
Have a university degree	0.187	(0.390)	0.182	(0.060)
Log income	9.550	(0.881)	9.537	(0.166)
Log income males	9.790	(0.821)	9.505	(0.486)
Log household income	10.195	(0.767)	10.060	(0.242)
N	30,465	30,465	2,709	2,709
Change in enjoyment of life before and after retirement (if improved)	0.464	(0.499)	N/A	N/A
Change in financial standard of living after retirement (if improved)	0.187	(0.389)	N/A	N/A
N	4,070	4,070	N/A	N/A

Income is adjusted using the 2002 Canadian Consumer Price index (CPI).

Table A.3: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling (Log Total Annual Income)

Dependent Variable: Years of Schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.688*** (0.059)	0.472*** (0.040)	0.411*** (0.051)
Age		-9.173** (4.121)	Dropped
Age^2		0.212** (0.085)	-0.266** (0.145)
Age^3		-0.002*** (0.001)	0.002** (0.001)
Age^4		0.000*** (0.000)	-0.000* (0.000)
Female		-0.096** (0.042)	-0.095** (0.042)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.156	0.404	0.458
F-stat of excluded instruments	137.71	110.04	66.70
N	2,703	2,703	2,703

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Equal Province/Cohort weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.4: Second-Stage Regression Results, Years of Schooling and Log Total Annual Income

Dependent Variable: Log Total Annual Income			
	(1)	(2)	(3)
Years of Schooling	0.092*** (0.014)	0.082** (0.020)	0.045* (0.026)
Age		-3.269** (0.884)	Dropped
Age^2		0.061** (0.018)	0.043 (0.052)
Age^3		-0.001* (0.000)	-0.000 (0.000)
Age^4		0.000* (0.000)	0.000 (0.000)
Female		-0.364*** (0.014)	-0.370*** (0.013)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.127	0.432	0.442
N	2,703	2,703	2,703

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Equal Province/Cohort weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.5: OLS Regression Results, Years of Schooling and Log Total Annual Income

Dependent variable: Log total annual income			
	(1)	(2)	(3)
Years of schooling	0.090*** (0.005)	0.088*** (0.005)	0.074*** (0.005)
Age		-3.679*** (0.856)	-3.591 (2.207)
Age^2		0.071*** (0.018)	0.067 (0.046)
Age^3		-0.001*** (0.000)	-0.001 (0.000)
Age^4		0.000*** (0.000)	0.000 (0.000)
Female		-0.368*** (0.014)	-0.368*** (0.014)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.127	0.447	0.510
N	2,703	2,703	2,703

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Equal Province/Cohort weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.6: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling (Retirement Status)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.719***	0.499***	0.354***
	(0.075)	(0.053)	(0.073)
Age^2		0.025***	-0.064
		(0.004)	(0.179)
Age^3		-0.000***	0.001
		(0.000)	(0.002)
Age^4		0.000***	-0.000
		(0.000)	(0.000)
Female		-0.247***	-0.244***
		(0.052)	(0.052)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.028	0.076	0.085
F-stat of excluded instruments	88.11	81.54	30.83
N	28,844	28,844	28,844

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.7: Second-Stage Regression Results, Years of Schooling and Retirement Status

Dependent variable: Retirement status	(1)	(2)	(3)
Years of schooling	-0.147***	-0.028**	0.008
	(0.027)	(0.011)	(0.018)
Age^2		0.008***	0.093***
		(0.001)	(0.021)
Age^3		-0.000***	-0.001***
		(0.000)	(0.000)
Age^4		0.000***	0.000***
		(0.000)	(0.000)
Female		-0.012	-0.005
		(0.008)	(0.009)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.007	0.223	0.229
N	28,844	28,844	28,844

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.8: OLS Regression Results, Years of Schooling and Retirement Status

Dependent variable: Retirement status	(1)	(2)	(3)
Years of schooling	-0.013***	0.001	0.001
	(0.002)	(0.001)	(0.001)
Age^2		0.007***	0.007***
		(0.001)	(0.002)
Age^3		-0.000***	-0.001***
		(0.000)	(0.000)
Age^4		0.000***	0.000***
		(0.000)	(0.000)
Female		-0.005	-0.004
		(0.007)	(0.007)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.008	0.314	0.385
N	28,844	28,844	28,844

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.9: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling (Retirement Age)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.607*** (0.077)	0.422*** (0.057)	0.247*** (0.080)
Age^2		0.035*** (0.005)	-0.187 (0.254)
Age^3		-0.001*** (0.001)	0.002 (0.002)
Age^4		0.000*** (0.000)	-0.000 (0.000)
Female		-0.198*** (0.064)	-0.198*** (0.064)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.023	0.061	0.072
F-stat of excluded instruments	61.54	55.44	20.52
N	18,879	18,879	18,879

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.10: Second-Stage Regression Results, Years of Schooling and Retirement Age

Dependent variable: Retirement age			
	(1)	(2)	(3)
Years of schooling	0.143 (0.320)	1.154 (0.471)	1.271 (0.952)
Age^2		0.084*** (0.023)	-0.108 (1.052)
Age^3		-0.001*** (0.004)	0.000 (0.010)
Age^4		0.000*** (0.000)	-0.000 (0.000)
Female		-4.438*** (0.295)	-4.647*** (0.331)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.006	0.007	0.009
N	18,879	18,879	18,879

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.11: OLS Regression Results, Years of Schooling and Retirement Age

Dependent variable: Retirement age			
	(1)	(2)	(3)
Years of schooling	0.302*** (0.031)	0.345*** (0.028)	0.320*** (0.029)
Age^2		0.149*** (0.016)	-0.301 (1.034)
Age^3		-0.002*** (0.000)	0.002 (0.009)
Age^4		0.000*** (0.000)	-0.000 (0.000)
Female		-4.792*** (0.252)	-4.836*** (0.257)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.009	0.106	0.118
N	18,879	18,879	18,879

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.12: Second-Stage Regression Results, Years of Schooling and Retirement Status by Age Group

Dependent variable: Retirement status			
	Age group 50 to 70	Age group 55 to 65	Age group 60 to 65
Years of schooling	0.116 (0.216)	0.013 (0.301)	2.476 (20.013)
Age^2	-0.007 (0.350)	0.081 (0.144)	-1.019 (9.180)
Age^3	0.000 (0.004)	-0.000 (0.001)	0.006 (0.051)
Age^4	-0.000 (0.000)	Dropped	Dropped
Female	0.040 (0.035)	0.049* (0.028)	0.334 (2.462)
Cohort-varying controls	✓	✓	✓
Provincial fixed effect	✓	✓	✓
Cohort fixed effect	✓	✓	✓
Survey year fixed effect	✓	✓	✓
F-stat of excluded instruments from first stage	18.58	12.43	6.02
adj. R^2	0.231	0.115	0.056
N	18,027	12,063	8,279

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.13: Second-Stage Regression Results, Years of Schooling and Retirement Age by Age Group

Dependent variable: Retirement age	Age group 50-70	Age group 55-65	Age group 60-65
Years of schooling	2.466 (2.077)	1.818 (1.691)	4.734 (6.703)
Age^2	0.771 (1.525)	-0.164 (2.400)	-7.074 (25.255)
Age^3	-0.560 (0.025)	0.001 (0.013)	0.038 (0.135)
Age^4	0.002 (0.084)	Dropped	Dropped
Female	-2.918** (0.395)	-2.270*** (0.332)	-2.047** (0.891)
Cohort-varying controls	✓	✓	✓
Provincial fixed effect	✓	✓	✓
Cohort fixed effect	✓	✓	✓
Survey year fixed effect	✓	✓	✓
F-stat of excluded instruments from first stage	7.01	5.07	3.37
adj. R^2	0.127	0.115	0.077
N	9,675	5,205	4,146

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.14: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling Include Immigrants (Retirement Status)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.733***	0.512***	0.364***
	(0.075)	(0.052)	(0.070)
Age^2		0.025***	-0.065
		(0.004)	(0.177)
Age^3		-0.000***	0.001
		(0.000)	(0.002)
Age^4		0.000***	-0.000
		(0.000)	(0.000)
Female		-0.271***	-0.269***
		(0.051)	(0.051)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.030	0.078	0.086
F-stat of excluded instruments	95.59	88.14	29.75
N	30,465	30,465	30,465

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.15: Second-Stage Regression Results, Years of Schooling and Retirement Status Include Immigrants

Dependent variable: Retirement status			
	(1)	(2)	(3)
Years of schooling	-0.143**	-0.028**	0.007
	(0.028)	(0.010)	(0.018)
Age^2		0.008***	0.086***
		(0.001)	(0.022)
Age^3		-0.001***	-0.001***
		(0.000)	(0.000)
Age^4		0.000***	0.000***
		(0.000)	(0.000)
Female		-0.013*	-0.004
		(0.008)	(0.009)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.006	0.242	0.246
N	30,465	30,465	30,465

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.16: OLS Regression Results, Years of Schooling and Retirement Status Include Immigrants

Dependent variable: Retirement status			
	(1)	(2)	(3)
Years of schooling	-0.013***	0.001	0.001
	(0.002)	(0.001)	(0.006)
Age^2		0.007***	0.007***
		(0.001)	(0.001)
Age^3		-0.000***	-0.001***
		(0.000)	(0.000)
Age^4		0.000***	0.000***
		(0.000)	(0.000)
Female		-0.005	-0.005
		(0.007)	(0.007)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey Year fixed effect			✓
adj. R^2	0.008	0.310	0.317
N	30,465	30,465	30,465

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.17: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling Include Immigrants (Retirement Age)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.629*** (0.076)	0.446*** (0.056)	0.287*** (0.079)
Age^2		0.035*** (0.005)	-0.172 (0.242)
Age^3		-0.001*** (0.001)	0.002 (0.002)
Age^4		0.000*** (0.000)	-0.000 (0.000)
Female		-0.239*** (0.063)	-0.238*** (0.062)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.024	0.063	0.074
F-stat of excluded instruments	68.01	63.52	20.38
N	19,881	19,881	19,881

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.18: Second-Stage Regression Results, Years of Schooling and Retirement Age Include Immigrants

Dependent variable: Retirement age			
	(1)	(2)	(3)
Years of schooling	0.075 (0.308)	1.012 (0.445)	1.029 (0.789)
Age^2		0.094*** (0.022)	-0.097 (1.017)
Age^3		-0.001*** (0.000)	0.009 (0.009)
Age^4		0.000*** (0.000)	0.000 (0.000)
Female		-4.320*** (0.280)	-4.568*** (0.321)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.004	0.006	0.007
N	19,881	19,881	19,881

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.19: OLS Regression Results, Years of Schooling and Retirement Age Include Immigrants

Dependent variable: Retirement age			
	(1)	(2)	(3)
Years of schooling	0.308*** (0.031)	0.348*** (0.027)	0.325*** (0.028)
Age^2		0.150*** (0.016)	-0.234 (1.006)
Age^3		-0.002*** (0.000)	0.001 (0.009)
Age^4		0.000*** (0.000)	-0.000 (0.000)
Female		-4.689*** (0.244)	-4.736*** (0.248)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey Year fixed effect			✓
adj. R^2	0.009	0.105	0.116
N	19,881	19,881	19,881

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.20: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling (Change in Quality of Life After Retirement, 2016 GSS)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.541*** (0.105)	0.226** (0.115)	0.206* (0.118)
Age^2		0.019* (0.011)	-0.074 (0.179)
Age^3		-0.000* (0.000)	0.001 (0.002)
Age^4		0.000* (0.000)	-0.000 (0.000)
Female		-0.283** (0.123)	-0.244 (0.145)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.024	0.061	0.076
F-stat of excluded instruments	36.72	23.88	16.21
N	3,887	3,887	3,887

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.21: Second-Stage Regression Results, Years of Schooling and Change in Enjoyment of Life Before and After Retirement

Dependent variable: Change in enjoyment of life			
	(1)	(2)	(3)
Years of schooling	-0.086**	-0.044	-0.021
	(0.027)	(0.077)	(0.098)
Age^2		0.004	0.001
		(0.002)	(0.001)
Age^3		-0.000	-0.000
		(0.000)	(0.000)
Age^4		0.000	0.000
		(0.000)	(0.000)
Female		-0.040	-0.038
		(0.030)	(0.055)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.021	0.039	0.040
N	3,887	3,887	3,887

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.22: OLS Regression Results, Years of Schooling and Change in Enjoyment of Life Before and After Retirement

Dependent variable: Change in enjoyment of life			
	(1)	(2)	(3)
Years of schooling	-0.026***	-0.020***	-0.019***
	(0.002)	(0.003)	(0.006)
Age^2		0.003*	0.003*
		(0.002)	(0.002)
Age^3		-0.000*	-0.000
		(0.000)	(0.000)
Age^4		0.000	0.000
		(0.000)	(0.000)
Female		-0.033*	-0.033*
		(0.019)	(0.019)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.028	0.062	0.078
N	3,887	3,887	3,887

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.23: Second-Stage Regression Results, Years of Schooling and Change in Financial Standard of Living After Retirement

Dependent variable: Change in financial standard of Lliving			
	(1)	(2)	(3)
Years of schooling	-0.003 (0.020)	-0.002 (0.068)	-0.002 (0.079)
Age^2		-0.001 (0.002)	-0.001 (0.001)
Age^3		0.000 (0.000)	0.000 (0.000)
Age^4		-0.000 (0.000)	-0.000 (0.000)
Female		-0.030 (0.024)	-0.030 (0.035)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.001	0.003	0.004
N	3,887	3,887	3,887

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.24: OLS Regression Results, Years of Schooling and Change in Financial Standard of Living After Retirement

Dependent variable: Change in financial standard of living			
	(1)	(2)	(3)
Years of schooling	-0.002	-0.004	-0.004
	(0.002)	(0.003)	(0.003)
Age^2		-0.001	-0.003
		(0.001)	(0.002)
Age^3		0.000	0.000
		(0.000)	(0.000)
Age^4		-0.000	-0.000
		(0.000)	(0.000)
Female		-0.016	-0.018
		(0.017)	(0.019)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.001	0.008	0.010
N	3,887	3,887	3,887

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.25: First-Stage Regression Results, Minimum School Leaving Age and Years of Schooling Include Immigrants (Change in Quality of Life After Retirement, 2016 GSS)

Dependent variable: Years of schooling			
	(1)	(2)	(3)
Minimum school leaving age	0.553*** (0.017)	0.250** (0.115)	0.210* (0.116)
Age^2		0.025** (0.018)	-0.084 (0.179)
Age^3		-0.000** (0.000)	0.001 (0.002)
Age^4		0.000** (0.000)	-0.000 (0.000)
Female		-0.313*** (0.112)	-0.321*** (0.120)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.024	0.050	0.052
F-stat of excluded instruments	35.72	26.96	19.20
N	4,070	4,070	4,070

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.26: Second-Stage Regression Results, Years of Schooling and Change in Enjoyment of Life Before and After Retirement Include Immigrants

Dependent variable: Change in enjoyment of life			
	(1)	(2)	(3)
Years of schooling	-0.068** (0.026)	-0.048 (0.070)	-0.031 (0.098)
Age^2		0.002** (0.002)	0.002 (0.001)
Age^3		-0.000* (0.000)	-0.000 (0.000)
Age^4		0.000* (0.000)	0.000 (0.000)
Female		-0.027 (0.030)	-0.039 (0.055)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.022	0.057	0.059
N	4,070	4,070	4,070

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.27: OLS Regression Results, Years of Schooling and Change in Enjoyment of Life Before and After Retirement Include Immigrants

Dependent variable: change in enjoyment of life			
	(1)	(2)	(3)
Years of schooling	-0.023***	-0.018***	-0.013***
	(0.003)	(0.003)	(0.004)
Age^2		0.002	0.001
		(0.002)	(0.003)
Age^3		-0.000*	-0.000
		(0.000)	(0.000)
Age^4		0.000*	0.000
		(0.000)	(0.000)
Female		-0.028	-0.025
		(0.020)	(0.022)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.023	0.057	0.062
N	4,070	4,070	4,070

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.28: Second-Stage Regression Results, Years of Schooling and Change in Financial Standard of Living After Retirement Include Immigrants

Dependent variable: Change in financial standard of living			
	(1)	(2)	(3)
Years of schooling	-0.004 (0.022)	-0.003 (0.065)	-0.003 (0.079)
Age^2		-0.001 (0.002)	-0.000 (0.004)
Age^3		0.000 (0.000)	0.000 (0.000)
Age^4		-0.000 (0.000)	-0.000 (0.000)
Female		-0.021 (0.026)	-0.027 (0.035)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.002	0.004	0.006
N	4,070	4,070	4,070

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.29: OLS Regression Results, Years of Schooling and Change in Financial Standard of Living After Retirement Include Immigrants

Dependent variable: Change in financial standard of living			
	(1)	(2)	(3)
Years of schooling	-0.003 (0.002)	-0.004 (0.002)	-0.005 (0.003)
Age^2		-0.001 (0.001)	-0.000 (0.002)
Age^3		0.000 (0.000)	0.000 (0.000)
Age^4		-0.000 (0.000)	-0.000 (0.000)
Female		-0.012 (0.017)	-0.013 (0.019)
Cohort-varying controls		✓	✓
Provincial fixed effect			✓
Cohort fixed effect			✓
Survey year fixed effect			✓
adj. R^2	0.001	0.008	0.009
N	4,070	4,070	4,070

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.30: Decomposition: Second-Stage Regression Results, Years of Schooling and Occupation Part I

Dependent variable: Occupation					
	Management	Adminstration	Science	Health	Education
Years of schooling	0.012 (0.060)	0.055 (0.073)	0.034 (0.053)	-0.050 (0.065)	-0.075 (0.099)
Age^2	0.093 (0.087)	0.127 (0.090)	0.004 (0.064)	-0.059 (0.073)	-0.068 (0.131)
Age^3	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.000)	-0.000 (0.000)	0.001 (0.001)
Age^4	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Female	-0.027*** (0.008)	0.158*** (0.011)	-0.049*** (0.008)	0.077*** (0.009)	0.054*** (0.016)
Cohort-varying controls	✓	✓	✓	✓	✓
Provincial fixed effect	✓	✓	✓	✓	✓
Cohort fixed effect	✓	✓	✓	✓	✓
Survey year fixed effect	✓	✓	✓	✓	✓
F-stat of excluded instruments from first stage	5.91	5.91	5.90	5.90	5.90
adj. R^2	0.016	0.016	0.015	0.015	0.015
N	7,911	7,911	7,911	7,911	7,911

Notes: The occupation questions only applies to individuals who identified themselves as currently working or worked anytime during the past 12 month at the time of the survey. Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.31: Decomposition: Second-Stage Regression Results, Years of Schooling and Occupation Part II

Dependent variable: Occupation					
	Art	Sales	Trades	Agriculture	Manufacturing
Years of schooling	-0.017 (0.033)	0.085 (0.089)	0.082 (0.106)	-0.051 (0.047)	-0.076 (0.064)
Age^2	0.157* (0.086)	-0.047 (0.131)	-0.165 (0.132)	0.057 (0.080)	-0.098 (0.072)
Age^3	-0.002* (0.001)	0.001 (0.001)	0.002 (0.001)	-0.001 (0.001)	0.001 (0.001)
Age^4	0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Female	0.002 (0.005)	0.033** (0.015)	-0.202*** (0.021)	-0.019*** (0.006)	-0.026** (0.009)
Cohort-varying controls	✓	✓	✓	✓	✓
Provincial fixed effect	✓	✓	✓	✓	✓
Cohort fixed effect	✓	✓	✓	✓	✓
Survey year fixed effect	✓	✓	✓	✓	✓
F-stat of excluded instruments from first stage	5.91	5.90	5.90	5.91	5.91
adj. R^2	0.016	0.016	0.015	0.016	0.015
N	7,911	7,911	7,911	7,911	7,911

Notes: The occupation questions only applies to individuals who identified themselves as currently working or worked anytime during the past 12 month at the time of the survey. Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.32: Decomposition: Second-Stage Regression Results, Years of Schooling and Other Socialecomic Factors

Dependent variable:	Married/ Common-law	General health	Mental health	Life satisfaction	Log total Annual income
Years of schooling	-0.007	0.025	0.024	0.175	0.031*
	(0.025)	(0.026)	(0.015)	(0.114)	(0.019)
Age^2	-0.020	0.046	0.015	0.338	0.075
	(0.033)	(0.031)	(0.016)	(0.143)	(0.054)
Age^3	0.000	-0.000	-0.000	-0.003	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Age^4	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.180***	0.017*	0.010*	0.123***	-0.442***
	(0.013)	(0.009)	(0.005)	(0.038)	(0.022)
Cohort-varying controls	✓	✓	✓	✓	✓
Provincial fixed effect	✓	✓	✓	✓	✓
Cohort fixed effect	✓	✓	✓	✓	✓
Survey year fixed effect	✓	✓	✓	✓	✓
F-stat of excluded instruments from first stage	24.50	25.11	24.36	23.00	30.91
adj. R^2	0.020	0.014	0.013	0.016	0.077
N	25,288	25,288	25,288	25,288	25,288

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table A.33: Decomposition: OLS Regression Results, Log Total Annual Income and Retirement Decisions

Dependent variable:		
	Retirement status	Retirement age
Log total annual income	-0.062*** (0.005)	1.581*** (0.145)
Age^2	0.113*** (0.025)	-0.924 (1.151)
Age^3	-0.001*** (0.000)	0.008 (0.011)
Age^4	0.000*** (0.000)	-0.000 (0.000)
Female	-0.036*** (0.007)	-4.314*** (0.294)
Cohort-varying controls	✓	✓
Provincial fixed effect	✓	✓
Cohort fixed effect	✓	✓
Survey year fixed effect	✓	✓
adj. R^2	0.339	0.115
N	25,288	16,587

Notes: Standard errors clustered at provincial and birth cohort level which are shown in parentheses. Survey weights are used. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Appendix B

B.1 Creation of mini panels

In this appendix we provide information on how we construct our LFS mini panels.

The LFS follows a rotating panel design where households stay in the sample for six consecutive months. Every month one-sixth of the sample (i.e. households) is replaced by households in a similar area. The LFS is officially designed to generate cross-sectional samples; it follows dwellings and not individuals. If, for example, an individual leaves the dwelling part way through the six-month window, he/she is beyond the scope of the survey. Similarly, an individual that joins a household late will only be asked labour market information as of the time he/she started living in the targeted dwelling/household.

The LFS does not have a single person identifier variable. For our period of interest, one can nonetheless uniquely identify individuals across monthly files using the following two variables: the LFS household identifier (HHLDID), and the LINE variable that (uniquely) identifies a person within the household. The HHLDID variable (also called DWELID) is in fact made up of 10 variables (PROV, PROV1, PSEUDOUI, FRAME, STRAFRAM, TYPE, CLUST, ROTATION, LISTLINE and MULT). Once combined, they generate the unique 18-digit household/dwelling identifier.

It should be noted that for earlier periods (i.e. prior to 1996), the creation of mini panels requires the use of different sets of variables. See Brochu and Green (2013) and Brochu (2021) for details.

B.1.1 Worker-flow computation

Labour-market flows are computed using two-month panels. We compute the separation rate as:

$$\text{outflows}_t = \frac{\sum_i \omega_{i,t-1}\mathcal{I}(E_{i,t-1}=1, E_{i,t}=0)}{\sum_i \omega_{i,t-1}\mathcal{I}(E_{i,t-1}=1)},$$

for all period t in the sample of interest, and where $E_{i,t}$ is a dummy taking value of one when individual i is employed in t, $\mathcal{I}$ is the indicator function, which takes the value of one when the expression in parenthesis is true and zero otherwise, and where $\omega_{i,t} \geq 0$ represents the LFS weights of individual i in period t. Similarly, we compute hiring flows as:

$$\text{inflows}_t = \frac{\sum_i \omega_{i,t-1}\mathcal{I}(E_{i,t-1}=0, E_{i,t}=1)}{\sum_i \omega_{i,t-1}\mathcal{I}(E_{i,t-1}=1)},$$

for all t in the sample. Using these outflow and inflow measures, excess reallocation flows are computed following:

$$\text{excessflows}_t = \text{inflows}_t + \text{outflows}_t - |\text{inflows}_t - \text{outflows}_t|,$$

for all t in th sample. Resulting measures of flows are reported in Tables B.2 and B.13. We compute flows in and out of the LFS category "at work" (excluding workers on vacation, parental leave, and absent due to labour conflicts), as in Table B.2, and by occupations, as in Tables B.16 and B.17, following the same approach (considering flows in and out of employment, and in and out of given occupation groups).

B.2 Figures and tables

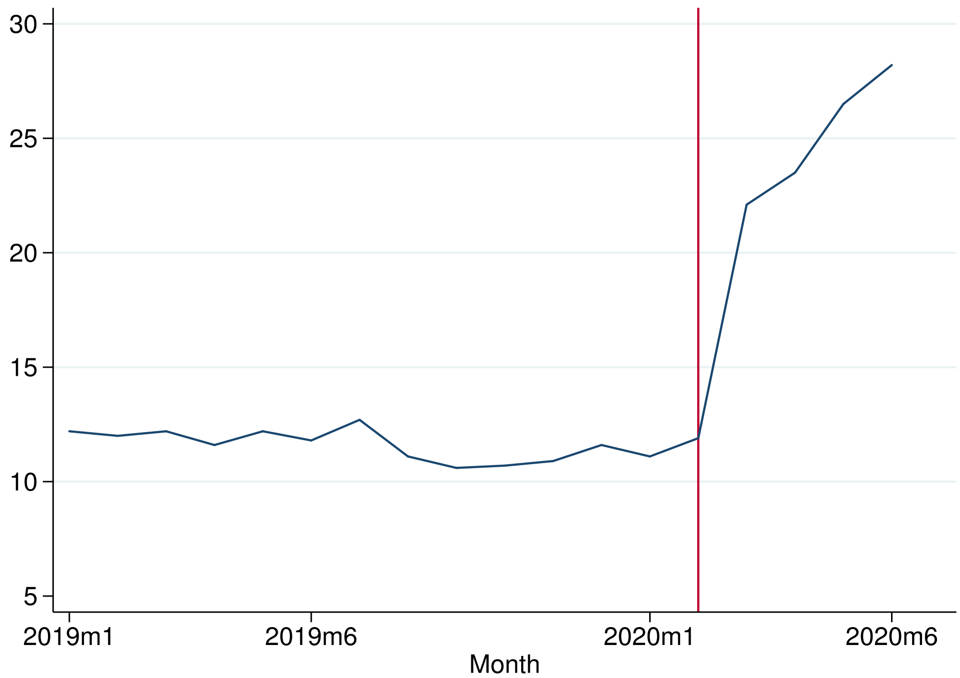

Figure B.1: Monthly household non-response rates (%). Data provided to the authors by Statistics Canada. Share of households in the monthly LFS with no available information. The red vertical line indicates February 2020. The data includes all LFS households in the ten Canadian provinces.

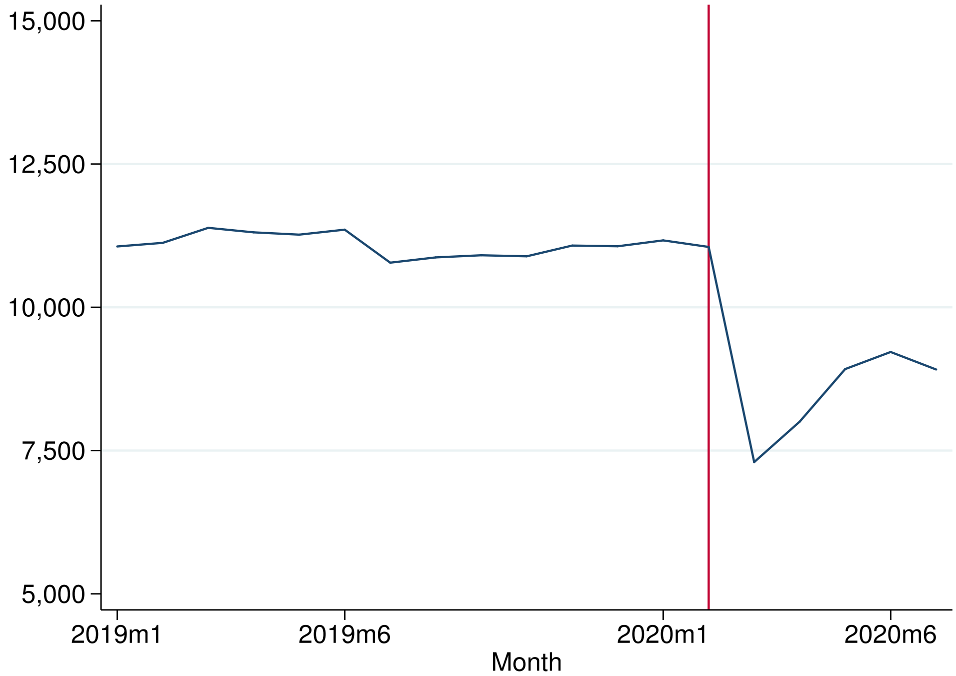

Figure B.2: Sample size of incoming rotations. Source: Authors' calculations. Number of individuals in the incoming rotation group in each month of the LFS. The red vertical line indicates February 2020. Monthly cross-sectional samples of individuals aged 20 to 64, from all ten Canadian Provinces, excluding full-time members of the armed forces. See Sections 2.3 and 2.4 for further information about samples used in the analysis.

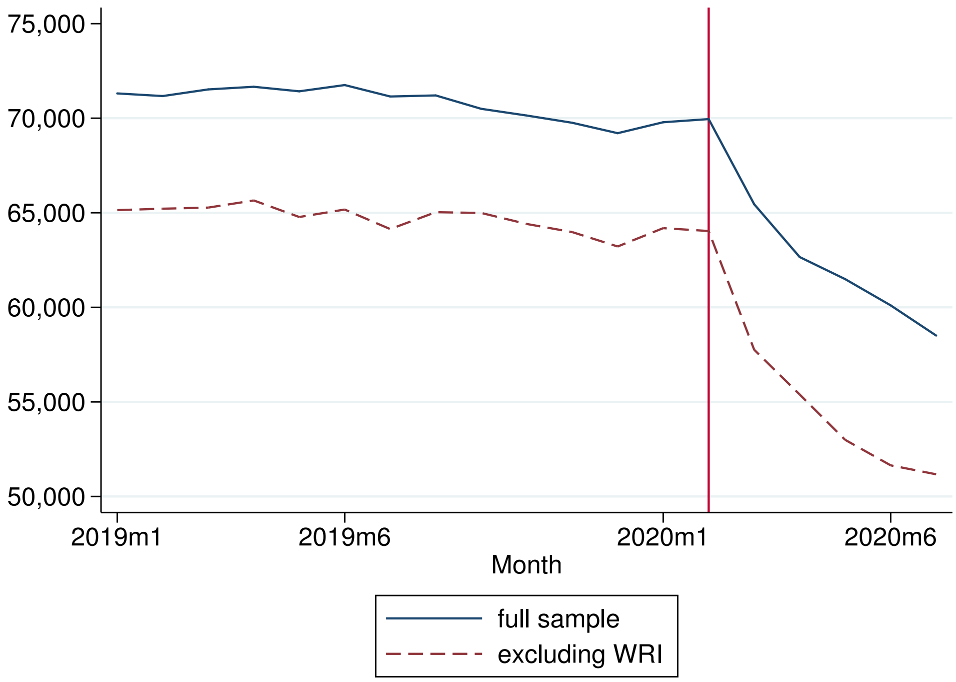

Figure B.3: Sample size with and without whole record imputation (WRI). Source: LFS and authors' calculations. The blue plain line shows the total number of individuals in the sample and the red dotted line shows the number of individuals after excluding those for whom LFS information has been fully imputed. The red vertical line indicates February 2020. Monthly cross-sectional sample of individuals of individuals aged 20 to 64, from all ten Canadian Provinces, excluding full-time members of the armed forces. See Section 2.3 and 2.4 for further information about WRI issues in the LFS.

Table B.1: Labour-force estimates for select samples, 2020

	February		April		June	
	all	no WRI	all	no WRI	all	no WRI
Employment rate (%)						
Cross-section	76.5	76.7	65.5	65.6	71.8	72.1
Consecutive-month panel	76.6	77.2	65.9	66.0	72.1	72.1
Three-month-span	76.8	77.2	66.5	66.4	–	–
Five-month-span	77.4	77.9	–	–	–	–
Unemployment rate						
Cross-section	5.5	5.4	12.9	13.0	11.0	10.7
Consecutive-month panel	5.5	5.3	12.7	12.5	10.9	10.7
Three-month-span	5.6	5.2	12.5	12.6	–	–
Five-month-span	5.2	4.9	–	–	–	–
Absent from work (employment share)						
Cross-section	7.3	7.4	20.9	20.7	11.8	11.9
Consecutive-month panel	7.3	7.4	20.7	20.4	11.8	11.8
Three-month-span	7.3	7.3	21.0	20.3	–	–
Five-month span	7.5	7.3	–	–	–	–
Temporary unemployment (labour-force share)						
Cross-section	0.3	0.3	6.9	7.1	3.2	3.1
Consecutive-month panel	0.3	0.3	6.9	6.8	3.1	2.9
Three-month-span	0.3	0.3	6.9	7.0	–	–
Five-month span	0.3	0.2	–	–	–	–
Participation rate						
Cross-section	81.0	81.0	75.2	75.4	80.7	80.7
Consecutive-month panel	81.1	81.5	75.5	75.5	80.9	80.7
Three-month span	81.3	81.5	75.9	75.6	–	–
Five-month span	81.6	81.9	–	–	–	–

Notes: select labour-force estimates, for a subset of the samples that are used in the analysis. "Cross-section" refers to the cross-sectional LFS monthly samples. "Consecutive-month" panel refers to the samples of individuals observed for two consecutive months; "three-month-span" panel refers to (1) the sample of individual observed in February and April (i.e. spanning three months from February) and (2) the sample of individuals observed in April and June; "five-month span" refers to the sample of individuals observed in February, April, and June. All samples are for individuals aged 20 to 64, excluding full-time members of the armed forces. In the case of the panel samples, the summary statistics are for the first month in which individuals are observed. "Temporary unemployment" refers to estimates of unemployed workers counted as being on temporary layoff, expressed in terms of the labour force. "Employed, at work" is for the share of employed workers who did not declare being absent from work. In all cases, we show the statistics for the entire sample (under column "all"), and after excluding individuals whose labour force information was fully imputed, i.e. with no whole record imputation (no WRI). All estimations are weighted.

Table B.2: Worker flows in and out of employment

	In and out of employment (%)									
	2020					2019				
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows	7.5	13.9	6.7	4.6	5.0	2.3	2.4	2.1	2.0	3.8
Inflows	2.5	3.8	9.8	10.2	5.5	2.3	3.2	3.8	2.5	2.4
Net change	-5.0	-10.1	3.2	5.6	0.5	0.4	0.8	1.7	0.6	-1.4
Excess flows	4.9	7.5	13.3	9.1	9.9	4.6	4.9	4.2	4.0	4.8
	In and out of employment, at work (including vacation, etc.) (%)									
	2020					2019				
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows	17.3	21.9	5.9	4.5	5.6	3.6	4.1	3.4	3.4	5.2
Inflows	3.2	6.5	15.6	15.8	8.2	3.9	4.8	5.5	3.9	3.8
Net change	-14.0	-15.4	9.8	11.3	2.6	0.3	0.8	2.1	0.5	-1.4
Excess flows	6.5	13.0	11.7	8.9	11.2	7.2	8.1	6.9	6.9	7.7

Notes: estimations of monthly employment inflows and outflows, based on two-consecutive-month LFS panels for individuals aged 20 to 64, excluding full-time members of the armed forces. Outflows for period $t-1$ and t is an estimation of the total number of workers employed in $t-1$ and non-employed in t, in terms of total employment in $t-1$. Inflows are computed similarly but are based on estimates of workers transiting from non-employment to employment. Excess flows are defined as total reallocation (i.e hiring + separation) from period $t-1$ to period t, minus the absolute value of the net change (i.e. —hiring - separation—). Flows in and out of the "employment, at work" category are estimated following the same approach. Note, however, that we include workers in the stock of reference who are absent due to vacation, parental leave, and labour conflicts, i.e. we do not consider flows associated with these motives. See appendix B.1.1 for details. All totals are estimated using samples of individuals observed for two consecutive months. All estimations are weighted.

Table B.3: Composition of flows in and out of employment, 2020

In and out of employment (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Temporary unemployment	32.5	37.8	32.8	24.5	12.8
Search unemployment	15.9	11.5	16.9	24.9	30.7
NILF	50.0	49.2	48.3	48.7	55.0
Inflows					
Temporary unemployment	4.2	26.8	40.2	35.7	29.7
Search unemployment	36.9	17.6	11.0	17.5	28.3
NILF	56.2	53.7	45.6	42.4	38.9

In and out of employment, at work (incl. vacation) (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Absent from work (excl. vac.)	58.2	43.8	45.0	47.0	39.9
Unemployment	21.5	29.8	30.5	25.1	25.7
NILF	20.3	26.4	24.5	27.9	34.4
Inflows					
Absent from work (excl. vac.)	39.1	66.1	47.1	41.2	43.9
Unemployment	30.4	15.7	30.9	36.3	36.1
NILF	30.5	18.3	22.0	22.5	20.0

Notes: analysis of the compositions of employment inflows and outflows. Two-consecutive-month samples for individuals aged 20 to 64, excluding full-time members of the armed forces. Total flows are computed as in tables B.2. Share of flows that are towards/from temporary unemployment (i.e. unemployment, on temporary layoff), search unemployment (i.e. unemployed workers declaring searching for a job), and non-participation. The bottom panel analysis the composition of flows in and out of the LFS category "employment, at work", but including workers on vacation, on parental leave, or absent due to labour conflicts. Similarly, the category "absent from work" exclude workers absent due to these motives. See notes in Table B.2 and Appendix B.1.1 for more details. All estimations are weighted.

Table B.4: Monthly transition probabilities of the out of work (%), 2020

	Apr-May	May-Jun	Jun-Jul
From temporary unemployment			
Employment	49.8	57.0	47.4
Employment, at work (incl. vac.)	35.4	46.8	34.7
Search unemployment	5.5	8.5	15.3
NILF	15.9	11.1	18.2
From absent from work (excl. vac.)			
Employment, at work (incl. vac.)	33.8	43.0	37.8
Search unemployment	2.6	3.9	6.5
NILF	11.4	8.7	11.3

Notes: estimation of monthly transition probabilities of workers who are either on temporary layoff or who are absent from work, except due to vacation, parental leave, or labour-conflict motives. Two-consecutive-month LFS panels for individuals aged 20 to 64, excluding full-time members of the armed forces. The category "employment, at work", includes workers absent due to vacation, parental leave, or labour conflicts. All estimations are weighted.

Table B.5: Monthly transition probabilities of the early-outbreak job losers (%)

	May	June	July
From non-employment in April			
Employment	37.0	58.0	64.1
Employment, at work (incl. vac.)	24.8	48.4	57.6
Search unemployment	14.1	12.7	12.5
NILF	31.3	17.5	17.9
From temporary unemployment in April			
Employment	52.5	71.4	75.9
Employment, at work (incl. vac.)	37.1	62.9	68.1
Search unemployment	4.6	6.5	7.5
NILF	12.0	7.2	9.0
From absent from work in April (excl. vac.)			
Employment	80.6	85.3	83.9
At work (incl. vac.)	33.1	57.1	66.9
Search unemployment	2.3	2.6	3.9
NILF	9.5	7.8	10.4
From NILF in April			
Employment	29.6	51.6	58.4
Unemployment	8.9	21.7	26.2

Notes: estimation of monthly transition probabilities of workers who lost their job between February and April, conditional on their detailed labour-force status in April. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces. Individuals employed in February, non-employed in April, and observed in May, June, or July. Transitions to "employment, at work" include transitions to absence due to vacation, parental leave, and labour conflicts. Transitions conditional on being absent from work exclude these motives. All estimations are weighted.

Table B.6: Probability of no occupation (NOC) or industry (NAICS) switch of the reemployed spring-2020 job losers (%)

	May	June	July
From non-employment in April			
Occupation, 2 digits	76.5	72.5	69.9
Occupation, 4 digits	71.5	66.7	62.5
Industry, 2 digits	88.0	84.6	83.2
Industry, 5 digits	77.1	73.1	71.0
From temporary unemployment in April			
Occupation, 2 digits	78.5	78.1	79.1
Occupation, 4 digits	73.6	74.1	73.0
Industry, 2 digits	89.5	90.6	91.2
Industry, 5 digits	77.5	80.8	76.3
From absent from work in April (excl. vac.)			
Occupation, 2 digits	78.4	75.6	76.6
Occupation, 4 digits	73.4	72.4	74.1
Industry, 2 digits	92.0	90.7	89.5
Industry, 5 digits	78.2	76.9	77.0

Notes: estimation of the probability of no occupation/industry switch conditional on reemployment after a job loss in the spring. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces. Individuals employed in February, non-employed in April, and employed again in May, June, or July. We consider the National Occupational Classification (NOC) and the North-American Industry Classification System (NAICS). "Absent from work" excludes absences due to vacation, parental leave, and labour conflicts. All estimations are weighted.

Table B.7: Transition probabilities conditional on LFS status in February (%)

	2020			2019		
	May	June	July	May	June	July
From employment						
Employment	84.2	88.4	88.4	96.5	96.2	94.0
Unemployment	8.1	6.4	5.4	1.6	1.6	2.2
Employment, absent from work	10.6	7.6	5.5	2.8	2.9	2.5
Temporary unemployment	4.6	3.0	1.6	0.1	0.1	0.3
From unemployment						
Employment	29.0	36.8	42.8	46.6	51.5	55.8
Unemployment	40.8	40.0	37.5	38.4	33.0	26.4
From NILF, wants a job						
Employment	15.0	18.5	16.2	25.9	25.9	31.5
Unemployment	17.6	21.8	18.8	14.4	14.3	16.3
From NILF						
Employment	10.9	13.1	14.4	13.7	16.3	17.0
Unemployment	7.9	8.9	9.0	5.4	5.1	4.8

Notes: estimation of transition probabilities of individuals in various labour force states as of February, over the May-to-June period. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces. Individuals non-employed in February and observed in May, June, or July, for 2019 and 2020. The category "NILF, wants a job" refer to individuals classified as not in the labour force in the LFS, but who declare wanting a job (either full-time or part-time). All estimations are weighted.

Table B.8: Employment-separation probability in the early phase of the outbreak

	To non-employment			To non-participation		
	(1)	(2)	(3)	(4)	(5)	(6)
Post COVID-19	0.142***	0.150***	0.111***	0.073***	0.085***	0.066***
	(0.003)	(0.010)	(0.016)	(0.002)	(0.008)	(0.013)
Select interaction terms with post COVID-19 indicator:						
Female		0.035***	0.036***		0.027***	0.023***
		(0.007)	(0.007)		(0.005)	(0.006)
20 to 29		0.080***	0.047***		0.054***	0.036***
		(0.010)	(0.010)		(0.008)	(0.008)
40 to 49		-0.024***	-0.007		-0.003	0.005
		(0.008)	(0.008)		(0.006)	(0.006)
50 to 64		-0.017**	0.006		-0.006	0.005
		(0.008)	(0.008)		(0.006)	(0.006)
Dropout		0.023	0.013		-0.003	-0.006
		(0.016)	(0.016)		(0.011)	(0.011)
College		-0.025***	-0.001		-0.017**	-0.003
		(0.009)	(0.009)		(0.007)	(0.007)
Bachelor degree +		-0.088***	-0.041***		-0.050***	-0.025***
		(0.016)	(0.016)		(0.010)	(0.011)
Female × Young child		-0.025**	-0.009		-0.015	-0.006
		(0.012)	(0.012)		(0.010)	(0.010)
Tenure: 0-11 months			0.053***			0.029***
			(0.010)			(0.008)
Tenure: 12-35 months			0.037***			0.016**
			(0.009)			(0.007)
Self-employed			-0.098***			-0.040***
			(0.008)			(0.007)
Union			-0.078***			-0.047***
			(0.007)			(0.005)
Constant	0.032***	0.042***	0.021***	0.018***	0.022***	0.013**
	(0.001)	(0.005)	(0.007)	(0.001)	(0.004)	(0.005)
adj. R^2	0.055	0.082	0.119	0.026	0.045	0.066
N	67,846	67,846	67,846	67,846	67,846	67,846

Notes: estimation of the effect of COVID-19 on job separations between February and April. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces, who are employed in February and observed in April, for 2019 and 2020. In columns (1)-(3), the dependent variable is an indicator for non-employment in April. In columns (4)-(6), it is an indicator for non-participation in April. Columns (1) and (4) show the result of regressions with only a dummy for 2020. Columns (2) and (5) are for regressions with socio-demographic variables (as of February) and their interactions with the 2020 indicator variable. Columns (3) and (6) also include job characteristics (as of February) and their interaction with the year-2020 indicator. Only select interacted variable coefficients point estimate and standard errors are shown in the table. Tables for the full set of coefficients are available upon request. All regressions are weighted. Huber-White robust standard errors shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table B.9: Reemployment probability of the early-outbreak job losers

	May		June		July	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.051**	-0.018	-0.072***	-0.032	0.004	0.029
	(0.021)	(0.024)	(0.028)	(0.032)	(0.039)	(0.047)
20 to 29	-0.048*	-0.015	-0.061	-0.026	-0.037	-0.030
	(0.028)	(0.029)	(0.038)	(0.038)	(0.053)	(0.054)
40 to 49	-0.009	-0.025	-0.060	-0.077*	-0.073	-0.083
	(0.032)	(0.031)	(0.040)	(0.040)	(0.058)	(0.059)
50 to 64	0.033	0.008	-0.051	-0.069*	-0.134**	-0.141***
	(0.029)	(0.029)	(0.036)	(0.037)	(0.053)	(0.054)
Dropout	0.017	0.011	0.021	0.017	-0.050	-0.051
	(0.040)	(0.039)	(0.048)	(0.047)	(0.085)	(0.085)
College	0.005	-0.002	0.031	0.025	0.069	0.067
	(0.024)	(0.024)	(0.031)	(0.031)	(0.044)	(0.045)
Bachelor degree +	-0.009	-0.011	0.002	0.012	-0.025	0.004
	(0.028)	(0.029)	(0.038)	(0.040)	(0.055)	(0.058)
Female × Young child	0.009	-0.001	-0.020	-0.029	-0.057	-0.047
	(0.041)	(0.041)	(0.054)	(0.052)	(0.072)	(0.072)
Tenure: 0-11 months		-0.078***		-0.110***		-0.016
		(0.026)		(0.034)		(0.049)
Tenure: 12-35 months		-0.067***		-0.029		0.028
		(0.025)		(0.033)		(0.047)
Self-employed		0.087**		0.085*		0.010
		(0.040)		(0.052)		(0.076)
Union		0.041		-0.024		0.009
		(0.027)		(0.034)		(0.047)
Constant	0.346***	0.399***	0.582***	0.689***	0.610***	0.742***
	(0.033)	(0.064)	(0.042)	(0.077)	(0.062)	(0.102)
adj. R^2	0.049	0.063	0.041	0.059	0.033	0.050
N	4,077	4,077	2,549	2,549	1,202	1,202

Notes: reemployment probability of spring-job-losers conditional on individual and job characteristics. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces, who are employed in February and non-employed in April, and observed in May, June, or July 2020. The dependent variable is an indicator for employment in May, June, or July. Columns (1), (3), (5) shows results of regressions with socio-demographic variables in February, and (2), (4), and (6) also include job characteristics in February. The table shows select coefficients point estimates and standard errors. Tables for the full set of coefficients are available upon request. All regressions are weighted. Huber-White robust standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table B.10: Transitions out of the labour force of the early-outbreak job losers

	May		June		July	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.030	0.001	0.013	-0.021	0.023	0.000
	(0.021)	(0.024)	(0.020)	(0.024)	(0.033)	(0.039)
20 to 29	0.003	-0.012	-0.028	-0.036	0.022	0.018
	(0.028)	(0.028)	(0.026)	(0.026)	(0.038)	(0.038)
40 to 49	0.048	0.047	0.053*	0.050	0.079*	0.068
	(0.031)	(0.031)	(0.031)	(0.031)	(0.045)	(0.046)
50 to 64	0.023	0.027	0.096***	0.090***	0.127***	0.125***
	(0.027)	(0.028)	(0.028)	(0.029)	(0.041)	(0.042)
Dropout	-0.026	-0.014	-0.009	-0.003	0.064	0.054
	(0.034)	(0.034)	(0.037)	(0.037)	(0.062)	(0.061)
College	0.020	0.022	-0.031	-0.032	-0.039	-0.046
	(0.023)	(0.024)	(0.023)	(0.023)	(0.033)	(0.034)
Bachelor degree +	0.024	0.022	-0.021	-0.036	0.048	0.005
	(0.029)	(0.030)	(0.029)	(0.031)	(0.046)	(0.049)
Female × Young child	0.031	0.033	0.097**	0.089*	0.105*	0.089
	(0.041)	(0.040)	(0.046)	(0.046)	(0.064)	(0.063)
Tenure: 0-11 months		0.032		0.013		0.007
		(0.025)		(0.024)		(0.036)
Tenure: 12-35 months		0.007		-0.027		-0.003
		(0.024)		(0.024)		(0.034)
Self-employed		0.053		0.027		0.126*
		(0.040)		(0.043)		(0.070)
Constant	0.312***	0.262***	0.191***	0.177***	0.159***	0.102
	(0.032)	(0.058)	(0.032)	(0.058)	(0.049)	(0.088)
adj. R^2	0.038	0.047	0.043	0.053	0.048	0.071
N	4,077	4,077	2,549	2,549	1,202	1,202

Notes: probability of transition out of the labour force of spring-job-losers, conditional on individual and job characteristics. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces, who are employed in February and non-employed in April, and observed in May, June, or July 2020. The dependent variable is an indicator for non-participation in May, June, or July. Columns (1), (3), (5) shows results of regressions with socio-demographic variables in February, and (2), (4), and (6) also include job characteristics in February. The table shows select coefficients point estimates and standard errors. Tables for the full set of coefficients are available upon request. All regressions are weighted. Huber-White robust standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table B.11: Employment probability of individuals unemployed before COVID-19

	May		June		July	
	(1)	(2)	(3)	(4)	(5)	(6)
Post COVID-19	-0.177***	-0.225***	-0.147***	-0.268***	-0.130***	-0.340**
	(0.024)	(0.077)	(0.032)	(0.099)	(0.047)	(0.148)
Select interaction terms with post COVID-19 indicator:						
Female		0.023		-0.003		-0.026
		(0.052)		(0.069)		(0.108)
20 to 29		0.003		-0.004		0.265*
		(0.071)		(0.094)		(0.150)
40 to 49		0.062		0.073		0.080
		(0.076)		(0.103)		(0.149)
50 to 64		0.070		0.133		0.030
		(0.070)		(0.091)		(0.134)
Dropout		0.093		0.140		0.076
		(0.079)		(0.102)		(0.168)
College		-0.006		0.075		0.100
		(0.061)		(0.078)		(0.116)
Bachelor degree +		0.030		0.072		0.053
		(0.068)		(0.089)		(0.133)
Female × Young child		0.169*		0.096		0.152
		(0.097)		(0.133)		(0.211)
Unemp.: 6-11 months		-0.122		-0.041		-0.003
		(0.087)		(0.120)		(0.179)
Unemp. : 12 months +		0.087		0.024		0.068
		(0.069)		(0.084)		(0.143)
Constant	0.466***	0.539***	0.515***	0.575***	0.558***	0.725***
	(0.018)	(0.056)	(0.023)	(0.074)	(0.034)	(0.108)
adj. R^2	0.033	0.080	0.021	0.095	0.016	0.078
N	3,005	2,899	1,886	1,813	853	818

Notes: estimation of the effect of COVID-19 on the employment probability of individuals unemployed before COVID-19. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces, who are unemployed in February and observed in May, June, or July, for 2019 and 2020. The dependent variable is an indicator for employment in May, June, or July. Column (1), (3), and (5) show results of regressions with an indicator for the year 2020 only. Columns (2), (4), and (6) are for regressions with socio-demographic variables in February, alone and in interaction with the year-2020 indicator. Only select interacted-variable coefficient point estimates and standard errors are shown in the table. Tables for the full set of coefficients are available upon request. All regressions are weighted. Huber-White standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table B.12: Employment probability of NILF individuals before COVID-19

	May		June		July	
	(1)	(2)	(3)	(4)	(5)	(6)
Post COVID-19	-0.029***	-0.073**	-0.032***	-0.031	-0.025	-0.037
	(0.008)	(0.032)	(0.011)	(0.042)	(0.017)	(0.067)
Select interaction terms with post COVID-19 indicator:						
Female		0.024		-0.008		-0.015
		(0.018)		(0.023)		(0.035)
20 to 29		-0.036		-0.097**		-0.011
		(0.033)		(0.044)		(0.067)
40 to 49		-0.001		-0.035		0.060
		(0.033)		(0.046)		(0.066)
50 to 64		0.023		-0.015		0.026
		(0.028)		(0.037)		(0.057)
Dropout		0.034*		0.048*		-0.000
		(0.020)		(0.027)		(0.038)
College		0.013		0.041		0.064
		(0.021)		(0.028)		(0.041)
Bachelor degree +		0.039		0.057*		0.084*
		(0.024)		(0.032)		(0.048)
Female × Young child		0.014		0.024		-0.061
		(0.029)		(0.041)		(0.058)
Constant	0.137***	0.212***	0.163***	0.230***	0.170***	0.301***
	(0.006)	(0.024)	(0.008)	(0.030)	(0.012)	(0.049)
adj. R^2	0.002	0.063	0.002	0.097	0.001	0.115
N	12,722	12,722	8,237	8,237	3,879	3,879

Notes: estimation of the effect of COVID-19 on the employment probability of individuals out of the labour force before COVID-19. Sample for individuals aged 20 to 64, excluding full-time members of the armed forces, who are NILF in February and observed in May, June, or July, for 2019 and 2020. The dependent variable is an indicator for employment in May, June, or July. Column (1), (3), and (5) show results of regressions with an indicator for the year 2020 only. Columns (2), (4), and (6) are for regressions with socio-demographic variables in February, alone and in interaction with the year-2020 indicator. Only select interacted-variable coefficient point estimates and standard errors are shown in the table. Tables for the full set of coefficients are available upon request. All regressions are weighted. Huber-White standard errors are shown in parentheses. * denotes statistical significance at 10%; ** significance at 5%; *** significance at 1%.

Table B.13: Worker flows in and out of employment, 2018

In and out of employment (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows	2.2	2.6	2.2	1.9	3.6
Inflows	2.4	3.0	3.5	2.6	2.6
Net change	0.2	0.3	1.4	0.7	-1.0
Excess flows	4.4	5.3	4.3	3.8	5.1
In and out of employment, at work (incl. vac.) (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows	3.7	4.0	3.2	3.1	4.9
Inflows	3.9	4.7	5.1	3.8	3.9
Net change	0.2	0.7	1.9	0.7	-1.1
Excess flows	7.3	8.0	6.5	6.1	7.7

Notes: estimations of monthly employment inflows and outflows, based on two-consecutive-month LFS panels for individuals aged 20 to 64, excluding full-time members of the armed forces. Outflows for period $t-1$ and t is an estimation of the total number of workers employed in $t-1$ and non-employed in t, in terms of total employment in $t-1$. Inflows are computed similarly but are based on estimates of workers transiting from non-employment to employment. Excess flows are defined as total reallocation (i.e hiring + separation) from period $t-1$ to period t, minus the absolute value of the net change (i.e. —hiring - separation—). Flows in and out of the "employment, at work" category are estimated following the same approach. Note, however, that we include workers in the stock of reference who are absent due to vacation, parental leave, and labour conflicts, i.e. we do not consider flows associated with these motives. See Appendix B.1.1 for details. All totals are estimated using samples of individuals observed for two consecutive months. All estimations are weighted.

Table B.14: Composition of flows in and out of employment, 2018

In and out of employment (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Temporary unemployment	11.9	3.5	3.4	1.7	6.4
Search unemployment	26.3	31.3	29.8	38.3	32.4
NILF	57.4	61.9	62.3	55.5	58.9
Inflows					
Temporary unemployment	5.4	8.2	2.7	4.4	1.4
Search unemployment	36.0	39.4	31.4	42.4	42.0
NILF	55.0	46.2	55.5	44.0	49.3

In and out of employment, at work (incl. vac.) (%)					
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Absent from work	47.6	42.7	41.5	44.2	32.3
Unemployment	24.9	23.5	22.8	26.6	28.9
NILF	27.5	33.8	35.7	29.2	38.8
Inflows					
Absent from work	44.1	38.8	32.6	35.6	36.6
Unemployment	27.2	34.0	31.1	38.6	34.4
NILF	28.8	27.2	36.3	25.8	29.1

Notes: Analysis of the composition of employment inflows and outflows. Two-consecutive-month samples, for individuals aged 20 to 64, excluding full-time members of the armed forces. Total flows are computed as in Tables B.2. Share of flows that are towards/from temporary unemployment (i.e. unemployment, on temporary layoff), search unemployment (i.e. unemployed workers declaring searching for a job), and non-participation. The bottom panel analysis the composition of flows in and out of the LFS category "employment, at work", but including workers on vacation, on parental leave, or absent due to labour conflicts. Similarly, the category "absent from work" exclude workers absent due to these motives. See notes in Table B.2 and Appendix B.1.1 for more details. All estimations are weighted.

Table B.15: Composition of flows in and out of employment, 2019

In and out of employment (%)	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Temporary unemployment	5.9	3.6	3.1	3.8	6.5
Search unemployment	30.4	33.3	40.5	31.2	32.2
NILF	58.3	57.5	53.7	61.8	59.2
Inflows					
Temporary unemployment	4.5	4.8	3.4	3.6	2.7
Search unemployment	36.0	38.3	35.9	40.2	39.6
NILF	54.2	48.6	50.8	47.3	52.2

In and out of employment, at work (incl. vac.) (%)	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Outflows					
Absent from work	46.1	46.2	46.1	47.2	30.9
Unemployment	25.0	24.9	26.4	21.6	29.8
NILF	28.9	28.9	27.5	31.1	39.3
Inflows					
Absent from work	44.0	37.4	31.9	38.5	40.7
Unemployment	27.2	34.5	34.8	34.2	30.7
NILF	28.8	28.1	33.4	27.2	28.6

Notes: Analysis of the compositions of employment inflows and outflows. See notes of Table B.14 for details.

Table B.16: Employment inflows and outflows, one-digit NOC occupation groups

	2020					2019				
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Management										
Separation	3.6	7.3	3.6	3.4	3.3	1.6	2.3	1.1	1.4	1.5
Hiring	2.0	2.2	4.4	4.6	2.4	1.1	1.3	1.9	1.0	1.0
Net change	-1.6	-5.1	0.8	1.3	-0.9	-0.5	-1.0	0.8	-0.4	-0.5
Excess flows	4.1	4.3	7.2	6.8	4.7	2.2	2.5	2.2	0.2	2.1
Business, finance and administration										
Separation	4.7	10.6	5.1	4.3	3.2	2.0	1.6	2.0	1.6	2.7
Hiring	2.2	2.4	7.0	6.4	4.0	1.8	2.6	2.8	2.3	1.4
Net change	-2.6	-8.2	1.9	2.1	0.8	-0.3	1.0	0.8	0.8	-1.2
Excess flows	4.3	4.8	10.2	8.5	6.4	3.6	3.2	3.9	3.2	2.9
Natural and applied sciences										
Separation	2.4	8.0	3.8	2.1	1.7	1.1	1.8	1.8	1.2	1.3
Hiring	1.9	1.2	5.8	3.0	2.1	1.4	1.6	3.5	1.9	1.9
Net change	-0.5	-6.9	2.1	0.9	0.4	0.3	-0.2	1.7	0.8	0.6
Excess flows	3.8	2.4	7.6	4.3	3.5	2.2	3.3	3.6	2.3	2.6
Health										
Separation	5.1	6.6	4.3	3.4	2.2	1.5	1.6	1.5	1.2	2.1
Hiring	1.5	2.5	4.4	7.5	3.7	1.6	2.0	2.2	2.5	1.9
Net change	-3.6	-4.1	0.1	4.1	1.4	0.5	0.4	0.7	1.3	-0.2
Excess flows	3.1	5.0	8.6	6.8	4.4	3.1	3.2	3.0	2.4	3.7
Education, law and social, community and government services										
Separation	9.5	8.8	6.4	4.6	12.7	2.2	2.2	2.8	1.8	12.1
Hiring	1.5	5.6	6.2	7.3	3.4	1.7	2.6	2.7	2.1	2.1
Net change	-7.9	-3.2	-0.2	2.7	-9.3	-0.5	0.4	-0.7	0.3	-10.0
Excess flows	3.0	11.1	12.4	9.1	6.8	3.5	4.3	5.5	3.6	4.3

Notes: estimations of monthly employment inflows and outflows by one-digit National Occupational Classification (NOC) occupation groups. Two-consecutive-month LFS panels of non-military, of age 20 to 64. Outflows for period $t-1$ and t is an estimation of the total number of workers employed in $t-1$ and non-employed in t, in terms of total employment in $t-1$. Inflows are computed similarly but are based on estimates of workers transiting from non-employment to employment. Excess flows are defined as total reallocation (i.e hiring + separation) from period $t-1$ to period t, minus the absolute value of the net change (i.e. —hiring - separation—). See Appendix B.1.1 for details. All totals are estimated using samples of individuals observed for two consecutive months. All estimations are weighted.

Table B.17: Employment inflows and outflows, one-digit NOC occupation groups (continued)

	2020					2019				
	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul
Arts, culture, recreation and sports										
Separation	10.7	15.9	10.1	8.6	6.7	2.7	3.7	2.7	2.2	4.8
Hiring	2.2	5.7	7.1	14.0	8.9	3.2	4.2	5.2	5.1	5.2
Net change	-8.6	-10.2	-3.0	5.3	2.2	0.6	0.5	2.4	2.9	0.4
Excess flows	4.3	11.3	14.2	17.3	13.3	5.3	7.4	5.5	4.4	9.6
Sales and services										
Separation	13.1	22.4	11.1	6.7	6.0	2.4	2.9	2.1	2.6	3.2
Hiring	2.5	5.6	14.5	19.0	11.2	3.0	3.7	4.0	2.7	3.4
Net change	-10.6	-16.8	3.4	12.3	5.2	0.5	0.8	1.9	0.1	0.2
Excess flows	5.0	11.1	22.2	13.5	12.0	4.9	5.7	4.2	5.2	6.3
Trades, transport and equipment										
Separation	7.1	20.0	8.4	4.7	4.0	3.3	3.3	2.6	2.5	3.3
Hiring	4.5	4.1	17.2	14.4	6.1	4.0	5.6	6.4	3.2	3.1
Net change	-2.6	-15.7	8.8	9.8	2.1	0.7	2.3	3.8	0.7	-0.3
Excess flows	8.9	8.3	16.8	9.3	8.0	6.6	6.5	5.3	4.9	6.1
Natural resources and agriculture										
Separation	8.8	16.1	5.3	6.6	7.9	6.6	6.5	2.8	4.9	6.9
Hiring	7.2	7.1	25.5	20.6	8.6	5.1	11.1	17.0	8.3	4.4
Net change	-1.5	-9.0	20.2	14.1	0.7	-1.6	4.6	14.2	3.4	-2.5
Excess flows	14.5	14.1	10.5	13.1	15.9	10.1	13.1	5.5	9.8	8.7
Manufacturing										
Separation	6.5	20.3	7.3	4.2	3.3	3.0	2.3	1.9	2.2	2.6
Hiring	1.9	3.9	19.4	12.5	5.5	1.9	3.0	3.7	1.8	2.3
Net change	-4.6	-16.4	12.0	8.2	2.2	-1.0	0.7	1.8	-0.4	-0.3
Excess flows	3.7	7.9	14.6	8.5	6.5	3.8	4.6	3.9	3.6	4.6

Notes: estimations of monthly employment inflows and outflows by one-digit NOC occupation groups. See notes of Table B.16 for details.

Appendix C

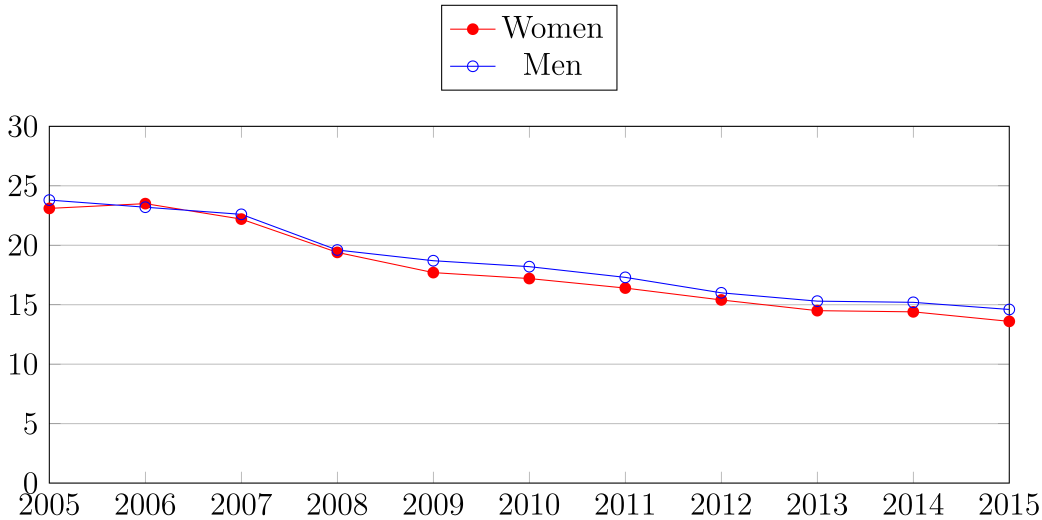

Figure C.1: Percentage of young workers aged 24 to 35 in 2005 who had a separation from main job, by gender, 2005 – 2015

Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

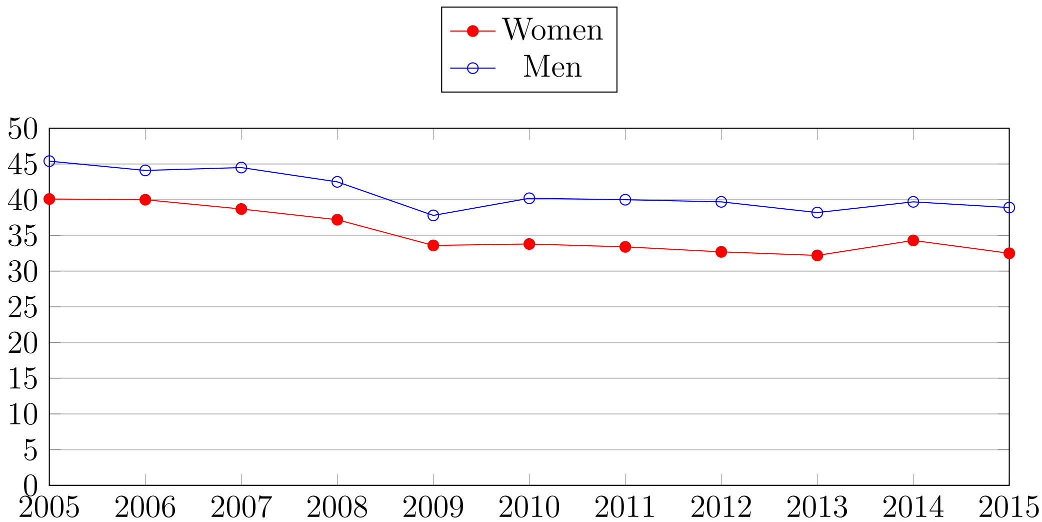

Figure C.2: Percentage of young workers aged 24 to 35 in 2005 who had a permanent separation from main job, conditional on having a separation, by gender, 2005 – 2015

Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

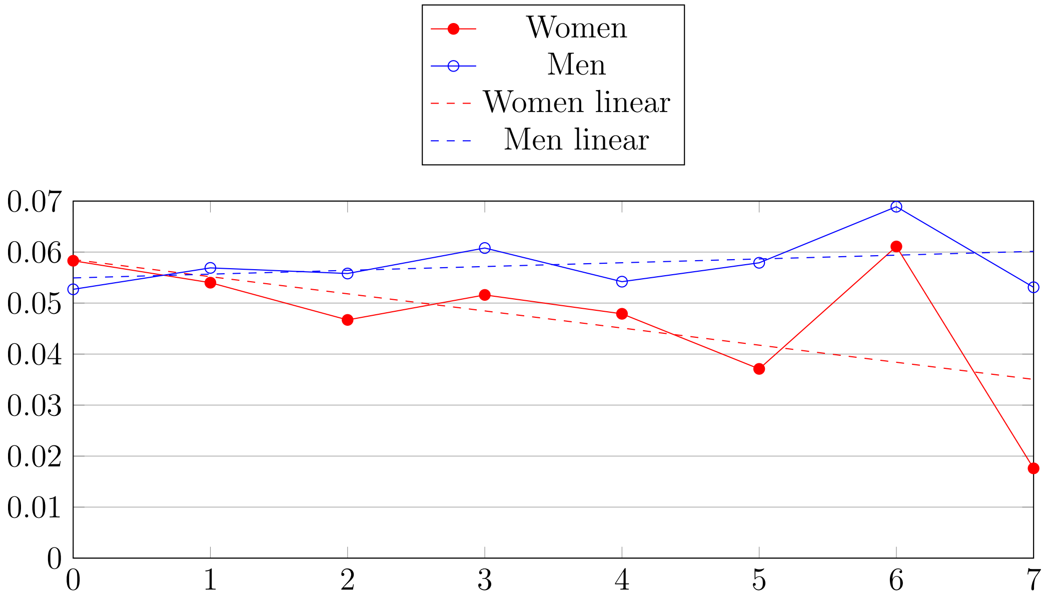

Figure C.3: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 who had a permanent separation from main job, conditional on having a separation, by gender and the total number (instance) of permanent separations, 2005 – 2015

Notes: Women linear and men linear indicates the trend line of the wage growth over the total number of separations.
Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

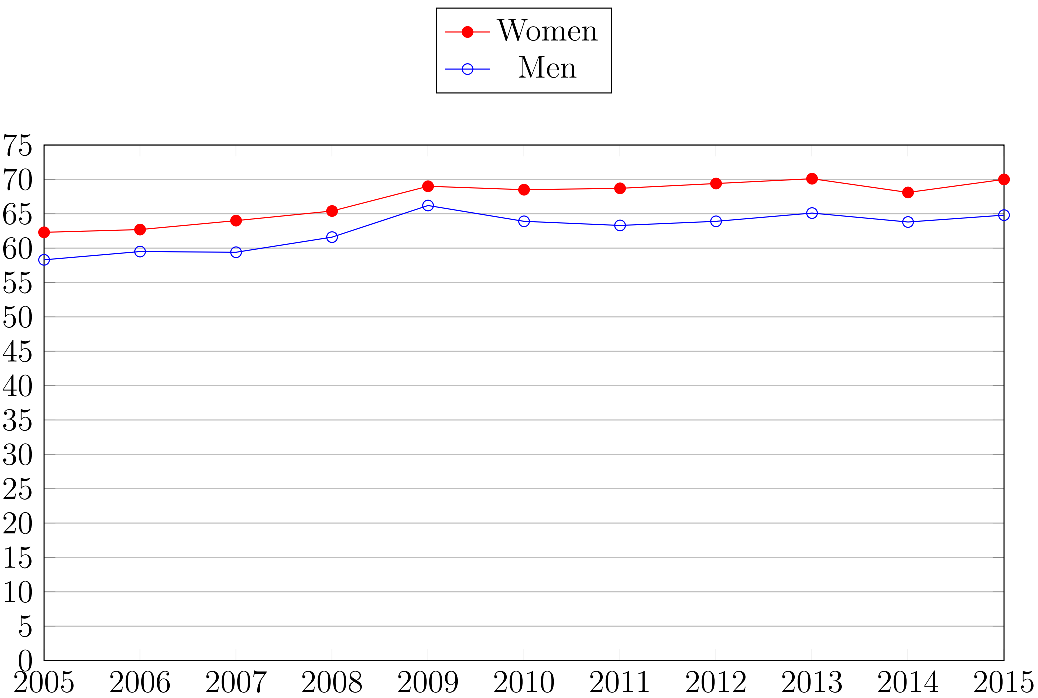

Figure C.4: Percentage of young workers aged 24 to 35 in 2005 that had a temporary separation from main job, conditional on having a separation, by gender, 2005 - 2015

Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

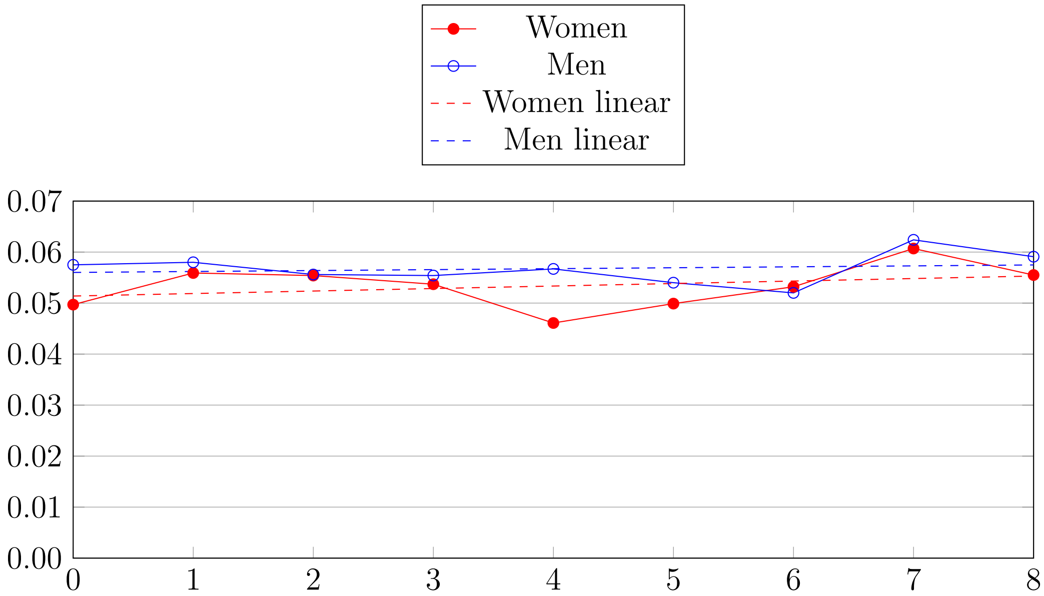

Figure C.5: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 who had a temporary separation from main job, conditional on having a separation, by gender and the total number (instance) of temporary separations, 2005 – 2015

Notes: Women linear and men linear indicates the trend line of the wage growth over the total number of separations.
Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

Table C.1: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005, by the status of permanent job separation and gender, 2005-2015.

	Status of permanent job separation		
	Female with no job separation	Female with at least one job separation	Differences by column
Weekly wage annual growth rate	5.95%	5.11%	0.84%***
	Male with no job separation	Male with at least one job separation	
Weekly wage annual growth rate	5.25%	5.72%	-0.47%**
Differences by row	0.70%***	-0.61%***	

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.2: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 by the status of temporary job separation and gender, 2005-2015.

	Status of permanent job separation		
	Female with no job separation	Female with at least one job separation	Differences by column
Weekly wage annual growth rate	4.74%	5.48%	-0.74%*
	Male with no job separation	Male with at least one job separation	
Weekly wage annual growth rate	5.50%	5.60%	-0.10%
Differences by row	-0.76%	-0.12%	

Notes: Data presented in percentage.*: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.3: Percentage of young workers aged 24 to 35 in 2005 who switched employer in-between the ten years' time frame, by gender, 2005–2015.

	Gender		
	Female	Male	Differences
Switched employer between 2005 and 2015	67.15%	68.61%	-1.46%***
Switched employer between 2005 and 2010	55.72%	56.75%	-1.03%***
Switched employer between 2010 and 2015	43.26%	45.26%	-2.00%***

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.4: Weekly wage growth of young workers aged 24 to 35 in 2005 by the changed employer and gender, 2005-2015.

	Changed employer		
	Female stayed with the same employer	Female switched employer	Differences by column
Weekly wage annual growth rate	5.15%	5.88%	-0.73%***
	Male stayed with the same employer	Male switched employer	
Weekly wage annual growth rate	5.23%	6.06%	-0.83%***
Differences by row	-0.08%	-0.18%**	

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.5: Percentage of young workers aged 24 to 35 in 2005 who changed occupation between 2005 and 2015, by gender.

	Gender		
	Female	Male	Difference
Changed occupation between 2005 and 2015	49.32%	48.47%	0.85%**

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.6: Weekly wage growth of young workers aged 24 to 35 in 2005 by changed occupation and gender, 2005-2015.

	Changed occupation		
	Female stayed within the same occupation	Female changed occupation	Differences by column
Weekly wage annual growth rate	5.91%	5.40%	0.51%***
	Male stayed within the same occupation	Male changed occupation	
Weekly wage annual growth rate	6.02%	5.54%	0.48%***
Differences by row	-0.11%	-0.14%**	

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.7: Percentage of young workers aged 24 to 35 in 2005 who stayed with the same employer and changed occupation between 2005 and 2015, by gender.

	Gender		
	Female	Male	Differences
Stayed with the same employer and changed occupation between 2005 and 2015	66.64%	57.45%	9.19%***

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.8: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 who stayed with the same employer by changed occupation and gender, 2005-2015.

	Changed occupation		
	Female stayed with the same employer and stayed within the same occupation	Female stayed with the same employer and changed occupation	Differences by column
Weekly wage annual growth rate	5.36%	5.05%	0.31%***
	Male stayed with the same employer and stayed within the same occupation	Male stayed with the same employer and changed occupation	
Weekly wage annual growth rate	5.41%	5.09%	0.32%***
Differences by row	-0.05%	-0.04%	

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.9: Percentage of young workers aged 24 to 35 in 2005 who switched employer and also changed occupation between 2005 and 2015, by gender.

	Gender		
	Female	Male	Differences
Switched employer and also changed occupation between 2005 and 2015	41.91%	44.58%	-2.67%***

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.10: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 who switched employers by changed occupation and gender, 2005-2015.

	Changed occupation		
	Female switched employer and stayed within the same occupation	Female switched employer and changed occupation	Differences by column
Weekly wage annual growth rate	6.08%	5.66%	0.42%***
	Male switched employer and stayed within the same occupation	Male switched employer and changed occupation	
Weekly wage annual growth rate	6.27%	5.83%	0.44%***
Differences by row	-0.19%	-0.17%	

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.11: The total number (instances) of permanent job separation of young workers aged 24 to 35 in 2005 by reasons of separation and gender, 2005-2015.

	The total number (instances) of permanent job separation by gender		
	Female	Male	Differences
Reason for permanent job separation			
Layoff	7.75%	12.92%	-5.17%***
Quit	15.60%	17.93%	-2.33%***
Parental or maternity leave	2.84%	0.19%	2.65%***
Other	7.48%	8.90%	-1.42%***
No reason	66.33%	60.06%	6.27%***

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%. The reason other includes separation from a business enterprise occurred during the year for ROE labour dispute, retirement, work-sharing program, other, parental leave, or compassionate care.

Table C.12: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 by reasons of separation and gender, 2005-2015.

	Weekly wage annual growth rate by gender		
	Female	Male	Differences
Reason for permanent job separation			
Layoff	5.15%	5.64%	-0.49%***
Quit	5.08%	5.88%	-0.80%***
Parental or maternity leave	3.74%	4.16%	-0.42%***
Other	4.85%	5.35%	-0.50%***
Overall wage growth	5.11%	5.72%	-0.61%***

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%. The reason other includes separation from a business enterprise occurred during the year for ROE labour dispute, retirement, work-sharing program, other, parental leave, or compassionate care.

Table C.13: The total number (instances) of temporary job separation of young workers aged 24 to 35 in 2005 by reasons of separation and gender, 2005-2015.

	The total number (instances) of temporary job separation by gender		
	Female	Male	Differences
Reason for temporary job separation			
Layoff	23.17%	32.30%	-9.13%***
Quit	4.52%	4.35%	0.17%
Parental or maternity leave	22.15%	2.42%	19.73%***
Other	11.45%	21.33%	-9.88%***
No reason	38.71%	39.60%	-0.89%*

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%. The reason other includes separation from a business enterprise occurred during the year for ROE labour dispute, retirement, work-sharing program, other, parental leave, or compassionate care.

Table C.14: Weekly wage annual growth rate of young workers aged 24 to 35 in 2005 by reasons of temporary separation and gender, 2005-2015.

	Weekly wage annual growth rate by gender		
	Female	Male	Differences
Reason for temporary job separation			
Layoff	5.25%	5.46%	-0.21%**
Quit	5.13%	5.53%	-0.40%**
Parental or maternity leave	5.51%	4.76%	0.75%**
Other	5.78%	5.76%	0.02%**
Overall wage growth	5.48%	5.60%	-0.12%**

Notes: Data presented in percentage. *: significant at 10%; **: significant at 5%; ***: significant at 1%. The reason other includes separation from a business enterprise occurred during the year for ROE labour dispute, retirement, work-sharing program, other, parental leave, or compassionate care.

Table C.15: Effect of labor market trajectories including type of separation, changed employer or occupation and reason for job separation on (log) weekly wage growth, conditional on observables, Equation (3.1).

	Female (1)	Male (1)	Female (2)	Male (2)	Female (3)	Male (3)
*T*ype of separation						
The total number of permanent separations	-0.002 (0.003)	-0.005** (0.002)	-0.002 (0.003)	-0.004** (0.002)	-0.004 (0.004)	-0.006** (0.003)
The total number of temporary separations	-0.002** (0.001)	-0.003*** (0.001)	-0.002** (0.001)	-0.002*** (0.001)	-0.002* (0.002)	-0.001 (0.001)
*C*hanged employer or occupation						
Switched employer between 2005 and 2015			0.010* (0.006)	0.007* (0.004)	0.011* (0.006)	0.007* (0.004)
Changed occupation between 2005 and 2015			0.002 (0.004)	-0.007* (0.004)	0.001 (0.004)	-0.007* (0.003)
*R*eason for job separation						
The total number (instances) of permanent separations due to layoff					0.002 (0.004)	0.002 (0.003)
The total number (instances) of temporary separations due to layoff					0.001 (0.002)	-0.001 (0.001)
The total number (instances) of permanent separations due to quit					0.003 (0.004)	0.002 (0.004)
The total number (instances) of temporary separations due to quit					-0.003 (0.006)	-0.004 (0.004)
The total number (instances) of permanent separations due to parental/maternity leave					-0.012 (0.008)	0.009 (0.033)
The total number (instances) of temporary separations due to parental/maternity leave					-0.004 (0.003)	-0.007 (0.006)
Control variables						
ED	✓	✓	✓	✓	✓	✓
EXP	✓	✓	✓	✓	✓	✓
FAM	✓	✓	✓	✓	✓	✓
OCCU	✓	✓	✓	✓	✓	✓
Other	✓	✓	✓	✓	✓	✓
BRR replications	500	500	500	500	500	500
adj. R^2	0.073	0.070	0.074	0.073	0.078	0.074
N	2,710	3,292	2,710	3,292	2,710	3,292

Notes: Other block contains the continuous measure of age, age square and the series of dummy variables for the province of residence derived from the 2016 Census. The number of times that an individual observed in the LWF file is also included in the block derived from the LWF. Standard errors are estimated using the bootstrap weight (BRR), which is shown in parentheses. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.16: part I – Gender difference in weekly wage annual growth rate.

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.006** (0.003)	-0.008*** (0.003)	-0.007** (0.003)	-0.007** (0.003)	-0.007** (0.004)	-0.009** (0.004)
Control variables						
ED		✓	✓	✓	✓	✓
EXP			✓	✓	✓	✓
FAM				✓	✓	✓
OCCU					✓	✓
Other						✓
BRR replications	500	500	500	500	500	500
adj. R^2	0.001	0.010	0.014	0.021	0.029	0.051
N	6,002	6,002	6,002	6,002	6,002	6,002

Notes: Other block contains the continuous measure of age, age square and the series of dummy variables for the province of residence derived from the 2016 Census. The number of times that an individual observed in the LWF file is also included in the block derived from the LWF. Standard errors are estimated using the bootstrap weight (BRR), which is shown in parentheses. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.16b: part II – Gender difference in weekly wage annual growth rate

	(7)	(8)	(9)	(10)	(11)	(12)
Female	-0.008** (0.004)	-0.008** (0.004)	-0.008** (0.004)	-0.008** (0.004)	-0.005 (0.004)	-0.005 (0.004)
Type of separation	✓	✓	✓	✓	✓	
Changed employer or occupation		✓	✓	✓	✓	
Reason for job separation (permanent + temporary)						
Layoff			✓	✓	✓	
Quit				✓	✓	
Parental/maternity					✓	✓
Control variables						
ED	✓	✓	✓	✓	✓	✓
EXP	✓	✓	✓	✓	✓	✓
FAM	✓	✓	✓	✓	✓	✓
OCCU	✓	✓	✓	✓	✓	✓
Other	✓	✓	✓	✓	✓	✓
BRR replications	500	500	500	500	500	500
adj. R^2	0.055	0.057	0.057	0.058	0.060	0.053
N	6,002	6,002	6,002	6,002	6,002	6,002

Notes: Other block contains the continuous measure of age, age square and the series of dummy variables for the province of residence derived from the 2016 Census. The number of times that an individual observed in the LWF file is also included in the block derived from the LWF. Standard errors are estimated using the bootstrap weight (BRR), which is shown in parentheses. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Table C.17: Oaxaca-Blinder decomposition of gender difference in weekly wage annual growth rate.

	(1)		(2)	
	Values	Percentage	Values	Percentage
*T*ype of separation				
The total number of permanent separations	-0.0016* (0.001)	-33.33%		
The total number of temporary separations	-0.0005 (0.000)	-16.67%		
*C*hanged employer or occupation				
Switched employer between 2005 and 2015	0.0005** (0.000)	16.67%		
Changed occupation between 2005 and 2015	-0.0000 (0.000)	0.00		
*R*eason for job separation				
The total number of permanent separations due to layoff	0.0005 (0.001)	16.67%		
The total number of temporary separations due to layoff	-0.0000 (0.001)	-0.00%		
The total number of permanent separations due to quit	-0.0000 (0.000)	-0.00%		
The total number of temporary separations due to quit	-0.0000 (0.000)	-0.00%		
The total number of permanent separations due to parental/maternity leave	-0.0015* (0.001)	-33.33%	-0.0016* (0.001)	-50.00%
The total number of temporary separations due to parental/maternity leave	-0.0006 (0.001)	-16.67%	-0.0005 (0.001)	-16.67%
Control variables				
ED	✓		✓	
EXP	✓		✓	
FAM	✓		✓	
OCCU	✓		✓	
Other	✓		✓	
BRR replications	500		500	
adj. R^2	0.055		0.046	
N	6,002		6,002	

Notes: Other block contains the continuous measure of age, age square and the series of dummy variables for the province of residence derived from the 2016 Census. The number of times that an individual observed in the LWF file is also included in the block derived from the LWF. Standard errors are estimated using the bootstrap weight (BRR), which is shown in parentheses. *: significant at 10%; **: significant at 5%; ***: significant at 1%.

Appendix D

D.1 Data Flow Chart

All number here in the chart are count of unique individuals

Census 2006

- n = 6,842,339, weighted counts = 31,612,897
- n = 640,285 young workers (aged 25-34 and employed in 2005), weighted counts = 3,191,563

Census 2016

- n = 8,651,677, weighted counts = 34,460,064

2006 LWF & Census linkage key file (2019 vintage)

2016 LWF & Census linkage key file (2019 vintage)

Census 2006 linked to LWF

- remain = 3,782,993 (55.29%) of unweighted counts of 2006 Census
- lost = 3,059,346 (44.71%) of unweighted counts of 2006 Census

Census 2016 linked to LWF

- remain = 6,751,009 (78.03%) unweighted counts of 2016 Census
- lost = 1,900,668 (21.96%) of unweighted counts of 2016 Census

Census 2006 linked to Census 2016 by LWF's longitudinal key casenum2019
remain = 840,216 (12.28%) of unweighted counts of 2006 Census
lost total = 8,853,570 (which 2,942,777 are from the 2006 Census, and 5,910,793 are from the 2016 Census)

Census 2006 linked to Census 2016 by LWF's longitudinal key casenum2019
pre additional restrictions: n = 840,216 (12.28%) of unweighted counts of 2006 Census
Using 2006 Census weights on the 840,216 = 4,012,579/31,612,897 (2006 Census long-form weighted counts of individuals) = 12.69% (only a proxy)
Using 2016 Census weights on the 840,216 = 3,234,153/34,460,064 (2016 Census long-form weighted counts of individuals) = 9.39% (only a proxy)

Continued

Continued

Merge to LWF T1+T4 file
keep if rectype_p = 1 (keep to cases where both T1 and T4/T4E record exists & keep if earnflag = 1 (keep to main job per person per year)
lost = 150,462, including unmatched cases between T1 and T4 merge285 young workers (aged 25-34 and employed in 2005), weighted counts = 3,191,563

There are some unmatched cases between these two conditions when merging T1 and T4:
The reason that the two restrictions when put together have some nomatches is because rectype_p = 1 refers to situation when there is matches between T1 and T4 and/or T4E. The no matches that we see is for records that can only be linked with T1 and T4E, but no T4 slip

Census 2006 linked to Census 2016 by LWF's longitudinal key casenum2019, after applying further restriction using rectype_p from the LWF person file and earnflag from the LWF job file
remain = 689,754 (10.08%) of unweighted counts of 2006 Census, 2.18% (only a proxy) of the weighted population of 2006 Census
lost = 150,462, (17.91%) of the 840,216 remaing individuals from above, which is 0.48% (only a proxy) of the weighted population of 2006 Census

Merge to Census — LWF weight file, note the Census — LWF linkage weight only applies for population aged 16 and above

Final analytical file (689,754)
Apply age restriction (age 25-34) and employment status (employed in 2005, lftag≥1 & lftag≤4):
Remain= **110,430**, using the Census — TAX admin weight, the weighted counts is:
713,880/31,612,897 (2006 Census weighted population) = 2.26%
713,880/3,191,563 (2006 Census weighted young workers population) = 22.37%
Keep to individuals with valid values for control variables in the regression analysis,
Remain= **6,002**, using the Census — TAX admin weight, the weighted counts is:
35,819/31,612,897 (2006 Census weighted population) = 0.113%
35,819/3,191,563 (2006 Census weighted young workers population) = 1.12%

D.2 Technical Annex

D.2.1 Derivation of the dependent variable

The weekly wage annual growth rate between 2005 and 2015 is defined as (ln(weekly wage 2015)-ln(weekly wage 2005)/10). This would give us ln (1+r (real weekly wage growth)), which is approximately r (when r is close to zero). Outliers of weekly wages larger than $1,000,000 are dropped from the calculation. Note that one of the reasons here that log weekly wage is not chosen as the main outcome variable in this study is due to the fact that using labor market trajectories measures which derived over ten years from the LWF on one-time wage differences observed in 2015 is subjective to bias. An average weekly wage annual growth is more representative of the outcome for the change in labor market trajectories for over ten years.

D.2.2 Definition of regressor variables, by block

ED:

ED block contains the series of dummy variables derived from the highest level of education from the 2016 Census. The original variable contains the following categories: no high school certificate or equivalency certificate without further schooling; no high school certificate or equivalency certificate with a registered apprenticeship or other trade certificate; no high school certificate or equivalency certificate with college, CEGEP or other non-university certificate; with high school certificate or equivalency certificate without further schooling; with high school certificate or equivalency certificate with a registered apprenticeship or other trade certificate; with high school certificate or equivalency certificate with college, CEGEP or other non-university certificate; with high school certificate or equivalency certificate with certificate below bachelor; with high school certificate or equivalency certificate with bachelor's degree; with high school certificate or equivalency certificate with certificate above bachelor; with high school certificate or equivalency certificate with degree in medicine, dentistry, veterinary medicine or optometry; with high school certificate or equivalency certificate with master's degree; with high school certificate or equivalency certificate with earned doctorate degree and not applicable (Institutional residents), which is recoded to missing values. The reference group used in the regression analysis is no high school certificate or equivalency certificate without further schooling. Note that in some cases, the highest level of education between the 2006 Census and 2016 Census decreased, possibly caused by collection error. Therefore, for the regression analysis, only the 2016 Census education variable is used.

EXP:

EXP block contains the continuous measurement of the total length of work experience by year and the square of work experience, derived from the year first worked from the LWF (first_year_worked). Outliers on experience caused due to data collection of the Census are adjusted. For example, for an individual 44 years old in 2015, the max exp should be around 30, and for an individual 25 years old in 2005, the max exp should be around 15.

FAM:

FAM block contains the series dummy variables for marital status and parent flag derived from the 2016 Census. Series of dummy variables for the presence of children by age group is also included in the block but is only available from the 2006 Census. The original variable of marital status (marst) contains the following categories: never married (including living common law); married; separated (including living common law); divorced (including living common law), and widowed (including living common law). The reference group selected for the regression analysis is never married (including living common law). The parent flag (parent) contains two categories: false - This person is not the parent of anyone else in the household; true - This person is the parent of another member of the household and not processed for family characteristics which are recoded to missing values. The reference group selected for the regression analysis is false - This person is not the parent of anyone else in the household. The original presence of children by age group variable (present_child) contains the following categories: no children present; at least one child < 2, none > 5, (possibly children 2-5 also); at least one child < 2, some > 5 (possibly children 2-5 also); none < 2, at least one child 2-5, none > 5; None < 2, at least one child 2-5, some > 5; none < 6, at least one child 6-14 (possibly children > 14 also); none < 15, at least one child 15-24 (possibly children > 24 also), none $<$ 25, at least one child 25 or more and lastly not applicable, which is recoded to a missing value. The reference group selected for the regression analysis is no children present.

OCCU:

OCCU block contains the series of dummy variables for occupation groups derived from the 2016 Census. The pension status flag and union status flag are also included in the block derived from the LWF. The original occupation variable (noc16brd), which are grouped based on the first digit of North American Occupation Classification Code (NOC) 2016, contains the following categories: management occupation; business, finance and administration occupation; natural and applied science and related occupation; health occupations; occupations in education, law and social, community and government services; occupations in art, culture, recreation and sport; sales and service occupation; trades, transport and equipment operators and related occupations; natural resources, agriculture and related production occupations; Occupations in manufacturing and utilities; did not work in 2015 and 2016 and not applicable, < 15 years in which the last two categories are recoded to missing values. The reference group selected for the regression analysis is management occupation. The pension status flag is derived from the yearly pension contribution (empe_padj_amt), which is assigned a value of 1 if individuals in the LWF made a pension contribution to a registered pension plan or deferred profit-sharing plan anytime between 2005 and 2015. Similarly, the union status flag is derived from the yearly union fees (union_dues_dues), which are assigned a value of 1 if individuals in the LWF had union fees dues any time between 2005 and 2015.

Other:

Other block contains the continuous measure of age, age square and the series of dummy variables for the province of residence derived from the 2016 Census. The number of times that an individual observed in the LWF file is also included in the block derived from the LWF. Similar to the case of education, there are some cases where individuals' age from the

2006 Census decreased in the 2016 Census, possibly due to collection error. Therefore, the age variable (age) from the 2006 Census is used (which is the age used to select young workers) to derive the age of individuals in 2016 instead of using the self-reported age in the 2016 Census. The original province of residence variable (pr) contains the following categories: Newfoundland and Labrador; Prince Edward Island; Nova Scotia; New Brunswick; Quebec; Ontario; Manitoba; Saskatchewan; Alberta; British Columbia; Yukon; Northwest Territories; Nunavut and not applicable, which is recoded to a missing value. The reference group selected for the regression analysis is Newfoundland and Labrador (default by Stata command using "xi:").

TP:

TP block contains the continuous measurement of the type, including the total number of separations, the total number of permanent separations and the total number of temporary separations. The LWF allows a distinction between permanent and temporary separations. In the LWF, a temporary separation is identified as an employee returning to his or her employer during the year of the separation or in the following year. When such a return does not occur, the separation is identified as permanent separation instead. More specifically, the total number of separations is derived from the individuals' main job firm ID, including both temporary and permanent separation on a year-to-year basis. The total number of permanent separations is derived from the LWF permanent reason variables by counting the total number of permanent job separations reported due to these reasons. Similarly, the total number of temporary separations is derived from the LWF temporary reason variables by counting the total number of temporary job separations reported due to these reasons.

CEO:

CEO block contains the two binary variables on switched employers between 2005 and 2015 and changed occupation between 2005 and 2015. Both variables are derived from the 2006 and 2016 Census. The switched employer variable is assigned a value of 1 if an individual's main job firm ID is different between 2005 and 2015 and 0 otherwise. The changed occupation variable is derived from the occupation group variable (noc16brd) between 2005 and 2015, where the variable is assigned a value of 1 if the occupation group is different between two Census and 0 otherwise.

RS:

RS block contains the continuous measurement of the total number of job separations by reason. The LWF distinguishes between various types of employee separations, allowing distinctions to be made between layoffs, quits and other separations. This is done through the ROE, which specifies the reason for the work interruption or separation. The list of variables in the block includes the total number of separations due to layoff, the total number of separations due to quit, the total number of separations due to parental/maternity leave and the total number of separations due to other reasons. All of these variables are derived separately by type of job separation (permanent and temporary). More specifically, each the total number of job separation by reason variable is derived from the LWF separation reason variables, combined with the detailed ROE separation reason variable which include all reasons for separation as a separate category listed in ROE. The total number of separations is then calculated by counting the total number of separations reported due to these reasons.

Table D.1: Description of the sample population for the descriptive statistics analysis

Variable	Women	Men	Overall
Control variables			
ED			
No high school certificate or equivalency certificate without further schooling	4.27%	8.00%	6.16%
No high school certificate or equivalency certificate with registered apprenticeship or other trade certificate	0.63%	1.97%	1.31%
No high school certificate or equivalency certificate with college, CEGEP or other non-university certificate	5.55%	0.52%	0.59%
With high school certificate or equivalency certificate without further schooling	22.17%	22.72%	19.72%
With high school certificate or equivalency certificate with registered apprenticeship or other trade certificate	5.95%	13.13%	9.60%
With high school certificate or equivalency certificate with college, CEGEP or other non-university certificate	29.11%	22.77%	25.89%
With high school certificate or equivalency certificate with certificate below bachelor	3.56%	2.41%	2.97%
With high school certificate or equivalency certificate with bachelor's degree	27.02%	19.35%	23.12%
With high school certificate or equivalency certificate with certificate above bachelor	2.80%	1.68%	2.23%
With high school certificate or equivalency certificate with degree in medicine, dentistry, veterinary medicine or optometry	0.79%	0.52%	0.65%
With high school certificate or equivalency certificate with master's degree	7.67%	5.95%	6.80%
With high school certificate or equivalency certificate with earned doctorate degree	0.93%	0.99%	0.96%
EXP			
Total years of work experience	11.27	10.92	11.10
Total years of work experience square	132.25	124.90	128.64
FAM			
Never married (including living common law)	32.44%	34.38%	33.43%
Married	54.79%	56.50%	55.66%
Separated (including living common law)	4.57%	3.43%	3.99%
Divorced (including living common law)	7.72%	5.50%	6.59%
Widowed (Including living common law)	0.48%	0.20%	0.34%
This person is the parent of another member of the household	74.63%	66.70%	70.60%
No children present	55.23%	64.67%	60.03%
None < 25, at least one child 25 or more	0.01%	0.02%	0.01%
At least one child < 2, none > 5, (possibly children 2-5 also)	12.48%	14.75%	13.63%
At least one child < 2, some > 5 (possibly children 2-5 also)	2.77%	2.54%	2.65%
None < 15, at least one child 15-24 (possibly children > 24 also)	0.33%	0.24%	0.29%
None < 2, at least one child 2-5, none > 5	10.37%	7.80%	9.06%
None < 2, at least one child 2-5, some > 5	8.1%	5.19%	6.62%

Continued on next page

Variable	Women	Men	Overall
None < 6, at least one child 6-14 (possibly children > 14 also)	10.72%	4.79%	7.70%
OCCU			
Management occupations	10.66%	16.69%	13.73%
Business, finance and administration occupation	23.28%	9.63%	16.34%
Natural and applied science and related occupation	3.92%	12.79%	8.43%
Health occupations	12.10%	2.44%	7.19%
Occupations in education, law and social, community and government services	20.64%	9.18%	14.81%
Occupations in art, culture, recreation and sport	3.16%	2.43%	2.79%
Sales and service occupation	15.07%	11.66%	13.33%
Trades, transport and equipment operators and related occupations	1.55%	23.89%	12.91%
Natural resources, agriculture and related production occupations	0.47%	2.55%	1.53%
Occupations in manufacturing and utilities	1.80%	5.40%	3.63%
Had union fees dues	54.20%	48.45%	51.27%
Had pension contribution	63.37%	59.23%	61.27%
Other			
Age in 2015	39.54	39.60	39.57
Age in 2015 square	1571.21	1576.57	1573.94
The total number of times individuals observed in LWF	9.41	9.41	9.41
Newfound land and Labrador	1.52%	1.45%	1.49%
Prince Edward Island	0.42%	0.35%	0.39%
Nova Scotia	2.70%	2.36%	2.53%
New Brunswick	2.39%	2.15%	2.27%
Quebec	24.56%	24.21%	24.38%
Ontario	37.99%	36.94%	37.46%
Manitoba	3.27%	3.44%	3.36%
Saskatchewan	3.04%	3.04%	3.04%
Alberta	11.52%	13.10%	12.33%
British Columbia	12.01%	12.44%	12.23%
Yukon	0.19%	0.16%	0.17%
Northwest Territories	0.23%	0.18%	0.20%
Nunavut	0.16%	0.17%	0.16%
N	55,131	55,299	110,430

Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.

Table D.2: Description of the sample population for the regression analysis

Variable	Women	Men	Overall
Control variables			
ED			
No high school certificate or equivalency certificate without further schooling	4.19%	12.54%	8.82%
No high school certificate or equivalency certificate with registered apprenticeship or other trade certificate	0.97%	5.18%	3.30%

Continued on next page

Variable	Women	Men	Overall
No high school certificate or equivalency certificate with college, CEGEP or other non-university certificate	0.51%	0.61%	0.57%
With high school certificate or equivalency certificate without further schooling	17.05%	22.74%	20.2%
With high school certificate or equivalency certificate with registered apprenticeship or other trade certificate	7.87%	24.88%	17.29%
With high school certificate or equivalency certificate with college, CEGEP or other non-university certificate	34.11%	18.81%	25.64%
With high school certificate or equivalency certificate with certificate below bachelor	3.68%	1.64%	2.55%
With high school certificate or equivalency certificate with bachelor's degree	22.39%	9.34%	15.16%
With high school certificate or equivalency certificate with certificate above bachelor	2.03%	0.58%	1.23%
With high school certificate or equivalency certificate with degree in medicine, dentistry, veterinary medicine or optometry	0.06%	0.04%	0.05%
With high school certificate or equivalency certificate with master's degree	6.59%	3.28%	4.76%
With high school certificate or equivalency certificate with earned doctorate degree	0.54%	0.34%	0.43%
EXP			
Total years of work experience	11.18	10.77	10.95
Total years of work experience square	128.26	119.76	123.56
FAM			
Never married (including living common law)	38.93%	49.26%	44.65%
Married	46.61%	41.62%	43.85%
Separated (including living common law)	6.05%	4.02%	4.92%
Divorced (including living common law)	7.92%	4.89%	6.24%
Widowed (Including living common law)	0.49%	0.22%	0.34%
This person is the parent of another member of the household	76.87%	63.82%	69.64%
No children present	49.98%	63.88 %	57.67%
None < 25, at least one child 25 or more	0.01%	0.00%	0.00%
At least one child < 2, none > 5, (possibly children 2-5 also)	15.82%	13.78%	14.69%
At least one child < 2, some > 5 (possibly children 2-5 also)	3.91%	2.58%	3.17%
None < 15, at least one child 15-24 (possibly children > 24 also)	0.52%	0.4%	0.45%
None < 2, at least one child 2-5, none > 5	9.15%	7.93%	8.47%
None < 2, at least one child 2-5, some > 5	6.92%	5.21%	5.97%
None < 6, at least one child 6-14 (possibly children > 14 also)	13.70%	6.22%	9.56%
OCCU			
Management occupations	7.32%	6.00%	6.59%
Business, finance and administration occupation	19.79%	6.01%	12.16%
Natural and applied science and related occupation	2.88%	6.42%	4.84%
Health occupations	6.15%	0.78%	3.18%
Occupations in education, law and social, community and government services	35.74%	7.07%	19.86%
Occupations in art, culture, recreation and sport	3.23%	2.57%	2.86%
Sales and service occupation	15.99%	10.26%	12.82%
Trades, transport and equipment operators and related occupations	4.18%	46.03%	27.36%

Continued on next page

Variable	Women	Men	Overall
Natural resources, agriculture and related production occupations	1.30%	6.92%	4.41%
Occupations in manufacturing and utilities	3.42%	7.94%	5.92%
Had union fees dues	68.17%	68.06%	68.11%
Had pension contribution	69.49%	63.49%	66.17%
Other			
Age in 2015	39.27	39.12	39.19
Age in 2015 square	1551.20	1538.64	1544.24
The total number of times individuals observed in LWF	9.96	9.79	9.87
Newfound land and Labrador	2.43%	2.60%	2.52%
Prince Edward Island	1.16%	0.68%	0.89%
Nova Scotia	3.53%	3.04%	3.26%
New Brunswick	3.03%	3.61%	3.35%
Quebec	32.58%	38.70%	35.97%
Ontario	31.39%	25.50%	28.13%
Manitoba	3.42%	2.65%	2.99%
Saskatchewan	2.85%	2.65%	2.74%
Alberta	8.70%	10.92%	9.93%
British Columbia	10.31%	9.43%	9.83%
Yukon	0.14%	0.13%	0.13%
Northwest Territories	0.32%	0.10	0.14%
Nunavut	0.14%	2.60%	0.12%
N	2,710	3,292	6,002

Source: Statistics Canada, 2019 Longitudinal Worker File linked to 2006 and 2016 census long-form data.