Three Essays on Wage Inequality: Evidence for the Role of Monopoly Power, Average City Rent, Regional Growth Cluster, and Interprovincial Migration in Canada (1995-2015)

By

Sadjad Kalhor

A thesis submitted to
the Faculty of Graduate Studies
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Economics
Faculty of Arts
University of Manitoba
Winnipeg, Manitoba

July 2020

Abstract

This dissertation is an empirical analysis investigating various determinants of real wage inequality in Canada. It includes a general literature review and three empirical studies.

The first two empirical essays aim to explore the link between wage inequality and industrial monopoly power, average city rent (ACR), and regional growth clusters (RGCs) within industries, as well as within Census Metropolitan Areas (CMAs). Canadian Census Microdata (1996, 2001, 2006, 2011, and 2016) will be used as the primary dataset in these essays. In the third essay, the effect of interprovincial migration on national wage inequality is examined using the 2016 Census PUMF.

A two-round empirical analysis is conducted to examine the effect of monopoly power (measured by Lerner's index derived from multifactor productivity dataset) on within-industry wage inequalities in Chapter 3. In the first round, wage inequalities (captured by Theil's index) are computed for two-digit industries (sectors) to uncover variations of wage inequality within industries. Within-industry T-values (Theil's index), then, will be used as a dependent variable in the second round in a pooled OLS framework. The same approach as Chapter 3 (i.e. two-round analysis) has been taken in Chapter 4 for CMAs instead of industries to disentangle the effect of ACR (an index representing high/low-pay industry composition of a CMA) and RGCs (representing major industrial clusters of a CMA) on within-CMA wage inequalities. To uncover the effect of interprovincial migration on inequality, a semiparametric approach is considered in which counterfactual wage densities and inequalities are estimated in the absence of internal migration.

The results show that a higher monopoly power is associated with a reduction in wage inequality within industries by 0.19%. With respect to within-CMA inequality, the estimated results exhibit that there is a significant relationship between wage inequality and average city rent. More importantly, the growth rate of RGCs also matter to inequality across CMAs. It is estimated that a faster rate of growth of RGCs tends to increase CMA wage (earnings) inequality by 0.51% in Canada. And, finally, the estimated counterfactual Theil's values indicate that interprovincial migration substantially reduces wage inequality in Canada. Besides, estimated wage densities show that migration exerts large and differing impacts on the lower portion of the wage distribution, whereas the effect sharply reduces and fades away in the upper portion of the wage distribution.

Keywords: Wage Inequality, Industrial Inequality, Spatial Inequality, CMA, Inequality Decomposition, Interprovincial Migration.

Acknowledgments

I would like to offer my sincere gratitude to my supervisor, Dr. Wayne Simpson, who has supported me throughout this research with his patience and knowledge. Without his encouragement and insightful thought, this work would not have been completed. I am also deeply grateful to my committee member Dr. Umut Oguzoglu for his constructive guidance and solutions he always provided. I also would like to extend my gratitude to my other committee member, Dr. Royce Koop, for insightful feedback he has offered.

I would also like to thank the Manitoba Research Data Center for providing facilities in accessing micro datasets that have been used in this project.

Dedication

It is an honour to dedicate this thesis to my lovely family and beautiful companion, Ronak, for their never-ending support throughout this journey.

Contents

List of Tables

List of Figures

Chapter 1

Overall Introduction

The subject of wage inequality has attracted a great deal of attention among economists and policymakers due to a rapid rise in inequality that began in the 1980s. While the Classical economists cheered for inequality as a source of economic development, the modern view is concerned that it is harmful to the economy (Galor, 2009). Such views can be best summarized, according to Cingano (2014), in the following manner;

- Inequality can be considered to be good because greater inequality is a reward to more talented and productive individuals and enterprises to take higher risks. Hence, it results in more innovation and efficiency. Additionally, inequality boosts capital accumulation through larger investment by the rich, since they have a lower propensity to consume. Moreover, redistribution of income causes a deadweight loss from taxation as resources have to be spent on bureaucracy and administration.
- On the other hand, inequality can be harmful to an economy for both social and economic reasons. From a social perspective, if the voters don't accept higher inequality, then they will request higher tax rates and regulations, or lose their

> trust in business and government, and in extreme cases, social unrest might take place. Furthermore, from an economic point of view, if human capital accumulation depends on income, then the poor would have less opportunity to become educated and develop their skills. This effect, in turn, harms economic growth through declining human capital.

Given the importance of wage inequality in the modern economy, this project is an empirical exercise to investigate annual real wage inequality fairly comprehensively for the case of Canada using multiple data sources, including Census 2001, 2006, 2011, and 2016[1]. When exploring the literature, a great number of insightful studies are available that have documented both theoretical and empirical aspects of inequality (e.g. Acemoglu (1999); Borjas and Ramey (1995); Card (2009); Kremer and Maskin (1996)).

However, to the best of my knowledge, an investigation of Canadian wage inequality across and within industries has not been reported in the literature. In addition, when considering inequality across locations (e.g. across provinces or metro areas), only few studies can be found for Canada. Moreover, it seems that no study has been conducted to quantify the effect of interprovincial migration on Canadian wage inequality. Therefore this exercise is an attempt to contribute to the literature by addressing the above-mentioned gaps in which a general literature review is discussed, followed by three empirical essays.

---

[1] Please note that only observations for the year prior to Census reference week is included in the statistical analysis (i.e. 1995, 2000, 2005, 2010 and 2015).

The first essay (Chapter 3) is a two-round empirical analysis that concerns wage inequality when industries are the unit of study. In the first round, the goal is to explore the pattern of wage inequality among workers (employed and self-employed, full-time and part-time) within and between Canadian industries. And the second round aims to explore the determinants of inequality within these industries. Thus, a pooled OLS regression is implemented in which the within-industry Theil's values from first-round serve as a dependent variable in the second stage. Literature suggests that industries with larger market concentration (monopoly power) pay a premium (Gera and Grenier, 1994), but it is not clear to what extent it affects wage inequality (e.g. Theil's index) in particular within the industry itself (i.e. among the workers of the sector). To address this challenge, first, sectoral monopoly power is estimated using a creative and insightful methodology introduced by Hall (2018) that employs aggregate data (multifactor productivity dataset) to extract Lerner's values for each industrial group to proxy for monopoly power. Next, the estimated Lerner's values will be included in the regression analysis to find out whether it has any impact on wage inequality within industries.

Regarding the spatial aspect of inequality, the effect of several determinants have been documented in the literature, but no study has specifically measured the effect of average city rent (ACR) and regional growth clusters (RGCs). ACR is an index introduced by Beaudry et al. (2012) that captures the composition of high-pay and low-pay industries in an economy. In their framework, they show that a larger composition of high-pay industries would boost the wages in other industries through general equilibrium. RGCs are also another element to consider when studying wage inequality at the CMA level as it is found to be an important determinant of wage inequality within CMAs. Identifying RGCs requires a statistical test that will be explained in Chapter 4. This variable is supposed to control for the impact of economic growth at the regional level on urban wage inequality. Therefore, the second essay is intended to investigate the effect of these factors, for the first time, as a potential determinant of inequality within CMAs.

In line with the first two essays, the third area of study concerns the effect of interprovincial migration on wage inequality, which has not been addressed in Canadian literature. This discussion aims to quantify the effect of interprovincial migration on wage distribution (using wage densities) and wage inequality (measured by Theil's values) across Canadian provinces. Equally important, another goal is to discover the implications that interprovincial migration has for specific skill groups (skilled vs unskilled proxied by education) with respect to wage destines and inequalities. To achieve these goals, a counterfactual framework is used that was developed by DiNardo and Lemieux (1996). In this method, wage inequality and densities are estimated in the absence of migration to obtain estimations of counterfactual wage densities and Theil's values. With the estimated output in hand, a comparison between observed wage densities (and inequality) with estimated wage densities (and inequality) is made to extract the effect of migration on wage structure in Canada and across skill groups.

Chapter 2

General Literature Review

The subject of income inequality or wage inequality (as a component of income) began to receive a great deal of attention among practitioners when rising inequality was detected in the 1980s. Since then, a great number of studies have attempted to explore the pattern of inequality along certain portions of income distribution as well as the causes of wage inequality among workers. What follows provides a general overview of selected papers that have shaped our general understanding of income (or wage) inequality in the U.S. or in Canada within the last four decades.

2.1 Past Studies of Income Distribution

Figure 2-1 illustrates the average Gini coefficient for OECD countries for mid 1980s and 2012. It demonstrates that rising inequality became almost pervasive among OECD countries as the OECD-average of Gini coefficient increased by 10% from 0.29 in the mid-1980s to 0.316 by the late 2000s. (OECD, 2011).

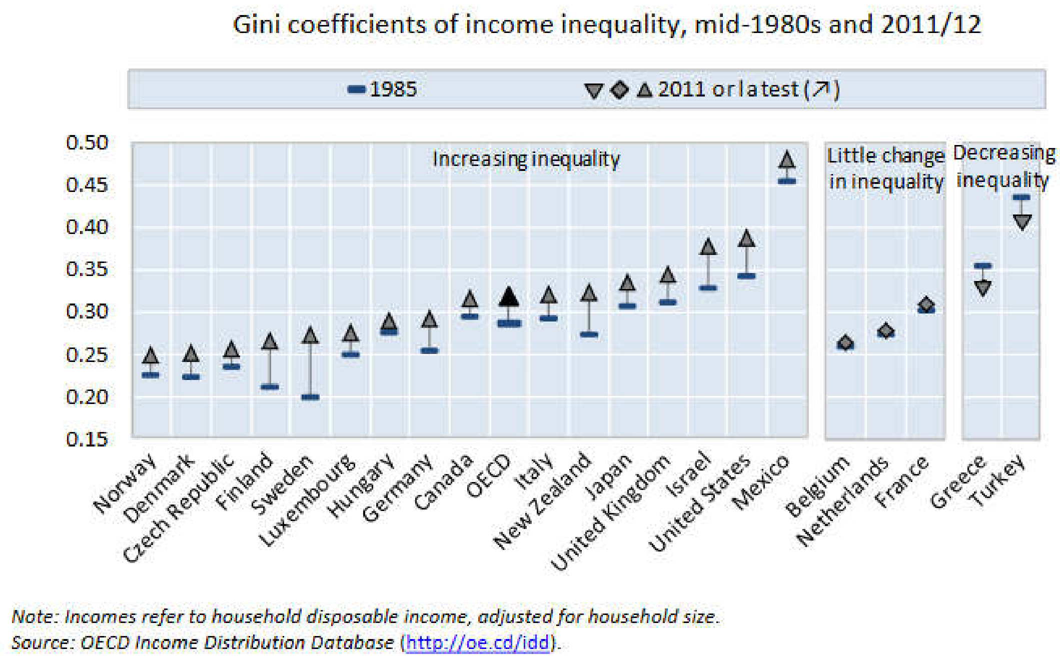

Figure 2-1. Inequality increased in most OECD countries

Focusing on Canadian inequality, Saez and Veall (2005) provide a comparative and historical analysis of the pattern of inequality for Canada and the U.S. They find that the pattern of inequality in Canada is very similar to that in the US, especially at the top of the income distribution. They show that both the U.S. and Canada experienced a sharp drop in the top 0.1% income share during World War II with no recovery before the 1970s. Since the 1980s, however, the top group reached its prewar levels.

Complementary findings are also documented by Yalnizyan (2011). She carries out a historical and statistical examination of Canadian income inequality from 1920 to 2010 and compares Canada today with the Gilded Age a century ago. She shows that Canada's top 1% accounted for 32% of income growth between 1997 and 2007, a period marked by moderate economic growth. However, the share of Canada's top 1% in the 1950s and 60s (a period with the fastest economic growth ever recorded) was only 8%.

2.2 Determinants of Inequality

A large number of factors can be found in the literature that have been examined as determinants of inequality or rising inequality since the 1980s. As Lemieux (2006) points out, one very well-known hypothesis in this field is Skilled-Biased Technological Change (SBTC), proposed in a series of studies during the 1990s (e.g. Berman et al. (1994); Juhn et al. (1993); Katz and Murphy (1992)). However, later, Card and DiNardo (2002) cast serious doubt on this widely-accepted view and conclude that while SBTC is somewhat able to explain the rising inequality during the 1980s, it fails to account for the much smaller change in the 1990's inequality despite the strong supporting evidence of technological advancements during that period.

In contrast with the SBTC hypothesis, which recognizes the demand side of the labour market as the determinant of rising inequality, a supply view has been proposed as well. For instance, Kremer and Maskin (1996) express that growth in the dispersion and the mean of skill levels is the explanation for rising inequality. Acemoglu (1999, 2002) also takes a supply approach and argues that technological change is an endogenous response to a rise in the relative supply of skilled and unskilled workers.

In addition to the SBTC hypothesis and the skill-supply view of inequality, earnings dispersion has been scrutinized from other aspects. Ongoing research includes the impact of international trade on national wage inequality. The idea here is that the trade of developed economies, which are relatively skill-abundant, with emerging economies, which are relatively scarce in skill, drives up the wage gap between skilled and unskilled workers within the skill-abundant country, thereby raising inequality (Harrison et al., 2011). This view can also be detected in Krugman (2008), who thinks of the changes in the pattern of trade since the early 1990s as an important source of rising inequality in the U.S. and other advanced economies.

Another possible factor that can influence the magnitude of inequality is gender pay discrimination documented in the literature (e.g. Blau and Kahn (2000) and Moyser (2017)). This challenge seems to be a more serious issue in the past as the female-male wage ratio was lowest around 60% in the early 1970s while as high as 88% in 2014 (Moyser, 2017).

There are also studies that have measured the effect of education of inequality. For instance, Boudarbat et al. (2014) document that the education premium increased significantly from 34% to 43% for Canadian men from 1980 to 2000. The same pattern also occurred for women but with a more modest rise. Supportive evidence is also found by Lemieux (2006b) who finds that most of the increase in wage inequality from 1973 to 2005 arises from a dramatic rise in the returns to post-secondary education.

Immigration may also play a role in wag inequality, although , in practice, the impact of immigration on labour market outcomes (including wage inequality) seems to be controversial. For example, Borjas (2003) develops a supply-view framework and estimates that immigrants negatively affect native wages, while using U.S. data Card (2009) finds that immigration has had a minimal effect on native wages, and its effect on inequality is low as well.

Labour market institutions, as in minimum wages and labour unions, have also been identified as contributing factors to wage inequality (Card et al., 2004; DiNardo and Lemieux, 1996; Legree et al., 2016). In particular, while labour unions were considered to have an increasing effect on wage inequality prior to the 1980s, the modern view believes that labour unions are expected to compress the wage gaps between high-wage and low-wage workers (Card et al., 2004). Unionization, however, seems to be less relevant to wage inequality among female workers (DiNardo and Lemieux, 1996).

Finally, noting that earnings are the product of the hourly wage rate and hours worked, Johnson and Kuhn (2004) investigate the role of hours worked and wage rates in inequality for male workers in the U.S. and Canada from 1981 to 1997. They find that most of the rise in earnings inequality is attributable to inequality in wage rates rather than to hours worked in both the U.S. and Canada. However, it seems that inequality in hours worked is starting to become more important as Checchi et al. (2016) find that the dispersion of hours worked are notably large across other countries, i.e., 15% for the US, 29% for the UK, 30% for France, and 34% for Germany.

2.3 Concluding Remarks

The discussion above was presented in a fashion to illustrate the overall evolution of research in this field of study. It indicates that inequality is a wide subject, and new streams of research are emerging to broaden our understanding of wage inequality from different perspectives. A recently established literature is the study of inequality across firms or industries (e.g., Card et al. (2018)) which explores how industry differentials contribute to overall wage inequality. Chapter 3 will proceed to this matter in more details and offers an empirical analysis of wage inequality for Canadian industries from 1995 to 2015. Furthermore, Chapter 4 will discuss another growing view of wage inequality, which focuses on the spatial aspect of wage inequality (e.g., Fong (2017)). An empirical analysis will also be provided in this chapter that examines the wage inequality within and between Canadian CMAs from 1995 to 2015 using Canadian Census 1996, 2001, 2006, 2011, and 2016. And lastly, in Chapter 5, a different stream of available studies will be discussed that deal with labor mobility and wage inequality. In particular, the effect of interprovincial migration on Canadian wage inequality will be under question in this chapter.

Chapter 3

Wage Inequality for Canadian Industries: Investigating the Effect of Monopoly Power (1995-2015)

3.1 Introduction

An abundant amount of research has been conducted on the economic question of wage inequality in developed countries, including the U.S and Canada since the 1980s, when a rise in wage inequality was detected in most developing and developed countries (Pavcnik, 2011).

Early studies mainly focused on observable characteristics, in particular on skill, as the determinants of inequality (Berman et al., 1994; Bound and Johnson, 1992; Card and Freeman, 1993; Katz and Murphy, 1992). In the meantime, other studies examined wage structure from various perspectives such as productivity dispersion (Dunne et al., 2004), labour market institutions (DiNardo et al., 1996; Lee, 1999), labour market structure (Acemoglu, 1999; Borjas and Ramey, 1995; Kremer and Maskin, 1996), global trade (Breau and Rigby, 2010; Wolfson and Murphy, 1998), immigration (Borjas et al., 1997; Card, 2009) and the gender gap (Blau and Kahn, 1994).

However, little study has investigated inequality from an industrial point of view (i.e., wage inequality across and within industries). In fact, wages are generated in industries, and the way that these wages are distributed among workers should vary across and within industries over time for a variety of reasons. Some industries pay high wages, while others pay lower wages (Gera and Grenier, 1994). Some industries have strong labour institutions, while others have weaker ones (e.g., unionization rate varies significantly across industries as confirmed by data[2]). Some industries are exposed to international competition, while others are more local and less influenced by global factors (Borjas and Ramey, 1995).

Studying inequality at the industry level is actually a growing view which concerns the role of firms or industries and their characteristics in inequality. This view was neglected until recently due to a lack of comprehensive microdata-sets by which the role of firms (or industries) could be better understood in earnings inequality (Song and Price, 2015). A recent study by Card et al. (2018) provides an insightful survey of two dominant approaches in quantifying the role of firms in wage structures, and also offer a theoretical framework which connects the two views.

---

[2] Table 14-10-0132-01 Union status by industry

In particular, inspired by the SBTC hypothesis and the skill-supply view of Kremer and Maskin (1996), Acemoglu (1999, 2002), and Dunne et al. (2004) draw attention to the potential role of firms in inequality. Furman and Orszag (2015) also provide more evidence by studying inequality at the firm level in the U.S. Using various datasets, they find that a rising fraction of firms are receiving *supernormal* returns over 10, 20, or 30 percent annually with much of the return going to capital rather than to labour services.

Another insightful study is that of Borjas and Ramey (1995). In their study, they develop an oligopoly model of trade and concentration to explain the variations in the college premium in manufacturing industries that produce durable goods across U.S metropolitan areas. To test their hypothesis, the authors use a panel dataset of U.S. manufacturing industries and find that wages in such industries are higher than the national average and that the college premium is low. In other words, industries with larger concentration tend to share the rents among their skilled and unskilled workers more equally. Also, when those industries are open to international import (i.e. more competition or less market power), the college premium rises.

Proceeding to the Canadian literature, Gera and Grenier (1994) use the 1986 Labor Market Activity Survey to test the effect of various factors on wage differentials. They find that a substantial portion of wage differentials was explained by rent sharing in the form of efficiency wages. Comparing the public and the private sectors Mueller (1998) uses quantile regression to measure the government wage premium across the wage distribution. The results show that workers at the lower end of wage distribution enjoy a higher government premium, while the effect fades away as the wage moves up the distribution. Similar evidence can be found in Palacios et al., (2016) who use the 2013 Labor Force Survey to measure the government-private sector wage differential. Their results show that government workers at both the federal and provincial levels were paid 9.7% on average higher than their counterparts in the private sector but, when union status is taken into account, the wage premium declines to 6.2%.

In summary, based on the discussion provided above, it can be understood that the industrial dimension of inequality is understudied in the literature, particularly in the Canadian case. More importantly, no study has measured wage inequality within and between Canadian industrial groups. Therefore, the first objective of this study concerns the magnitude and pattern of real wage inequality by industry. A second objective is to test the effect of potential determinants of within-industry inequalities where an interesting variable (industrial Lerner's index) turns out to be a significant determinant of within industry wage inequality.

To accomplish these goals, a two-round empirical analysis is considered. In the first step, within- and between-industry Theil's values are computed using individuals' wages from Census observations (excluding observations from PEI province) to uncover the overall real wage inequality pattern by industrial groups (industries are defined based on North American Industry Classification (NAICS) 2007 at the two-digit level).

In the second round, the computed within-industry Theil's values from the first stage are regressed over a number of independent variables to quantify the determinants of inequality across industries. A novel variable in the regression analysis is Lerner's index, which attempts to measure the effect of monopoly power on the distribution of wages[3] within an industry. Interestingly, it is found that industries with larger market power tend to distribute wages more equally. I believe this finding can be linked to rent sharing theories (e.g. in the form of efficiency wages within the industries), which is also noted in Gera and Grenier (1994), where they use it as an explanation on why substantial industrial wage differentials exist among Canadian industries. They find that a great deal of Canadian industrial wage differentials are explained by rent sharing, whereas other factors such as the gender gap, skill gap, and labour institutions are weak determinants of interindustry wage differentials.

In addition to Lerner's index, the regression equation is including other control variables constructed for industries using multiple data sources. These variables are comprised of labour union rate, export exposure, import exposure, innovation in industry (measured by the share of R&D and Computer expenditures in the total input of industry), job characteristics, and demographic characteristics such as place of birth to proxy minority effects. These control variables have been found to be relevant in the inequality literature (to be discussed shortly below). The estimated coefficients turn out to be consistent with what the literature suggests. The only exceptions are a) the innovation variable whose significance is sensitive to model specification and has a negative sign, and b) the skill variable that has a consistent sign, but its significance is not robust to model specification.

---

[3] wages and earnings are used interchangeably during the discussion.

Section 3.2 proceeds to data description and construction of variables. Section 3.3 explains the procedure for estimating Lerner's index for individual industries. Theil's decomposition technique is briefly described in section 3.4. Empirical findings are provided in section 3.5. And section 3.6 concludes.

3.2 Data Description and Variables

Multiple data sources have been utilized to conduct the regression analysis in this study. Each data source serves to construct a variable or a set of variables of relevance that will be directly or indirectly included in the regression analysis. What follows provides a brief description of datasets and their contribution to the analysis.

- **Input-Output tables from Statistics Canada to construct trade variables and an innovation index:**

  Input-output data from 1997 to 2015 are used to construct trade-exposure variables, i.e. export exposure and import exposure, for NAICS industries that will enter in the regression equations separately as dummy variables.

  To calculate the export and import exposure indices for a given industry, I followed the formula provided by Statistics Canada, which computes the trade share of each commodity in the output table[4]. Once the trade share of each commodity is obtained, it would be possible to calculate a weighted average of commodity

[4] Please refer to section (vi) in https://www150.statcan.gc.ca/n1/pub/15-201-x/2010001/technote-notetech1-eng.htm

trade shares for a given industry where weight is the share of that commodity in the output of that industry.

To describe in more details, with this method, first, the import and export shares for each given commodity (in the output table rows) are calculated as shown below:

1. Import share of a commodity = (Total imports / Total domestic availability) x 100
2. Export share of a commodity = (Total exports / Total domestic availability) x 100

The domestic availability of a commodity is defined as total production less exports plus imports (assuming no inventory change). Next, for a given industry, the share of each commodity in the output of that industry is calculated to construct a weight variable. By multiplying the import share and export share of all the commodities with the corresponding weights (from the output table), I derived the weighted summation of those multiplications for that given industry resulting in new indices.

With the new indices just calculated, I used their medians (as thresholds) to generate dummy variables. Those industries whose export and import indices are above the median are assigned a value equal to one, whereas the rest of industries take on a value of zero. The industry classifications in the input-output tables are different from that of the Census. To convert the input-output industries into 20 industrial groups based on the 2007 North American Industry Classification (NA-

ICS 2007), I used the concordance table provided by the Statistics Canada website[5]. Also, input-output industries that can be converted to NAICS are available only from 1997 onwards. Considering that the regression analysis is based on Census years (1996, 2001, 2006, 2011, and 2016), 1997 import and export indices are used for 1996.

In addition to trade variables, the input table is used to derive the innovation index for a given industry from 1997 to 2015. Following the literature (Hall, 2014), using the input-definition of innovation (i.e. R&D expenditures, as well as computer usage in an industry), the innovation index is equal to those expenditures divided by the total inputs of that industry. Unfortunately, the definition of R&D in the Canadian input-output tables is not consistent over time, as it has changed since 2000, and there has been no revision to the input-output tables according to Statistics Canada. In this case, the estimated coefficients in the regression analysis need to be treated with caution.

- **Multifactor productivity accounts from Statistics Canada[6], Commodity Price Indices from the Bank of Canada[7], and Military expenditure from IndexMundi[8] for sectoral Lerner's index estimation:**

  As stated in section 3.1, market power and rent sharing have implications for inequality. As a result, Lerner's index will be needed when measuring the effect of

---

[5] https://www.statcan.gc.ca/eng/statistical-programs/document/1303_D7_T9_V1#tb1n_1
[6] https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610021701
[7] https://www.bankofcanada.ca/rates/price-indexes/bcpi/
[8] https://www.indexmundi.com/facts/canada/military-expenditure

market concentration (monopoly power) on earnings inequality within industries. The methodology for Lerner's index calculations closely mimics that of Hall (2018). In his method, there is no need for plant-level data to calculate the market power of industries. On the contrary, his method requires only aggregate data from KLEMS multifactor productivity accounts at the industry level to calculate Lerner's index for a given industry. Detailed explanations on data and methodology are provided in section 3.3, along with Lerner estimation for Canadian industries from 1985 to 2015. Unfortunately, Canadian multifactor productivity accounts are available for only 17 industrial groups, whereas there are 20 industrial groups in the Census (which is the database of my study). I will explain in more detail how these Lerner's values will be used in the regression analysis, where 20 industrial groups exist based on NAICS 2007.

- **Labor Force Survey (LFS) from Statistics Canada to construct the union share variable:**

  In the Census dataset, there is no data available to learn about the union status of a worker. As a result, I utilized the monthly survey of the labour force (LFS) from 1997 to 2015 to calculate the share of unionized workers for a given industry to match with Census years (i.e., 1995, 2000, 2005, 2010, and 2015). [9] Since LFS is a monthly survey, I used the average union rates over a 12-month period in a given year. In the survey, the union status of workers has been categorized into

[9] Note that the union status of workers was not collected for prior years, therefore 1997 union shares will be used for 1995 to match with Census years.

three groups; a) covered by union b) not in a union but covered by a collective agreement and c) no coverage. I combined categories a and b in the construction of union shares.

- **Census to construct demographic variables and wages:**

  Census 1996, 2001, 2006, 2011, and 2016 have been utilized to construct a series of demographic variables. The sample of observations included self-employed workers and employed workers who have held a part-time job or a full-time job as their main job in the year prior to Census reference week (i.e. observations for only 1995, 2000, 2005, 2010, and 2015). The Census also reports the annual wages (wages, salaries, and bonuses) of workers along with the province of respondents. Therefore, the provincial consumer price index rather than GDP deflator is used just to include more accurate values as real wages. These real wages are used for inequality decompositions (later in sections 3.4 and 3.5) where within-industry and between-industry Theil's values are generated.

  With the chosen sample, for 20 industrial categories (based on NAICS 2007 definition), I created skill shares across all the industries. Skill is measured by education degree attained (bachelor's and above) following the literature (Card and DiNardo, 2002). This variable is included to control for education effects on inequality. I also created male shares across all industries. Such a variable is meant to control for any gender effects that may exist in relation to inequality (Blau and Kahn, 2000). Two more control variables are also added to the regression as suggested by the literature: a) the Canadian-born share, measured as the share of workers who were born in Canada versus those who were born outside of Canada, a variable meant to capture any potential influence that immigration may have on inequality (Picot and Hou, 2016; Warman and Worswick, 2016), and b) the full-time ratio, measured as the share of workers who reported holding a full-time job versus those who reported holding a part-time job as their main job. A part-time worker is generally paid below a full-time worker as a result of the differences in their working hours (Simpson, 1986). Also, there has been a tendency, mainly due to labour market deregulation, for employers to expand their temporary and part-time workforce (Peters, 2014). As a result, such a variable needs to be included in an inequality regression equation to control for the effects of job status (OECD, 2008).

3.3 Lerner Estimation

In this section, I attempt to estimate the market power of industries, which will be used in the regression analysis as a variable of interest in section 3.5. The relation between market power and wage inequality is not known *a priori*, but at least it is known that a wage premium exists in industries with large concentration (Belman, 1988; Borjas and Ramey, 1995; Gera and Grenier, 1994). To discover the exact relationship, therefore, a regression estimation is necessary that will be addressed in section 3.5.

To estimate the Lerner's values, as a proxy for market power, I have mimicked Hall's methodology (Hall, 2018). I need to mention that Hall has made his Matlab codes publicly available on his website[10]. In his method, Hall uses the micro definition of Lerner's index to derive an equation that requires only aggregate data for estimation of Lerner's index. Cost is defined as $C = \sum_i (w_i x_i)$ with technology $y = Af(x)$. Given that Lerner's formula is $\mathcal{L} = \frac{p-mc}{p}$, with some extra algebra he produces an interesting equation as follows:

$$\Delta \log(\mathrm{y}) - \sum_i \alpha_i \Delta \log(x_i) = \mathcal{L}\Delta \log(\mathrm{y}) + (1-\mathcal{L})\,\Delta \log(\mathrm{A}) \qquad \text{equation } 3-1$$

where $\alpha_i = \frac{w_i x_i}{py}$. The left-hand side of this formulation is actually the Solow residual calculated in the productivity accounts (Hall 2018). With a time-series or panel data, one can treat $\mathcal{L}$ as a parameter to be estimated. The only issue left to address is the correlation between productivity and output. To control for such issues, Hall uses instrumental variables. These instrumental variables need to be a driver of y but not affected by y. Military purchases of equipment, military purchases of ships, military purchases of software, military expenditure on R&D, and oil prices are the instrumental variables that Hall has used in his study.

Following Hall's method, I will use the same procedure to estimate Lerner's value for two-digit industries based on NAICS 2007. Having said that, I will have to use a different set of instrumental variables simply because a breakdown of military expenditures is not available for Canada; only total expenditures are available through IndexMundi. As a result, commodity prices and total military expenditures will be used as instrumental variables in

[10] https://web.stanford.edu/~rehall/Recent_Unpublished_Papers.html

this study. Commodity prices include an energy price index, a metal and minerals price index, a forestry price index, an agriculture price index, and a fish price index. All of these commodity price indices are available on the Bank of Canada website[11]. Multifactor productivity data is also available online on the Statistics Canada website[12]. Such data is available for only 17 industrial groups based on the North American Industry Classification at the two-digit level.

Figure 3-1 illustrates the instrumental variables that were used in the estimation process. In the first stage of regression with the OLS model, the instruments were able to account for 67% of the variation in output growth from 1986 to 2015. In Hall's work, the corresponding value is 59%.

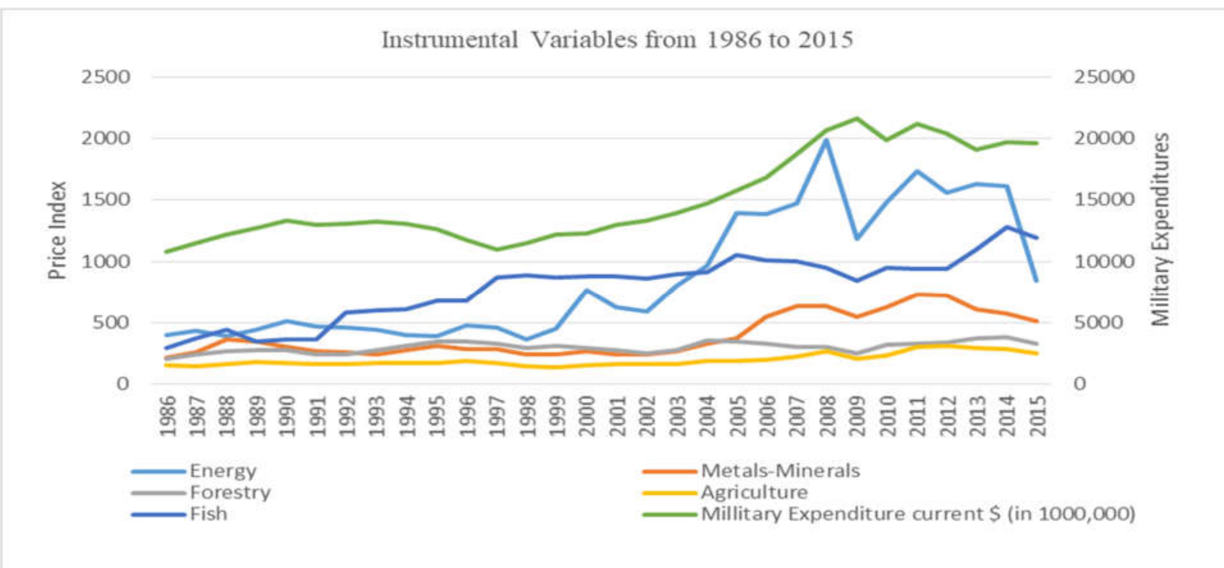

Figure 3-1. Instrumental variables for GDP

[11] https://www.bankofcanada.ca/rates/price-indexes/bcpi/

[12] https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=3610021701

The next step concerns the estimation of Lerner's index. Table 3-1 reports the estimated values for 17 NAICS industries. Similar to Hall's findings, some of my estimated Lerner's values are negative, whereas the true values of Lerner's must be non-negative by definition. Hall argues that such a problem arises due to data noise and measurement errors. To disentangle the measurement errors, he specifies that $L = \mathcal{L} + \eta$ where L is the estimated values of Lerner's, $\mathcal{L}$ is the true values, and η captures sampling error. To discover the effect of sampling error and data noise on the estimated values, he assumes that the true values of Lerner's index are distributed as beta(ν, β) with density proportional to $\mathcal{L}^{(\nu-1)}(1-\mathcal{L})^{(\beta-1)}$ $and\ \mathcal{L} \in [0,1]$.

Figure 3-2 shows the histogram for measurement errors from my results for Canada. It is evident that there is a good deal of noise around zero, and also around -0.3, consistent with Hall's findings and leading to his conclusion that all the negative values are attributed to measurement errors. Furthermore, Figure 3-3 plots the fitted values versus actual values (the convolution of $\mathcal{L}$ and η) with a pretty good fit.

Table 3-1: Estimated Lerner's values by two-digit sectors

Sector name	**Lerner index**	**Bootstrap standard error**
Finance, insurance, real estate and renting and leasing	-0.38	0.32
Administrative, waste management and remediation services	-0.13	0.14
Construction	-0.08	0.05
Mining and oil and gas extraction	0.03	0.23
Accommodation and food services	0.04	0.25
Health care and social assistance (except hospitals)	0.10	0.28
Manufacturing	0.14	0.02
Professional, scientific and technical services	0.16	0.11
Information and cultural industries	0.17	0.09
Other services (except public administration)	0.17	0.18
Wholesale trade	0.26	0.09
Retail trade	0.29	0.11
Arts, entertainment and recreation	0.32	0.27
Transportation and warehousing	0.39	0.13

Educational services (except universities)	0.44	0.17
Agriculture, forestry, fishing and hunting	0.48	0.08
Utilities	0.87	0.38

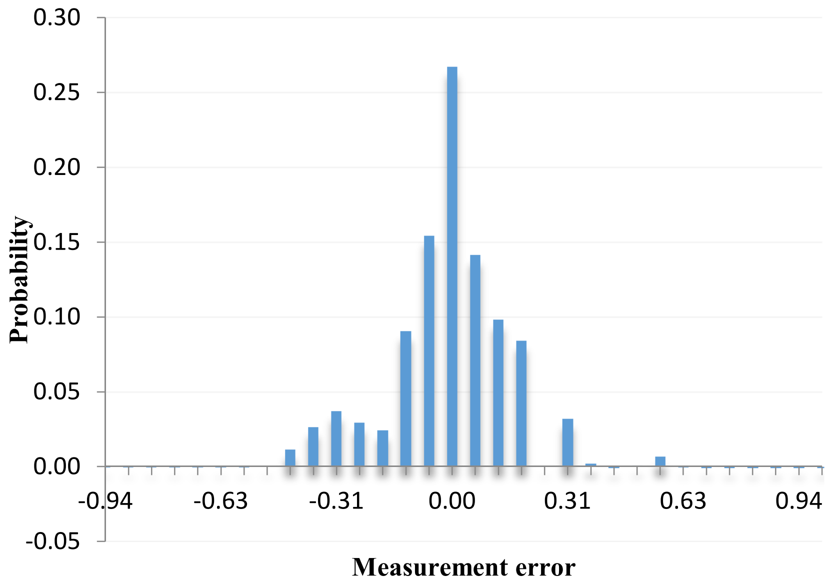

Figure 3-2. Distribution of measurement errors

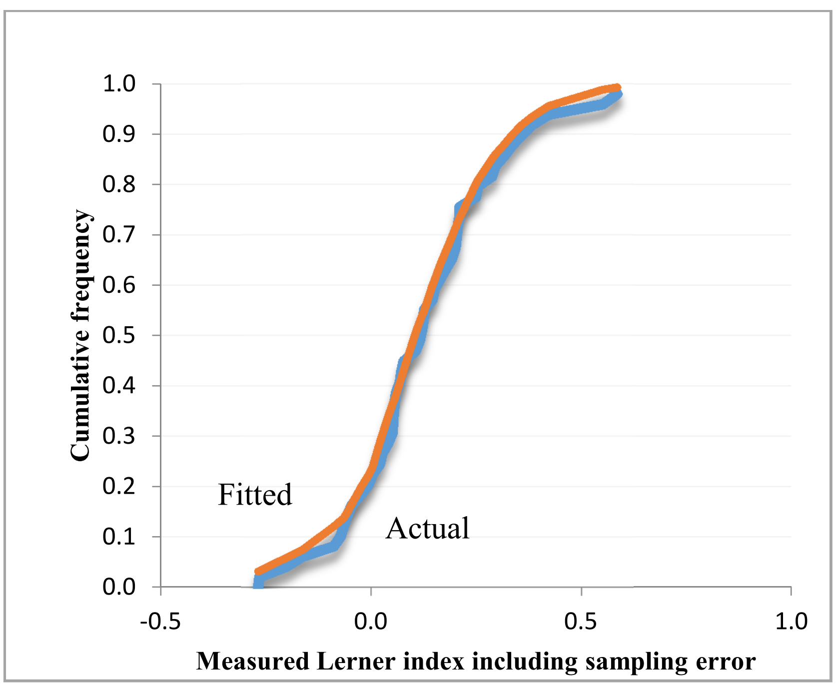

Figure 3-3. Actual cumulative frequencies of estimates and calculated cumulative distribution functions from the statistical model

In addition to Lerner's values estimated, Hall's method also estimates a growth coefficient for each industry to reflect the average variation of sectoral Lerner values over time. With Canadian data, Table 3-2 presents the growth coefficients for the same industries. Of the 17 sectors, 12 sectors have been experiencing a decline in market power over time, including the utility sector, which had the highest Lerner's value in the previous table. On the other hand, monopoly power has been rising for the other five industrial sectors, i.e. Manufacturing, Finance, insurance, real estate and renting and leasing, Construction, Wholesale trade, and Health care and social assistance.

Table 3-2: Estimated growth rate of monopoly power by sector

Sector name	Growth coefficient (ψ)	Standard error
Utilities	-0.05516	0.02523
Other services (except public administration)	-0.02554	0.01439
Transportation and warehousing	-0.02393	0.00802
Arts, entertainment and recreation	-0.0214	0.03115
Administrative, waste management and remediation services	-0.01665	0.01357
Mining and oil and gas extraction	-0.01533	0.01575
Educational services (except universities)	-0.01488	0.01823
Accommodation and food services	-0.01124	0.01438
Professional, scientific and technical services	-0.00638	0.0094
Agriculture, forestry, fishing and hunting	-0.00343	0.00712
Retail trade	-0.00189	0.00942
Information and cultural industries	-0.00101	0.00699
Manufacturing	0.002426	0.00193
Finance, insurance, real estate and renting and leasing	0.003675	0.01431
Construction	0.004139	0.0052
Wholesale trade	0.011964	0.00982
Health care and social assistance (except hospitals)	0.029689	0.02363

With the Lerner's values and the growth coefficients just estimated for an individual sector, it is possible to calculate the Lerner's values from 1995 to 2015 (to match it with Census years). Given that the time period for Lerner estimation is from 1986 to 2015, the year 2000 is selected as the base year (because it is in the middle of the time period) to which the estimated Lerner's values in Table 3-1 are assigned. Subsequently, the Lerner's values are calculable for other years using the growth coefficients.

These newly generated values will be used in my main regression analysis. The only issue left is that here I have only 17 industry groups, whereas in the Census 20 industry categories are reported. More specifically, Multifactor productivity data combines the Finance and insurance industry (NAICS 51) with Real estate and rental and leasing industry (NAICS 53). Also, Management of companies and enterprises (NAICS 55) along with Public administration (NAICS 91) are missing in Multifactor Productivity accounts that cause missing values in the Lerner's variable. For the regression analysis, the same Lerner's values for both NAICS 51 and NAICS 53 are considered. As for the missing industries (i.e. NAICS 55 and NAICS 91), I used the average of all the estimated Lerner's values to fill the rest of the missing values.

3.4 Theil's Decomposition Method

To break down inequality by industry, Theil's technique introduced by Shorrocks and Wan (2005) is applied to industrial groups. Alternative approaches such as regression-based decomposition, the coefficient of variation (CV), variance of log wages, 90/50 and 50/10 percentiles, and the Gini index are also available in the literature to quantify the industry aspects of income inequality, but the most consistent and commonly-used inequality indicator for decomposition purposes belongs to the entropy family popularized by Theil (Akita, 2003; Shorrocks and Wan, 2005).

Decomposition of inequality by industry requires partitioning the sample into a set of groups, and then calculating two components of aggregate inequality: 1) a within-group component which captures the inequality within the group (or between the members of the group, here the industrial sector as a group); and 2) a between-group component which captures the inequality due merely to variation across groups (Shorrocks and Wan, 2005).

The industry decomposition of inequality is typically conducted by using Theil indices, T and L. The T index uses income shares as weights while the L index uses population shares as weights. The T index is, thus, more sensitive to the changes in richer regions, while the L index is more sensitive to the changes in poorer regions (Akita, 2003). What follows will concisely explain the decomposition method. Please note that the original method developed by Shorroks and Wan (2005) uses spatial units as subgroups for decomposition purposes, but their method is applicable to any type of subgroup, including industrial subgroups (Galbraith and Hale, 2014). As a result, I will use the original decomposition method to quantify the contributions of industries to inequality. According to Theil's indices, the earnings inequality of an economy is defined in equations (3-2) and (3-3).

$$T = \frac{1}{N}\sum_i \frac{y_i}{\mu}\ln\frac{y_i}{\mu} \qquad \text{(equation 3-2)}$$

$$L = \frac{1}{N}\sum_i \ln\frac{\mu}{y_i} \qquad \text{(equation 3-3)}$$

$$\mu = \sum_i \frac{y_i}{N} \qquad \text{(equation 3-4)}$$

where T and L are the Theil inequality indices, N is the total population, y_i is the income of individual i, and μ is the mean income of the total population.

I will start with the decomposition of the first index, T. As Shorroks and Wan state, the intuition behind this decomposition is as follows. At the individual level in an equal economy, the individual's share of total income should be equal to the individual's share of the total population. At the industry level in an equal economy, an industry's share of income (resources) in total income (resources) should be equal to the industry's share in total population (employment). Put differently; if an industry makes up 10% of the total employment, in an equal economy, it should own 10% of all available resources across all industries.

Plugging equation (3-4) in (3-2) and some simple manipulation leads us to the following equation:

$$T = \sum_i \frac{y_i}{Y} \ln\left(\frac{y_i/Y}{1/N}\right) \qquad \text{(equation 3-5)}$$

Clearly, T is comparing the income share of an individual (y_i/Y) with the population share of that single individual (1/N). Assuming that the total population is partitioned into j industries, it can be shown that the industrial decomposition of overall inequality, T, can be achieved through the following:

$$T = \sum_j \frac{Y_j}{Y} T_{WR}^j + \sum_j \frac{Y_j}{Y} \ln\left(\frac{Y_j / Y}{N_j / N}\right) \qquad \text{(equation 3-6)}$$

$$\text{where } T_{WR}^j = \sum_i \frac{y_{ij}}{Y_j} \ln\left(\frac{y_{ij} / Y_j}{1 / N_j}\right) \qquad \text{(equation 3-7)}$$

In equation 3-6, T_{WR}^j is the inequality within only industry j; the first term in the right-hand side is the overall within-industry inequality, T_{WR}; and the second term is the between-industry inequality T_{BR}. Therefore, expression $\frac{T_{BR}}{T}$ enables us to measure to what extent the industrial disparity contributes to overall inequality. Furthermore, since the decomposition is an additive function, it is also possible to extract the inequality share of a single industry as within- and between-components of inequality.

With respect to the L index, Shorrocks and Wan (2005) propose a method by which the inequality index can be decomposed into both within and between components. The following counter-factual questions can illustrate the intuition behind this decomposition; How much inequality would occur if industrial income differences are the only sources of inequality?" Put differently, to capture the between-component of inequality, the authors replace the income of each individual with the mean income of their respective industry. In this regard, the only source of inequality comes from industrial disparities resulting in mean-income differences across industries.

Let the set of individuals be partitioned into m subgroups N_k $(k=1,2,..,m)$, with corresponding income vector y^{k}, mean incomes μ_k, population sizes n_k, and population shares $\gamma_k = n_k / n$ where n is the total population. Thus, it can be shown that the L index is decomposable to the following within and between elements:

$$L = \frac{1}{n}\sum_{k=1}^{m}\sum_{i\in N_k}\ln\frac{\mu}{y_i}$$

$$= \sum_{k=1}^{m}\frac{n_k}{n}\frac{1}{n_k}\sum_{i\in N_k}\ln\frac{\mu_k}{y_i} + \frac{1}{n}\sum_{k=1}^{m}\sum_{i\in N_k}\ln\frac{\mu}{\mu_k} \qquad \text{(equation 3-8)}$$

$$= \sum_{k=1}^{m}\gamma_k L_W^k + \sum_{k=1}^{m}\gamma_k \ln\frac{\mu}{\mu_k} = L_W + L_B$$

where $L_W = \sum_{k=1}^{m}\gamma_k L_W^k$ is the within-group component of inequality, and $L_B = \sum_{k=1}^{m}\gamma_k \ln\frac{\mu}{\mu_k}$ is the between-group contribution to overall inequality, which is obtained by replacing the income of each person with the mean income of the corresponding group. As in the T index, the L index is additively decomposed. Thus, it enables us to quantify the industry contribution of any single industry to overall inequality.

3.5 Empirical Analysis

The empirical process is composed of two phases. In the first stage, individual wages deflated by the provincial consumer price index are considered to calculate the T index for all observations. Simultaneously, overall inequality is broken down by industry into two components, a between-industry component and a within-industry component. The results of the first round are presented in Figure 3-4, Table 3-3, and Table 3-4. In the second round, a pooled OLS regression is conducted in which within-industry wage inequality (obtained from the first stage) serves as a dependent variable. Independent variables constitute several relevant determinants (including Lerner's index) that have been generated from the data sources mentioned above. Results of OLS regression are reported in Table 3-5.

3.5.1 Inequality Trend and Decomposition by Industry

Figure 3-4 plots the evolution of overall wage inequality across Canada (excluding observations from PEI) from 1995 to 2015. As shown, the T-value increased from 0.73 in 1995 to 0.82 in 2005 (12% growth). By 2010, the T-value had decreased back to its 2000 level at 0.76 before increasing to 0.79 in 2015. The graph also demonstrates the breakdown of T-index by industry, i.e., within-industry and between-industry inequalities. As is evident, most of the overall inequality has to do with the within-component of inequality. The contribution of within-component and between-component to overall inequality is given in Table 3-3. It is shown that 90% of the overall inequality is attributable to the within-industry component, whereas the share of the between-component is around 10% in each Census year. But it is notable that the between-industry inequality has increased by 10% from 1995 to 2000, remained unchanged through 2010, and increased by another 10% from 2010 to 2015.

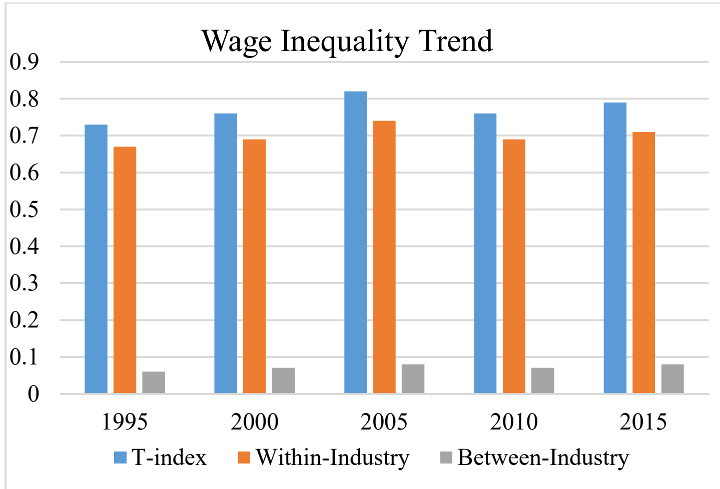

Figure 3-4. Theil's index with decomposition across Canada

Table 3-3: Share of within- and between-industry components in overall inequality

Year	Within-Industry	Between-Industry
1995	91%	9%
2000	90%	10%
2005	90%	10%
2010	90%	10%
2015	89%	11%

In addition, within-industry T-values are provided in Table 3-4. The average T-values of all the industries for each year are also provided at the bottom of the table. Over the course of 20 years, Agriculture, forestry, fishing, and hunting sector, as well as Mining (mining and oil extraction) sector, have had large levels of earnings inequality. And their inequality trend has remained relatively stable over time as there is only a 5% and 8% decline in their T-values by 2015, respectively. Large T-values can also be observed in the Management of companies and enterprises sector, although its measure of inequality has sharply dropped by 2005. Despite a 32% decline in its T-value, earnings inequality remains to be high in Management of companies and enterprises sector as the corresponding 2015 T-value is

0.85. The Finance and insurance sector, along with Real estate, rental and leasing experienced a low level of inequality in 1995, but these sectors have undergone a sharp increase in their T-values of 42% and 46% by 2015, respectively. Also, their T-values reached high levels at 0.86 and 0.97. Arts, Entertainment, and Recreation also have high T-values with a relatively small upward change in the 20-year time period. Despite a high T-value for the Accommodation and Food Services sector in 1995, it ranked below the average level in 2015. Furthermore, the Retail trade sector has constantly scored high T-values, with a relatively stable trend. On the contrary, there exist several other sectors that have had low T-values over time, including Information, the Professional, Scientific, and Technical sector, Public Administration, Health Care and Social Assistance, Educational Services, and Transportation and Warehousing.

Table 3-4: Within industry wage inequality

NAICS Industry (two-digit level)	1995	2000	2005	2010	2015	20-year growth rate
Agriculture, Forestry, Fishing and Hunting	1.83	1.88	1.87	1.81	1.74	-5%
Mining, and Oil Extraction	1.26	1.32	1.51	1.42	1.17	-8%
Utilities	0.56	0.60	0.62	0.69	0.60	8%
Construction	0.77	0.76	0.90	0.82	0.75	-3%
Wholesale Trade	0.63	0.63	0.72	0.61	0.69	9%
Information	0.44	0.54	0.56	0.53	0.53	19%
Finance and Insurance	0.59	0.74	0.87	0.71	0.84	42%
Real Estate and Rental and Leasing	0.66	0.71	0.96	0.77	0.97	46%
Professional, Scientific, and Technical	0.58	0.57	0.65	0.59	0.61	5%
Management of Companies and Enterprises	1.25	0.82	1.04	0.96	0.85	-32%
Administrative and Waste Management	0.65	0.71	0.76	0.72	0.75	15%
Educational Services	0.61	0.59	0.61	0.60	0.58	-6%
Health Care and Social Assistance	0.67	0.68	0.67	0.66	0.63	-6%
Arts, Entertainment, and Recreation	0.95	0.93	0.92	0.98	1.02	7%
Accommodation and Food Services	0.87	0.85	0.89	0.81	0.76	-12%
Other Services (except Public Administration)	0.75	0.73	0.71	0.71	0.69	-7%
Public Administration	0.53	0.53	0.50	0.52	0.49	-8%
Manufacturing	0.68	0.70	0.74	0.71	0.73	6%
Retail Trade	0.82	0.85	0.92	0.83	0.89	8%
Transportation and Warehousing	0.61	0.61	0.66	0.61	0.65	6%
Average of all industries	0.79	0.79	0.85	0.80	0.80	0.04

3.5.2 OLS Regression Analysis at Industry Level

Regression model:

At this point, the second stage of estimation is implemented in which the dependent variable is the log of within-industry inequality values provided by the T index, as presented above[13]. The explanatory control variables include demographic and non-demographic variables. Given that the industry fixed effects are already extracted in the first stage, a pooled OLS regression (rather than a panel regression) is implemented in the second step with the following equation:

$$\begin{aligned}\log(T_{jt}) = a + \beta_1 time + \beta_2 lerner_{jt} + \beta_3 unionshare_{jt} + \beta_4 importexposure_{jt} \\ + \beta_5 exportexposure_{jt} + \beta_6 innovationshare_{jt} + \beta_7 naicsgrowth_{jt} \\ + \delta demography_{jt} + \varepsilon_{jt}\end{aligned}$$

where T_{jt} represents the T index for industry j at time t, $time$ is a time dummy variable, $lerner$ proxies for sector monopoly power, $unionshare$ measures the share of union members for industry j at time t, $importexposure$ measures the extent with which industry is exposed to international import, $exportexposure$ controls the exposure of industry to export, $innovationshare$ measures the share of computer and R&D expenditures in industry j at time t, $naicsgrowth$ measures the output growth of a given industry over

[13] Scatter plots of the dependent variable with independent variables are available in Appendix I

time, and $demography$ is composed of demographic variables (male share, skill-share proxied by education share, foreign-born-share, full-time share) and their interactions. Parameters $a, \beta, and\ \delta$ are to be estimated.

Regression results:

The results of pooled OLS regression are provided in Table 3-5. Two sets of models are reported; model I and model II. Model I is the basic model, whereas the second model includes the interaction of demographic variables. The reason that I am adding the interactions is that a group of workers with a specific characteristic may have an advantage over another group with a different characteristic. For instance, a skilled worker may have a higher chance of finding a full-time job versus an unskilled worker (Foley and Green 2016). Moreover, the chances that a Canadian-born worker holds a fulltime job could be higher compared to a foreign-born worker integrating into the Canadian labour market. Or, there may exist wage differentials between Canadian skilled workers versus foreign-born skilled workers, and therefore such a gap may impact the T-values.

Table 3-5.

Independent variables	**Model I**	**Model II**
	Coefficient (Robust Std. Err.)	Coefficient (Robust Std. Err.)
Time	0.0567*** (0.01)	0.063*** (0.016)
Lerner	-0.195** (0.08)	-0.227** (0.10)
Union share	-0.008*** (0.001)	-0.004** (0.001)
Export exposure	0.231*** (0.06)	0.2*** (0.06)
Import exposure	0.09** (0.04)	0.09*** (0.03)
Innovation index	-0.0003*** (0.0001)	-0.0001 (0.0001)
NAICS growth	0.074	0.14

	(0.12)	(0.10)
Male share	-0.002	-1.7***
	(0.001)	(0.45)
Canadian-born share	0.034***	-0.783***
	(0.005)	(0.232)
Full-time share	-0.002	-0.624***
	(0.002)	(0.233)
Skill share	0.002	-4.07***
	(0.001)	(0.84)
Constant	-2.74***	60.8***
	(0.48)	(18.3)
Male share X Canadian-born Share		0.022***
		(0.005)
Male share X Full-time share		0.018***
		(0.005)
Canadian-born Share X Full-time share		0.008***
		(0.003)
Male share X Canadian-born Share X Full-time share		-0.0002***
		(0.00006)
Male share X Skill share		0.097***
		(0.021)
Canadian-born Share X Skill share		0.053***
		(0.01)
Male share X Canadian-born Share X Skill share		-0.001***
		(0.0002)
Full-time share X Skill share		0.045***
		(0.009)
Male share X Full-time share X Skill share		-0.001***
		(0.0002)
Canadian-born Share X Full-time share X Skill share		-0.0005***
		(0.0001)
Male share X Canadian-born Share X Full-time share X Skill share		0.00001***
		($3.1e^{-06}$)
R-square	0.73	0.81
F(11,68)	17.5	
F(22,57)		21.79

*Statistical significance of the estimates for 10%
**Statistical significance of the estimates for 5%
***Statistical significance of the estimates for 1%

Table 3-6: Margin effects of demographic variables after accounting for interaction terms

Variable	dy/dx	p-value
Male share	-0.003	0.06
Canadian-born share	0.032	0.00
Full-time share	-0.01	0.02
Skill share	-0.004	0.08

Focusing on the first model, most of the estimated coefficients are statistically significant at the conventional confidence intervals. Starting from the first variable, the estimated coefficient indicates that inequalities within industries increase by 5.6% on average every five years, which confirms the rising trend of Canadian earnings inequality over time as documented in (Gray et al., 2003; Pavcnik, 2011) even when other explanatory variables are considered.

The next variable is the Lerner index. Consistent with rent sharing explanation (Borjas and Ramey, 1995; Furman and Orszag, 2015; Gera and Grenier, 1994), the estimated coefficient turns out to be negative. It indicates that a larger monopolistic power of an industry is partially correlated with lower inequality within industries. The estimated parameter indicates that one basis point increase in the Lerner's index (i.e., a change of 0.01 in Lerner's value) is associated with a 0.19% decrease in within-industry inequality.

Assuming that rent-sharing and efficiency wages can, too, be practiced at sector level (not just at firm levels), and by considering only the returns to labour input (i.e., ignoring the returns to capital), the estimated coefficient implies that industries with larger monopoly power tend to share the labour portion of the monopolistic gains with their workforce more equally.

The union variable also seems to reduce inequality as implied by the estimated parameter. A one unit increase in the percentage of workers covered by unions is associated with a 0.008% decrease in within-industry inequality. This finding is consistent with what is generally expected from the literature (Card et al., 2004). Trade variables are next in the queue.

Both of them positively impact inequality, a result that is generally expected by the literature (Krugman, 2008). Industries with more exposure to exports or imports would experience larger inequality. Industries with greater exposure to exports experience a 0.23% increase in their within inequality, whereas industries with more exposure to imports experience a 0.09% higher within inequality.

The innovation index, surprisingly, seems to have a negative impact on inequality. In other words, industries with a larger share of R&D and computer expenditures in their inputs experience a lower level of inequality. The significance of the innovation variable, however, is sensitive to model specification. In comparison with the literature, this finding is inconsistent because the literature generally expects a positive relationship (Breau et al., 2014). As mentioned earlier, this may have to do with the way the innovation has been defined. Following the literature (Hall, 2014), the input definition of innovation was chosen as a proxy, i.e. R&D and computer expenditures, whereas previous studies have used output definitions of innovation (e.g. patent data). Moreover, innovation has been measured across geographic areas in previous studies, whereas my unit of study is the industry. Last but not least, the definition of R&D has changed over time, which might cause measurement errors in my estimations as well. As a result, the innovation coefficient needs to be treated with caution.

The next variable is the rate of industry growth which is measured by the growth rate of industrial output over time. It is expected that a larger growth rate would result in a higher level of inequality as the gains from growth may not be equally distributed. Even though the sign of the coefficient matches with expectations, the estimated coefficient turns out to be statistically insignificant. The other factors above matter more according to my results.

Next comes the demographic variables. The results are mixed and sensitive to model specification. Male ratio, to control for gender effects, is only marginally significant at 10% confidence interval. The estimated parameters for the skill share and fulltime share variables are also statistically insignificant. However, all of these variables become statistically significant when interaction terms for the demographic variables are included in the second model. More explanations will be provided for these variables once I get to the second model.

With respect to the effect of place of birth, the estimated parameter is robust to model specification. Industries with a larger share of Canadian-born workers would experience a larger level of inequality. According to the estimated coefficient, a one basis point increase in the share of Canadian-born workers in an industry would cause the within industry inequality to rise by 0.034%. Few alternative interpretations can be made about this positive relation. The relation is positive perhaps because earnings inequality is large among Canadian-born workers of an industry due to unobservable characteristics. Alternatively, the reason may be associated with the sufficiently large earnings gaps that exist between immigrants and native-born workers (Block et al., 2019). Another interpretation could be made about this relation if we looked at it the other way around. As the estimate shows, an increase in the share of foreign-born (rather than native-born) workers is tagged with a lower level of earnings inequality within that industry. This relationship sounds convincing when one comes to realize that the economic outcomes of immigrants have been improving as a result of the changes to immigration policies since the 2000s (Picot and Hou, 2016).

A second model of regressions has been carried out as a supplement that allows for the interaction of demographic variables. Table 3-5, Model II reports the estimation results with new variables added. Although the magnitude of the estimated coefficients somewhat varies across the two models, all the non-demographic variables are robust to new specifications in terms of both significance and coefficient signs. The only exception is the innovation index, which was significant in the first model but became insignificant in the second model.

As for the demographic variables, the skill share and fulltime share variables became statistically significant when the interaction variables are included in the model. With the interaction variables included in the regression model, the "margins" syntax in Stata is considered to measure the average marginal effect of the demographic variables[14]. These margin effects have been reported in Table 3-6. The margin effects are significant at conventional 5% and 10% confidence intervals, as shown by the p-values.

With respect to the relation between within-industry wage inequality and skill share (share of workers with bachelor's degree or above), the sign of the variable is consistent with

[14] The margins command estimates margins of responses for specified values of covariates and presents the results as a table. It is a handy syntax in particular when regression equation carries a set of interaction variables. The marginal effect of variable x on variable y (i.e., dy/dx) in my calculations are set at the means of x and y.

SBTC argument suggesting that improvements in human capital at industry level are associated with a reduction in within-industry wage inequality. One basis point increase in the share of skilled workers within an industry (i.e., a change of 0.01 unit in skill share) decreases the earnings inequality by 0.004% within that industry. The effect seems to be small also consistent with recent studies by Foley and Green (2016), and Fortin and Lemieux (2014) who leave only small room for the role of skill in Canadian earnings inequality, mainly due to a boom in the energy regions which benefited mostly low skilled workers. If that's the case, then it perhaps makes sense that the estimated coefficient is insignificant in the first model, as the time period of my study also coincides with the energy boom era.

Considering the male-female effect, a one basis point increase in the share of male workers in an industry causes the within-industry inequality to decrease by 0.003%. In other words, in industries dominated by male workers, wages are distributed more equally than in industries with fewer men and more women. This finding implies that either inequality is larger among female workers so that a rise in their share in the workforce drives up the inequality, or that a sufficiently large gender gap may exist between male and female workers (Blau and Kahn, 2000) such that a rise in the share of female workers causes an unequal distribution of wages due to the gender gap. A combination of these two factors may also be relevant.

And lastly, the effect of part-time work is positive on within industry earnings inequality. One basis point increase in the share of workers with part-time job status would increase the within-industry earnings inequality by 0.01%. Although only a few Canadian studies are available with respect to this subject (relation between part-time work and inequality), it seems that a positive relationship between earnings inequality and part-time job status is expected in the literature (Peters, 2014; Simpson, 1986; OECD, 2008), an expectation consistent with my finding.

3.6 Conclusion

This empirical study intends to investigate various determinants of real wage inequality for two-digit Canadian industries for the years 1995 to 2015 using Census 1996, 2001, 2006, 2011, and 2016. Empirical results demonstrate that industry differentials account for about 10% of Canadian wage inequality in each census year, while the rest of inequality were attributed to variations within industries. Using the computed within-industry inequalities (measured by Theil's index) as a dependent variable, an OLS regression was undertaken to test the impact of several demographic and non-demographic determinants. Multiple sources of data have been used to create variables of interest. What follows briefly summarizes the findings and provides insightful implications.

A variable, not examined in the literature (at an aggregate level), is Lerner's index as a proxy for sectoral monopoly power. The marginal impact of monopoly power on wage inequality is found to be negative. That is, the payments to labour input are more equally distributed among workers in industries with larger market power. And a larger Lerner's value can essentially be achieved through either an increase in the average price that the industry is selling its product for or a decrease in the marginal cost of production in that industry where both of them can be influenced by domestic and global forces.

International trade is also another factor that affects wage inequality within-industries as shown by the regression results. The estimated parametrs exhibit that exposure to both imports and exports would worsen the earnings distribution within the industry and the the effect of exports is even stronger than imports. In contrast, labour unions seem to be directing the earnings dispersion toward an equal way of distribution. My findings related to trade variables, and labour unions are consistent with literature as well.

Industries with a larger share of full-time employments would observe a lower level of inequality. Also, employing more foreign-born workers is linked to a reduction in wage inequality within industries. In addition, industries dominated more by skill workers would have a more equal earnings distribution. As for the gender variable, i.e. share of male workers in a given industry, the estimated value reflects a more equal distribution of wages within industries with larger male share.

Overall, it can be concluded that there exist several factors to consider when evaluating wage inequality at industry levels. For instance, in the mining and oil industry, from Figure 3-1, we can observe how the prices of energy, and metal and minerals have changed over time, which influences the Lerner's value of this industry. At the same time, trade exposure (mostly export exposure than import) of this industry varies when there is enough change in the price of the industry's product. Lerner's effect on within-industry wage inequality is negative, whereas it is positive for the trade-exposure index. Which effect is stronger is a question that should be addressed in future research.

Although my goal was not to offer specific policy suggestions, my study suggests that it is important to consider how earnings inequality is shaped by the structure of industries, since different types of industries have different structures of employment and different distributions of wages. Thus, the industrial dimension of earnings inequality can be an aspect to consider when economic policies are designed for specific industries.

For instance, competition policies in industries with internal economies of scale that affect the ease of mergers and acquisitions to alter market power in the affected industries could also increase earnings inequality within that industry. Of course, such policies would have to be weighed against their potential adverse impact on the efficiency of the industry. In this case, to offset the resulting efficiency losses partially if not fully, further regulation might be needed to direct the industry resources toward larger spending on R&D, which in turn appears to dampen the earnings inequality as well.

Policies are often concerned with industries that are exposed to international trade, perhaps as a consequence of new trade agreements. For instance, from Table 3-4, we have seen that Mining has had consistently large levels of inequality over the 20-year time period of our study. This sector is strongly affected by international trade, but the gains from trade have been distributed unequally. To lead the industry toward a more equal wage distribution, policymakers might focus on several factors. Strengthening the presence of labour unions (and labour market regulations in general) in a sector of this nature might reduce wage inequality. Alternatively, sector-specific or more general policies to encourage greater expenditures on R&D and computers could also dampen earnings inequality levels in these exposed sectors.

Chapter 4

Wage Inequality for Canadian Metro Areas: Investigating the Effect of Average City Rent and Regional Growth Clusters (1995-2015)

4.1 Introduction

It is well documented that earnings (wage) inequality have been growing since the 1980s in Canada (Fortin and Lemieux, 2014; Green and Sand, 2015). Consequently, a large body of studies has been conducted to quantify the magnitude of Canadian earnings inequality as well as to explore the causes. A substantial portion of these studies has examined inequality and its determinants from a national perspective, whereas little study has been devoted to the spatial aspect of wage inequality as indicated by Florida and Mellander (2016); and Fong, (2017). .

From a theoretical point of view, there exist a number of interesting studies that attempt to provide theoretical connection between urbanization and wage gap. For instance, in the urban-labour economics literature, spatial sorting is a well-known theory which indicates that high ability workers, as well as good-performing firms, tend to locate in large cities where they can gain from both consumption and production amenities (Glaeser and Maré 2001; Florida 2003; Moretti and Thulin 2012; Moretti 2013). This tendency leads to a larger and more productive city, which enables the workers and firms to generate higher levels of earnings (Andersson et al., 2014; Combes et al., 2008; Matano and Naticchioni, 2011; Mion and Naticchioni, 2009; Moretti, 2004).

Additionally, Haworth et al. (1978) develop a “monopoly hypothesis” to explain how urban growth and urban size affect income distribution. They argue that following the increases in city size and growth, two groups of individuals benefit disproportionately, thereby contributing to increasing inequality: 1) those owning fixed assets, like landowners, 2) certain monopolists who are likely insulated from the competition in business or employment. For the first group, increasing the population drives up the prices for local fixed assets, which in turn provides some rents for the owners. For the second group, monopolists are able to benefit not only from obstacles to the entry of new businesses but also to gain additional advantages due to a larger market.

Although few studies can be found in the literature, history of the urban dimension of inequality can be dated back to 1978 when Haworth et al. (1978) found a positive relationship for income inequality with urban growth and city size. Since then, growing interest has been evolved among practitioners to probe more into this subject by testing different influencing factors at various geographical levels.

What follows sets out an overview of selected studies for wage and income inequality at the geographic level (including metropolitan areas). And this review will attempt to draw out the specific factors emphasized in each study with a view to incorporating them in a more general empirical analysis.

On the U.S. side, Chakravorty (1996) investigates income inequality across US metro areas for total, white and black populations using 1990 Census data, and finds that the structure of income inequality among races varies across metro areas. Further findings are discovered by Florida and Mellander (2016) who investigate wage inequality (measured by Theil's index) as well as income inequality (quantified by the Gini coefficient) for US metropolitan areas in 2010. They find that wage inequality explains only 15% of the income inequality. They also indicate that these two are quite different in determinants. Their findings illustrate that skill, human capital, technology, and metro size are strong determinants of wage inequality whereas the same factors are only weakly linked to income inequality. On the contrary, the geographic variation of income inequality is more associated with the tax rate, race, and de-unionization.

Korpi (2007) tests “human capital theory” and “central place theory” to explore whether the size of local labour markets is a source of inequality. Using the Swedish dataset from 1990 to 2002, the empirical results show that there is evidence for a positive relationship between local labour market size and wage inequality. More importantly, the effect of local labour market size is larger at the top of the wage distribution (i.e., larger cities get to experience a larger inequality at the top of the distribution).

On the Canadian side, Finnie (2001) uses the Longitudinal Administrative Database (LAD) to highlight the importance of the provincial aspect of earnings inequality over the 1982-1994 period. The findings illustrate that there is a great deal of discrepancy in earnings inequalities across provinces, with the Atlantic provinces experiencing the larges levels of inequality from 1982 to 1994. Similar evidence is provided by Heisz (2016) who examines the trends of income inequality across Canada and shows that inequality has been rising in all Canadian provinces where Ontario has experienced the largest rise in its Gini coefficient between 1985 and 2011.

Furthermore, Marchand (2012, 2015), and Fortin and Lemieux (2014) find further evidence, from 1971 to 2012, that energy price is an important factor contributing to provincial differences in wage inequality.

Considering metro areas as unit of study, Moore and Pacey (2003) find that inequality has been rapidly rising in CMAs compared to non-CMAs from 1981 to 1996. The authors also show that a tendency of immigrants to live in CMAs has significantly and increasingly contributed to the changes in inequality among Canadian households from 1981 to 1996.

Bolton and Breau (2012) and Breau et al. (2014) test the effects of various variables on wage inequality at the city level, and find that there is a direct relationship between city size and the inequality level. In addition, other factors such as innovation, deindustrialization, and share of visible minorities also played a role in inequality differentials at city level.Similarly, Walks et al. (2014) study inequality and polarization in Canadian CMAs and find that inequality and polarization have increased within and between CMAs.

And lastly, using the Canadian Income Survey (CIS), Fong (2017) investigates the inequality trends at the CMA level from 1982 to 2014. Gini coefficients for market income inequality show that metropolitan areas are rapidly becoming more unequal, while non-metropolitan areas are becoming more equal. In particular, Toronto, Montreal, Vancouver and Calgary have experienced the fastest growth in their inequality. The study also examines the effect of tax transfers on CMA inequality and shows that although after-tax income inequality is smaller than market income inequality, the former is still large and rising as well.

According to the above-mentioned studies, no study has investigated the effect of average city rent and regional growth cluster (RGC) on wage inequality within CMAs. This essay aims to explore whether these potential determinants are affecting CMA inequality (measured by Theil's index)? If so, what is the magnitude of the effects?

With respect to average city rent, Beaudry et al. (2012) demonstrate that industrial composition (in terms of high-pay industry and low-pay industry measured by average city rent index) of a city matter to the wage structure of that city. They show that a larger composition of high-paying industries in a local economy will boost the wages in all the other sectors in that economy. Building on a search and bargaining theory with multiple sectors and local markets included in the model, it is illustrated that a shift in the composition of industries away or toward high-paying jobs would come with substantial spillover effects, thereby influencing the wage structure of the urban economy.

Regarding the RGC variable, it is intended to capture the effect of regional (CMA) economic growth (captured by those industrial clusters that have contributed to the regional growth of that CMA; i.e. RGCs) on within-CMA wage inequality. A full description of RGC identification is provided in section 4.2 following Slaper et al. (2018). An increase in the growth rate of regional growth clusters may dampen or exacerbate CMA wage inequality, depending on how equal gains from growth are distributed among workers.

To achieve the above-mentioned goals, a two-round empirical analysis is conducted that in the first round overall inequality will be quantified using Theil's index across all CMAs and subsequently decomposed by CMAs (to obtain within- and between-CMA Theil's values). The decomposition results show that most of the earnings inequality across CMAs is attributable to the within-component rather than to the between-component (which explains only around 2% of overall earnings inequality in each year of the study). To further explore the within-CMA inequality variation, in the second round, the Theil's values for each CMA from the first stage are extracted and regressed over a number of explanatory variables to quantify the effect of ARC and RGCs as well as other control variables from the literature.

It is found that ACR negatively impacts the within-CMA inequality; i.e. CMAs with a larger composition of high-pay jobs would experience a lower level of inequality, keeping other factors constant. Regional Growth Clusters (RGCs) also found to be an important determinant of within-CMA inequality. It is shown that a faster rate of growth of industrial clusters tends to increase wage inequality by 0.51%. All the other control variables are estimated more or less consistent with what the literature generally expect. More details are provided in section 4.4.

Section 4.2 proceeds to data description and construction of variables. Section 4.3 reports the findings for inequality across and within CMAs. Pooled OLS regression is conducted in section 4.4 to test the effect of potential determinants of within-CMA inequality. And section 4.5 concludes.

4.2 Data Description and Variables

The datasets in this study largely overlap with those in Chapter 3. Similarly, each data source serves to construct a variable or a set of variables of relevance that will be directly or indirectly included in the regression analysis. What follows provides a brief description of the datasets and their contribution to the analysis.

- **Census data to construct economic and demographic variables:**

  Census data for 1996, 2001, 2006, 2011 and 2016[15] are the primary data sets in this study. As cited in Chapter 3, a set of demographic variables are necessary as control variables in the regression analysis. To capture the potential gender effects on inequality, a male share is calculated for CMAs as the share of male workers in a CMA[16]. Skill share is also computed as the share of workers with a bachelor's degree or above. To control for immigration effects on inequality, the share of foreign-born workers is measured as the share of workers who were born outside of Canada versus those who were born in Canada. And finally, the share of workers with part-time job status is constructed for each CMA to capture the effect of job status on earnings inequality.

[15] Please note that the sample is taken from only employed and self-employed individuals who reported annual earnings greater than zero in the year prior to the Census reference week (i.e. only from 1995, 2000, 2005, 2010, and 2015).

[16] Please note that only employed and self-employed individuals age 15 and over with full-time or part-time job status are included in the sample. Their earnings (or wages) are defined in the Census as wages, salaries, and bonuses.

In addition to demographic variables, four more variables were created using the Census: CMA size, an oil-region dummy, average city rent, and regional growth cluster. What follows is a brief clarification on why these variables are needed and how they are computed.

CMA size is a potential determinant of inequality, as the literature suggests that there is a positive relation between city size and inequality (Baum-Snow and Pavan, 2013). It will be measured by the size of the working force of age 15 years old and over. The oil-region dummy variable is meant to control for the effects of the energy boom or bust on inequality (Fortin and Lemieux, 2014; Marchand, 2015). Those CMAs that are located in Newfoundland, Saskatchewan, and Alberta are considered to be in oil regions.

Average city rent is a novel variable that, to my best knowledge, has never been tested in the literature. As explained in the Introduction, average city rent is an index that Beaudry et al. (2012) introduced to quantify the impacts that industrial composition of a city (in terms of high-pay industries, and low-pay industries) has on the wage structure of that city. A high value of the index means that the city's employment is mostly concentrated in high-paying industries. As shown in their model, high-pay jobs in an industry are supposed to boost wages in all other industries as a spillover effect. In this case, inequality may decline if the spillover effects are dominant, something that is going to be tested in the regression analysis. On the other hand, if the spillover effects are weak, then an increase in wage inequality would be expected.

Following Beaudry et al. (2012), to build the average-city-rent index, I first calculate the industrial relative mean wages at the national level, where the Manufacturing sector is the base industry. In the second step, a weighted average of industrial relative wages is computed to construct the index for each CMA. The weights are the employment share of each industry in a CMA; i.e. $\omega_{ij} = \frac{employment_{ij}}{employment_j}$ where i is industry and j indicates the CMA.

Regional growth clusters (RGCs) are also constructed using Census data. Before proceeding to the definition of RGCs, it is necessary to indicate that this variable is supposed to capture the impact of economic growth at the regional level. This growth indicator has been found to be relevant (but not tested yet) in the literature. For instance, Osberg (2018) is concerned that economic growth would not solve the issue of rising inequality in Canada (i.e. national level), and therefore encourages a shift in policy to address the issue. In another study, Bolton and Breau (2012) use the median wages of the labour force to quantify the effect of regional economic development on earnings inequality at the city level. They find that economic development has a dampening effect but fades away over time.
In this study, RGCs will be used to capture any potential effect of regional economic growth on within-CMA wage inequality. A full description of RGCs is provided in Slaper et al. (2018). In summary, RGCs are industrial clusters in cities that are a major force of employment growth. To be identified as an RGC, a sector needs to meet four statistical requirements:

1. The employment of the cluster has been increasing.

2. The cluster must be growing in relative importance in the city (i.e. the shift-share ratio for cluster cl, SS_{cl}, must be greater than one)
3. SS_{cl} must be less than 1 plus 2 standard deviations of all the SS_{cl} of the city.
4. The employment share of the cluster must be greater than 0.005 at the end of the time period (i.e. 2016 in this study).

The reason that I am choosing RGCs as my economic growth variable has to do with the urban aspect of the analysis. As an alternative, CMA's GDP data could be used to capture the growth effect, but with RGCs, I can be more specific and focus on a few industries that represent the features of a regional growth cluster. Besides, GDP data at the CMA level are available only from 2009 onward, whereas I need a longer period of coverage.

- **Input-Output tables from Statistics Canada to construct trade variables and an innovation index:**

  Using input-output data from 1997 to 2015, additional control variables are constructed to follow the literature.

  a) Export exposure and import exposure indices for each CMA: These indices are actually a modified version of those that were used in Chapter 3 in order to reflect the urban aspect of trade variables.

  To modify the indices from the prior chapter, the employment shares of industries in each CMA were calculated to create the weight of each industry in the total employment of a CMA; i.e. $\omega_{ij} = \frac{employment_{ij}}{employment_j}$ where i represents the industry, and j indicates the CMA. Having indices in one hand, and the industry weights on the other hand, a weighted average of import exposure and export exposure are calculated across industries within each CMA. The final output, subsequently, is meant to reflect the degree of exposure to international trade that a CMA has.

b) Innovation index: Similarly, with innovation index constructed in the previous chapter for industries, a weighted average was calculated to generate an innovation index for each CMA, and ω_{ij} was used as the industry weight.

- **Data on Unemployment Rate from Statistics Canada[17]:**

  Unemployment rates are necessary to be included in the regression analysis to control for any impacts that local unemployment has on inequality (Mocan, 1999). Although this study examines wage inequality among employed individuals, the unemployment rate may still affect the wages of employed workers as it might affect the bargaining power of workers in the local economy. If those workers are mostly low-wage individuals, then it is expected that earnings inequality would increase should the unemployment rate rises.

  Unemployment rates are available at the CMA level on a monthly basis from 2001 onwards. With seasonally adjusted unemployment rates, a simple average of rates over a 12-month period is used to obtain annual rates. CMA unemployment rate data is unavailable for the year 1995; therefore, I used the 2000 values to

[17] https://www150.statcan.gc.ca/t1/tbl1/en/cv.action?pid=1410029401

complete my data set (i.e., CMA unemployment rates for 1995 and 2000 are identical).

- **Government of Canada's Hourly Minimum Wage data to construct minimum wage variable:**

  Minimum wages vary across provinces as they are set by provincial legislation. It is expected that a higher minimum wage rate would boost the wages of individuals at the lower tail of the distribution, and therefore mitigate earnings inequality within local economies (Fortin and Lemieux, 2014). Data on hourly minimum wages are available by province since 1965 on the Government of Canada's website[18]. Those CMAs that share a province, therefore, will be assigned the same minimum wage[19].

- **Labor Force Survey (LFS) from Statistics Canada to construct the union share variable:**

  As described in the previous chapter, labour unions are considered to be a determinant of wage inequality (Card et al., 2004). To control for union effects, the monthly survey of the labour force (LFS) from 1997 to 2015 was employed to calculate the share of unionized workers for a given industry. Since LFS is a monthly survey, I used the average shares over a 12-month period in a given year. In the survey, the union status of workers has been categorized into three groups; a) covered by union b) not in a union but covered by a collective agreement and c)

[18] http://srv116.services.gc.ca/dimt-wid/sm-mw/rpt2.aspx?GoCTemplateCulture=en-CA

[19] Minimum wages will be deflated by provincial consumer price index in the regression analysis.

no coverage. Categories (a) and (b) are combined in the construction of union shares.

4.3 Inequality Trend and Decomposition by CMA

To measure inequality and break it down by CMA, I will use the spatial decomposition approach developed by Shorrocks and Wan (2005). A comprehensive description of the decomposition method was provided in Chapter 3. Briefly speaking, with Shorrocks and Wan's decomposition methodology, first, Theil's index is calculated as overall inequality across all observations (where observations are individual real annual earnings[20] across all CMAs), and decomposed into two groups: a) between-CMA inequality, and b) within-CMA inequality.

Decomposition results demonstrate that most of the earnings inequality across CMAs is attributable to within-CMA inequality rather than to between-CMA inequality. Put differently, total wages have been distributed among CMAs in proportion to their populations, whereas earnings have been distributed unequally within the CMAs.

Table 4-1 and Figure 4-1 report the decomposition results. The blue column in the chart represents the T-values of all the observations in the sample. It shows that overall earnings inequality (i.e., inequality across all CMAs) steadily rose from 0.37 in 1995 to 0.49 in 2005. Then it decreased to 0.41 in 2010, followed by an increase back to 0.48 in 2015. With

[20] Earnings are defined as wages, salaries, and bonuses, which have been deflated by provincial consumer price index. Also note that earnings and wages will be used interchangeably hereafter.

respect to decomposition by CMA, the between component accounts for only a small portion of the inequality, whereas the within-component is very large. The contribution of each component to inequality remains stable over time.

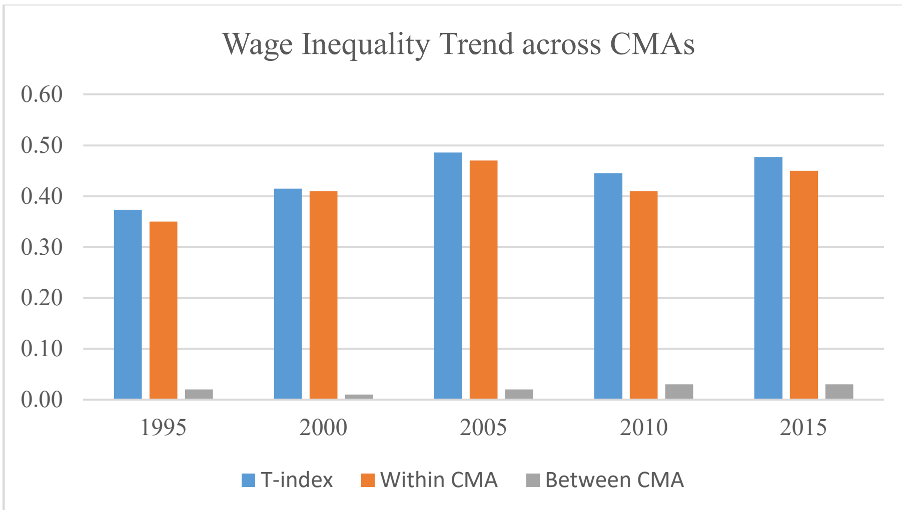

Figure 4-1. Overall, within, and between wage inequality across CMAs over time

Table 4-1: Within and between CMA inequality

Year	Overall inequality across CMAs	Within-CMA	Between-CMA
1995	0.37	0.35	0.02
2000	0.42	0.41	0.01
2005	0.49	0.47	0.02
2010	0.44	0.41	0.03
2015	0.48	0.45	0.03

Moreover, Figure 4-2 illustrates the earnings inequalities for individual CMAs over time[21]. A great deal of variation in within T-values can be observed both across CMAs and within CMAs over time. In particular, Calgary, Edmonton, and Saskatoon experienced a spike in 2005, recording the highest earnings inequality in their history. Those CMAs that are in

[21] Table of values are provided in Appendix II

the same province seem to have a similar pattern and level of inequality. For instance, Victoria, Abbotsford, and Kelowna in British Colombia (BC) have T-values in the same range, although Vancouver has the largest t-values in BC. Of all the CMAs, Calgary, Toronto, Vancouver, and Montreal have always had T-values above average.

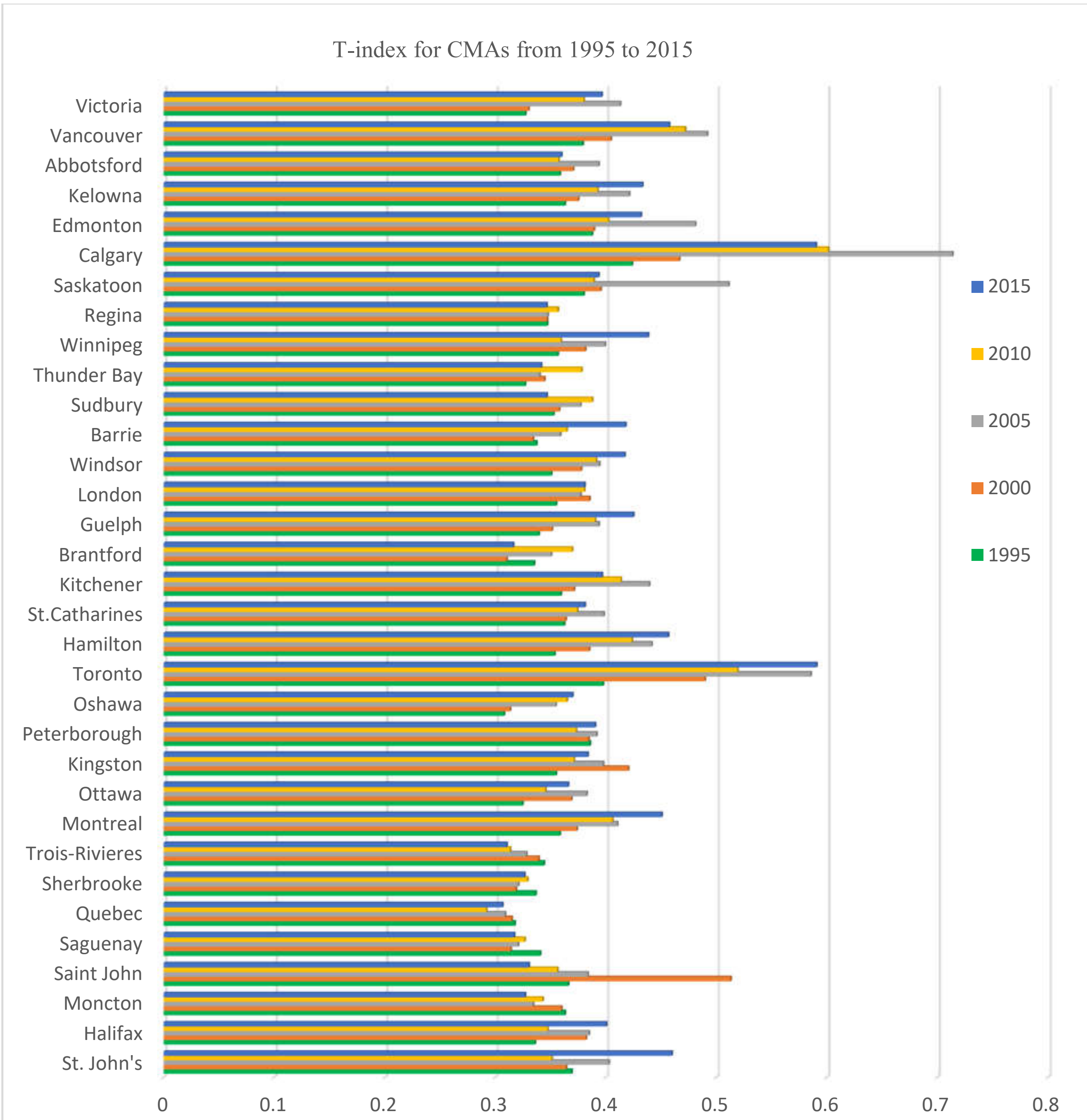

Figure 4-2. Within CMA Theil's values over time

4.4 Regression Analysis

To conduct the regression analysis, several variables have been constructed using the above-mentioned data sources. In summary, all the demographic variables (such as gender, education, job status (full time or part-time), and place of birth) come from the Census as well as CMA size. Trade variables, as well as the innovation index, come from the input-output tables from Statistics Canada. The union variable was calculated using LFS data available from Statistics Canada. Unemployment rates are collected from Statistics Canada. Minimum wage data are gathered from the Government of Canada's website.

Similar to the previous chapter, the independent variable is the log of within-CMA T-values that were presented in Figure 4-2. Given that the effect of CMA differentials on inequality have been already extracted in the first step, I will use a pooled OLS regression in the second step with the following equation[22]:

$$\begin{aligned}\log(T_{jt}) = a &+ \beta_1 time + \beta_2 \log(AverageCityRent_{jt}) + \beta_3 RGCgrwoth_{jt} \\ &+ \beta_4 \log(CMAsize_{jt}) + \beta_5 CMAminWage_{jt} + \beta_6 importexposure_{jt} \\ &+ \beta_7 exportexposure_{jt} + \beta_8 \log(innovation_{jt}) + \beta_9 unionshare_{jt} \\ &+ \beta_{10} oilregion + \delta demography_{jt} + \varepsilon_{jt}\end{aligned}$$

where T_{jt} represents Theil's values for CMA j at time t, $time$ controls for time effect, $AverageCityRent$ quantifies to what degree CMA j is composed of high-pay jobs,

[22] Scatterplots of T index with other control variables are provided in Appendix III

$RGCgrwoth$[23] captures the impact of RGCs, $CMAsize$ controls for the size of CMAs, $CMAminWage$ controls for minimum wages, $importexposure$ and $exportexposure$ control for trade effects on inequality, $innovation$ proxies how innovative a CMA is, $unionshare$ measures the share of union members for a given industry, $oilregion$ is dummy variable to distinguish between CMAs with oil and CMAs with no oil reserves, and $demography$ is composed of demographic variables (male share, skill share proxied by education ratio, foreign-born share, and full-time share) for CMA j at time t, and ε_{jt} is the error term. Parameters $a, \beta, and\ \delta$ are to be estimated.

The estimates are presented in Table 4-2. In the first model (Model I), only the key variables are included in the regression, while the second model (Model II) adds other control variables. While some variables are robust to model specification, others are sensitive. The effect of time is positive in both models and significant at conventional confidence intervals. It shows that earnings inequality within CMAs is increasing by 6.3% every five years, a finding that is consistent with a trend of rising inequality in CMAs (Gray et al., 2003).

Table 4-2: OLS estimation results (log of T index as the dependent variable)

Independent Variables	**Model I** **Coefficient** **(Robust Std. Err.)**	**Model II** **Coefficient** **(Robust Std. Err.)**
Time	0.0900** (0.0410)	0.0638*** (0.0171)
Log of ACR	-0.2578*** (0.0822)	-0.1461** (0.0724)
RGC growth	1.3425*** (0.4954)	0.5161*** (0.1935)
Log of CMA size	0.3720*** (0.0829)	0.1779** (0.0695)
Minimum wage (real)	-0.0674***	-0.0461**

[23] Growth rate is measured by employment growth of the RGCs.

	(0.0341)	(0.0196)
Export exposure	0.0992**	0.0374
	(0.0489)	(0.0229)
Import exposure	-0.0404	0.0046
	(0.0509)	(0.0259)
Log of innovation	-0.0182	-0.0270***
	(0.0129)	(0.0058)
Union share		-0.0063***
		(0.0015)
Oil-region dummy		0.1204***
		(0.0337)
Foreign-born share		0.4308***
		(0.1355)
Part-time share		0.9909**
		(0.4922)
Male share		-0.7105
		(1.0029)
Skill share		1.4264***
		(0.4013)
Unemployment rate		0.0178***
		(0.0059)
Manufacturing share		0.3265
		(0.2863)
Cons	-2.5***	-1.3107**
	(0.36)	(0.5628)
R-squared	0.611	0.8787
Prob > F	0.00	0.00
F(8, 123)	11.94	
F(16, 115)		107.87

*Statistical significance of the estimates for 10%
**Statistical significance of the estimates for 5%
***Statistical significance of the estimates for 1%

Average city rent is also robust and significant in both models, although the estimated co-efficient is cut by half in the second model when additional controls are present. It reads that average city rent is negatively related to within-CMA inequality, and this relationship is statistically significant according to the second model. This negative relation implies that a large mix of high-pay jobs in a city can lift the wages in other sectors sufficiently higher to partly offset the industrial differentials and contribute to a decline in inequality. It is, I believe, one interesting application of the index in Beaudry et al. (2012) to wage inequality. They state that a concentration of high-paying jobs in a city would boost the wages in other sectors through general equilibrium. The implication for earnings inequality is that, if this spillover effect is negligible, then inequality is expected to rise, whereas inequality would decline should the spillover effect be strong enough. My findings support the latter.

RGC growth (measured by the employment growth of RGCs) also matters to earnings inequality within CMAs. The sign and significance of the estimated coefficient are robust, but the magnitude of the coefficient is sensitive to model specification. It is shown that the faster the regional growth clusters grow, the larger would be earnings inequality. Also, a one basis point increase in the growth rate of the RGC is associated with a 0.51% increase in earnings inequality within CMAs. The magnitude of the effect is relatively large. This finding is interesting as it supports Osberg (2018), who believes that a higher rate of economic growth is not able to solve the rising inequality concern. In the context of this study, it is interpreted that the gains from the growth of industrial clusters are distributed largely unequally among the workers.

With respect to real minimum wages, the estimated parameter seems to be relatively robust in both coefficient and significance, and it is negative, as is generally expected from the literature (Fortin and Lemieux, 2014; Lee, 1999). It is also found that a $1 increase in minimum wages (equivalent to 12% increase relative to the mean of minimum wages during the period) is accompanied by a 4.6% decrease in earnings inequality within CMAs. Trade variables, however, seem to be irrelevant when looking from an urban aspect at earnings inequality, as they are insignificant at the 5% or 10% significance levels. The best results have been reported in the table of results, in which the trade exposure variables show up as dummy variables[24].

By contrast, the oil region variable seems to be a better explanatory determinant of wage inequality than the trade variables mentioned above. It is significant and indicates that those CMAs within oil regions experience, on average, a 12% higher level of earnings inequality, assuming other factors constant. It is also worth mentioning that my data set covers a time span that overlaps with the energy boom era when oil prices were high. With the recent changes in the economic conditions, however, I expect to observe a decline in wage inequality in oil-region CMAs (among employed and self-employed workers), although this requires further research in the future.

It is also interesting that the log of innovation index turned significant when the control variables were included. It seems that when the innovation index rises, earnings inequality tends to decline. Having that said, this variable needs to be treated with caution. The innovation index was calculated based on an input definition of innovation (i.e. share of R&D and computer expenditures), which carries noise because the definition of R&D has changed over time.

Labor unions also matter to wage inequality. The estimated coefficient is significant and in line with what the literature has found (Belman, 1988; Card et al., 2004; Farber and

---

[24] In the estimation process, trade exposure index for a CMA is equal to zero if it is less than the median across all CMAs, and equal to one if otherwise.

Western, 2001). It indicates that those CMAs that have a larger share of covered workers[25] would observe lower levels of inequality. With a one basis point increase in union share (i.e., a change of 0.01 unit), it is expected that wage inequality within a CMA falls by 0.006%.

With respect to other control variables, a larger share of foreign-born workers in a CMA is associated with a higher level of earnings inequality within that CMA. More specifically, a one basis point change in the share of foreign-born workers is associated with a 0.43% change in the T-index. This relation is also in line with what Block et al. (2019) have found. They demonstrate that unemployment rates are usually higher among racialized workers, and foreign-born male workers have earned 22% less than their Canadian-born counterparts since 2005. In addition, wage discrimination is more acute for female immigrants as they have received 59 cents for every dollar that a non-racialized man earns (Block et al., 2019).

Job-status also matters to wage inequality within CMAs. Economies with a larger share of part-time workers are expected to realize a notably larger earnings inequality. The magnitude of the effect is relatively large and indicates that for a one basis point change in the share of part-time workers, there would be a 0.99% change in inequality. Peters (2014) shows that there has been a tendency among employers toward market deregulation, in particular, an interest in employing more part-time workers. In this case, it could be con-

[25] covered by union or collective agreement as reported by respondents.

cluded that further market deregulation toward expanding the size of the part-time workforce comes with a large upside that influences earnings inequality within CMAs, something to consider if the goal is to control wage inequality at CMA levels.

With respect to skill share (or education ratio), the estimated effect is significant and relatively large as well. It can be read that one basis-point change (i.e., a change of 0.01) in the level of skill share comes with a 1.42% increase in the earnings inequality within CMAs. This evidence supports the literature, which expects a positive effect. For instance, Acemoglu (1999, 2002) argues that although a rise in the proportion of skilled workers compresses the returns to skill and reduces inequality, it has a larger effect, which can generate job polarization, thereby increasing the unemployment of unskilled workers and rising inequality[26]. This finding also supports Foley and Green (2016), who argue that increasing investments in education (in particular at university level) is not an effective policy response to rising inequality.

In addition to all the determinants above, the unemployment rate of a CMA seems to be another factor affecting wage inequality of the CMA. The estimated coefficient is significant and consistent with the findings in the literature (Beach et al., 2006; Mocan, 1999). One basis point increase in the unemployment rate tends to increase the CMA's earnings inequality by 0.018%. Apparently, a higher rate of unemployment is reducing the bargaining power of workers, in particular those in the lower end of the wage distribution. Also, a

---

[26] More discussion of this paper is available in General Literature Review.

larger unemployment rate may potentially hurt the casual workers more than others by reducing working hours, thereby contributing to larger earnings inequality.

4.5 Conclusion

Using multiple data sources, this empirical study examines annual real wage inequality for CMAs in Canada for the years 1995, 2000, 2005, 2010, and 2015. Briefly, the empirical analysis is composed of two rounds. In the first step, a sample of individuals (full-time and part-time, employed and self-employed) is chosen to conduct a decomposition analysis in which overall earnings inequality (across CMAs) is measured, and simultaneously decomposed by CMAs. It is found that not much inequality exists among CMAs, whereas most of the observed inequality is accounted for by the variations within CMAs.

To further explore inequality, in the second stage, the potential determinants of wage inequality are examined using pooled OLS regression where the dependent variable is the earnings inequality of CMAs from the first round (i.e. within-CMA inequalities). And the independent variables constitute a wide range of conventional relevant factors, as well as two new variables, namely, average city rent and RGC. What follows is a short review of empirical results.

Average city rent (ACR) is a novel variable that is investigated for the first time. A full description of this factor is provided in section 4.2. It is intended to quantify the composition of high-pay and low-pay jobs in a CMA. It is found that ACR is adversely related to wage inequality in the sense that a 1% increase in the composition of high-pay jobs in a

CMA is associated with a rise in wages in other sectors sufficient to decrease inequality by 0.14%.

The growth rate of major industrial clusters in a CMA (referred to as regional growth clusters or RGCs) also matters to within-CMA inequality, and the effect is positive. For an industry to be identified as an RGC of a CMA, four statistical requirements need to be met by that industry (a full description is available in section 4.2). It is estimated that a faster rate of growth of industrial clusters is linked to an increase in wage inequality of 0.51%.

Both union rate and minimum wages tend to reduce wage inequality, although it varies more in response to a one-unit change in the minimum wage than in the union rate. More specifically, a $1 increase in minimum wages (equivalent to 12% increase relative to the mean of minimum wages during the period) is accompanied by a 4.6% decrease in earnings inequality within CMAs. And with a one basis point increase in union rate, it is expected that earnings inequality within a CMA falls by 0.006%.

The innovation index is also tested in this paper. Measured by the share of R&D and computer expenditures, the estimated parameter is sensitive to model selection, but it seems that a larger expenditure on R&D and computers tends to mitigate the inequality, although only by 0.02%.

Immigration (defined as pertaining to those who are born outside of Canada) was also tested as a control variable. A larger share of foreign-born workers in a CMA corresponds to a higher level of wage inequality within that CMA by 0.43%, relatively large effect.

Job-status (measured as the share of part-time vs full-time workers) is also correlated with wage inequality within CMAs that has received very little attention in the literature. Not surprisingly, the results indicate that for a one basis point increase in the share of part-time workers (i.e., a change of 0.01 in the share of part-time workers), there would be a 0.99% rise in inequality.

When looking at wage inequality from an urban perspective (as opposed to the previous essay, which considered the industrial dimension), the estimated coefficient is large and significantly positive for the skill-share variable.

Last but not least, the unemployment rate of a CMA is also a determinant of wage inequality within CMAs. Although only employed and self-employed individuals are included in the sample of observations, it seems that a higher rate of unemployment may reduce the bargaining power or alternative job choices of workers, thereby causing the wage inequality to rise.

Rudimentary policy implications can also be found in this investigation. ACR and RGC are two interesting determinants of within-CMA wage inequality. ACR implies that attracting high-paying firms (industries) would alleviate the level of wage inequality within CMAs. And even a more fundamental way is to create these high-paying jobs by providing suitable infrastructures. RGCs can also be considered for tax regulations when an alleviating inequality is under consideration. An alternative implication concerns growth policies where a more balanced growth rate (between RGCs and non-RGCs) can be considered to mitigate within-CMA inequality.

Chapter 5

Effect of Interprovincial Migration on Canadian Wage Structure: Evidence from 2016

5.1 Introduction

Interprovincial migration in Canada has been in decline, as documented in Saunders (2018). Figure 5-1 maps a historical trend of interprovincial migration in Canada since the 1970s. Despite a declining trend, internal migration has been viewed to be a powerful factor affecting the Canadian economy in general, and local economies in particular. For instance, Coulombe (2006) states that positive net migration in a province comes with aggregate human capital gains in the long run for that province, and also for Canada in general. In another study, Sharpe et al. (2007) study the role of internal migration on aggregate output and labour productivity in Canada from 1987 to 2006. Interestingly, they find that interprovincial migration accounted for 6.2% of labour productivity growth in Canada in 2006. These studies show that there should exist a potential relationship between internal migration and wage distribution.

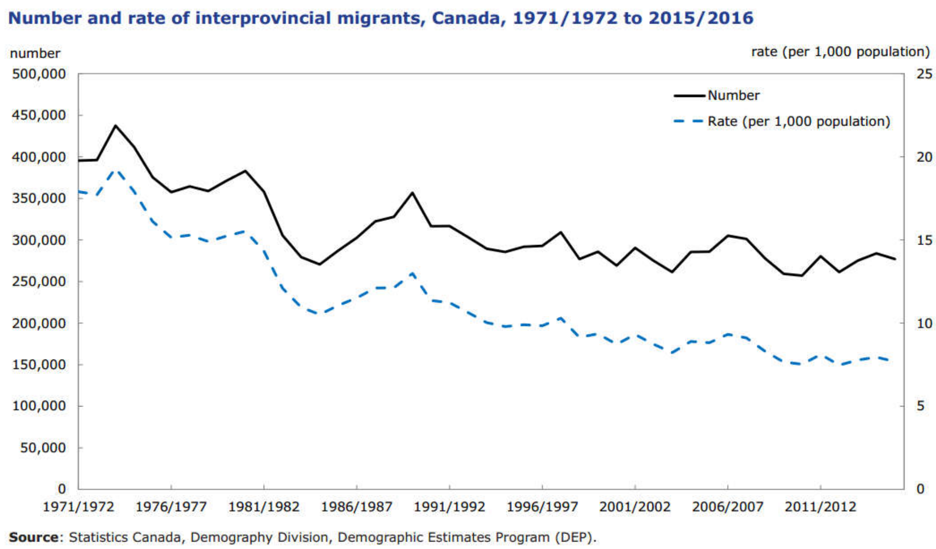

Figure 5-1. Migration pattern in Canada since the 1970s

In fact, there are only few studies that have concentrated on wage inequality in the context of interprovincial migration. One recent discussion is that of Mirjam et al. (2019). It is an empirical investigation that uses a counterfactual framework to quantify the effect of internal migration on the wage structure of migrants in the West African Economic and Monetary Union (UEMOA). Similarly, Phan and Coxhead (2010) develop a theoretical framework based upon the gravity model and find that internal migration is a determinant of relative wage inequalities across provinces in Vietnam during the country's transition.

Having said that, there exist theories that can be used to link the internal migration to wage inequality from theoretical point of view. In general, the choice of migration and its implications for the earnings distribution can be considered to be a specific case of Roy's model (Roy, 1951), as Heckman and Honore (1990) point out. In the context of Roy's model, individuals self-select themselves between two skill-specific sectors based on their comparative advantage to maximize their income. In this setting, Heckman and Honore (1990) show that the pursuit of comparative advantage, in the form of self-selection between sector 1 or sector 2, is accompanied by a decrease in overall and within wage inequality in contrast to a situation where individuals are assumed to be randomly assigned to the sectors. Similarly, Kanbur and Rapoport (2005) introduce agglomeration and spillover forces to conventional models of migration and find that the effect of migration on wage inequality could go either way depending on how the education function is formulated in the model.

Furthermore, insightful studies are available in the literature that have examined migration choices, and their determinants. Extending Roy's model to migration, Robinson and Tomes (1980) use the 1971 Census microdata and show that migrants tend to be self-selective. And the migration choice is impacted by potential wage gains.

Similarly, Coulombe (2006) conducts a comprehensive empirical analysis to investigate the determinants of migration and the distributional effects of migration on human capital in Canada. Using a pooled time-series cross-sectional approach and data since 1977, the author finds that migration is dependent on factors such as differences in long-run unemployment rates, and labour productivity which are associated with wage gains in the literature.

Unless it is put into an empirical test, according to the above-mentioned studies, it may be safe to say that the inequality effects of migration are not known *a priori*, although a mitigating effect is generally expected. Therefore, the goal of this research is to measure any potential impact of interprovincial migration on the wage distribution in Canada and compare the results with available opposing theories.

To test the impact of internal migration, a semiparametric counterfactual model is used that was developed by DiNardo et al. (1996). The concept behind this model is simple but complex to compute (more on this in section 5.3). In a counterfactual analysis, we are interested to learn about the wage structure in the absence of internal migration. Once estimated, a comparison between observed (factual) wage structure and corresponding unobserved (counterfactual) wage dispersion can be made to better understand the effect of internal migration.

Briefly speaking, using the 2016 Census PUMF, the findings show that wage inequality would have been almost twice larger in the absence of interprovincial migration (i.e. 0.79 vs 0.40). In addition, internal migration exerts differing impacts along the wage distribution. The largest impacts can be observed in the middle and lower portion of wage distribution; i.e. the mass of workers rises largely in the middle of the distribution, increases in the lower end but by a smaller size, and decreases in the mid-lower portion. Also, internal migration seems to have less effect on the upper portion of the wage density, and the effect gradually disappears in the higher end of the distribution. Moreover, the impact of migration is more or less in line with the above-mentioned effects across skills because in the absence of migration, a wider wage density curves, and larger Theil's values would have prevailed for both skilled and unskilled workers.

Section 5.2 overviews dataset. Section 5.3 presents the counterfactual framework, followed by a report of findings in section 4.4. And section 5.5 concludes.

5.2 Data Overview

The main data set for this study comes from the 2016 Census Public Use Microdata File (PUMF). Migrants are defined as those workers who have changed their province of residence within the last five years of the reference day of the Census. On the contrary, non-migrants are those workers who did not change the province of residence in the same period.

Observations in the sample include only individuals age 15 or over who have reported being in the labour force on the reference day of the Census. In this sample, there are 424,744 observations, of which 13,521 are migrants (3.2% of the sample). Annual wages, salaries, and bonuses are considered to assess the wage structures. These wages are deflated by the provincial consumer price index (2002$) to reflect a more accurate level of real wages.

Table 5-1 demonstrates a two-way tabulation of migrants by age and destination province. Similar to findings in the literature, individuals age 20 to 44 years old comprise over 75% of the migrant population. Of all the migrants, almost 70% of the migrants have chosen among the three provinces of Alberta, Ontario, and BC, with Alberta accommodating by far the largest share of the migrant population (30%) in this sample.

Table 5-1: Province of current residence of migrants by age groups (2016)*

	Province of residence										
Age	**NL**	**PEI**	**NS**	**NB**	**QC**	**ON**	**MB**	**SK**	**AB**	**BC**	**Total**
15 to 17	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.3%	0.1%	0.6%
18 to 19	0.0%	0.0%	0.1%	0.0%	0.1%	0.4%	0.1%	0.1%	0.6%	0.3%	1.7%
20 to 24	0.2%	0.1%	0.8%	0.4%	0.9%	2.6%	0.4%	0.7%	4.4%	2.7%	13.4%
25 to 29	0.3%	0.1%	1.1%	0.7%	1.5%	4.7%	0.6%	1.3%	8.0%	4.5%	23.2%
30 to 34	0.4%	0.1%	1.0%	0.5%	1.2%	4.5%	0.7%	0.8%	5.0%	4.0%	18.5%
35 to 39	0.3%	0.1%	0.7%	0.4%	1.0%	2.9%	0.4%	0.7%	3.6%	2.2%	12.5%
40 to 44	0.3%	0.1%	0.5%	0.4%	0.6%	2.1%	0.3%	0.5%	2.5%	2.0%	9.3%
45 to 49	0.2%	0.1%	0.3%	0.2%	0.4%	1.7%	0.3%	0.4%	2.1%	1.4%	7.1%
50 to 54	0.2%	0.1%	0.3%	0.4%	0.4%	1.1%	0.2%	0.4%	1.7%	1.1%	6.0%
55 to 59	0.2%	0.0%	0.2%	0.2%	0.2%	0.9%	0.1%	0.3%	1.2%	1.0%	4.4%
60 to 64	0.1%	0.0%	0.2%	0.1%	0.2%	0.4%	0.1%	0.2%	0.6%	0.6%	2.4%
65 to 69	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.1%	0.2%	0.7%
70 to 74	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%
75 to 79	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Total	2.3%	0.7%	5.2%	3.4%	6.5%	21.5%	3.4%	5.3%	30.1%	20.2%	100%

*Migration to the territories is small and ignored.

It is also worthwhile to touch on the wage structure of workers in different sub-samples using graphs. Figures 5-2 and 5-3 illustrate the wage kernel density of male workers and female workers by education (bellow bachelor's vs bachelor's or above) for both migrants and non-migrants, respectively. Yellow lines are included in the graphs to show the mode of migrants in their corresponding wage distribution.

Considering male migrants, having a bachelor's degree or above somewhat improves the wage distribution of male migrants by pushing a larger mass toward the middle. The education effect seems to be stronger for female migrants because it not only increases the mass at the mode but also provides a higher return, to some extent, compared to other female migrants who have education below a bachelor's degree.

Also, the variance of the wage densities widely varies across sub-groups, a sign of varying within wage inequalities[27]. In comparison with stayers, the wage structure of migrants appears to be wider (red curves vs blue curves). Therefore it may be expected that wage inequality will be larger in migrant groups compared to non-migrants. This is a question that will be addressed in section 5.3.

[27] In a normally distributed wage density, a larger variance is associated with larger inequality (DiNardo et al., 1996).

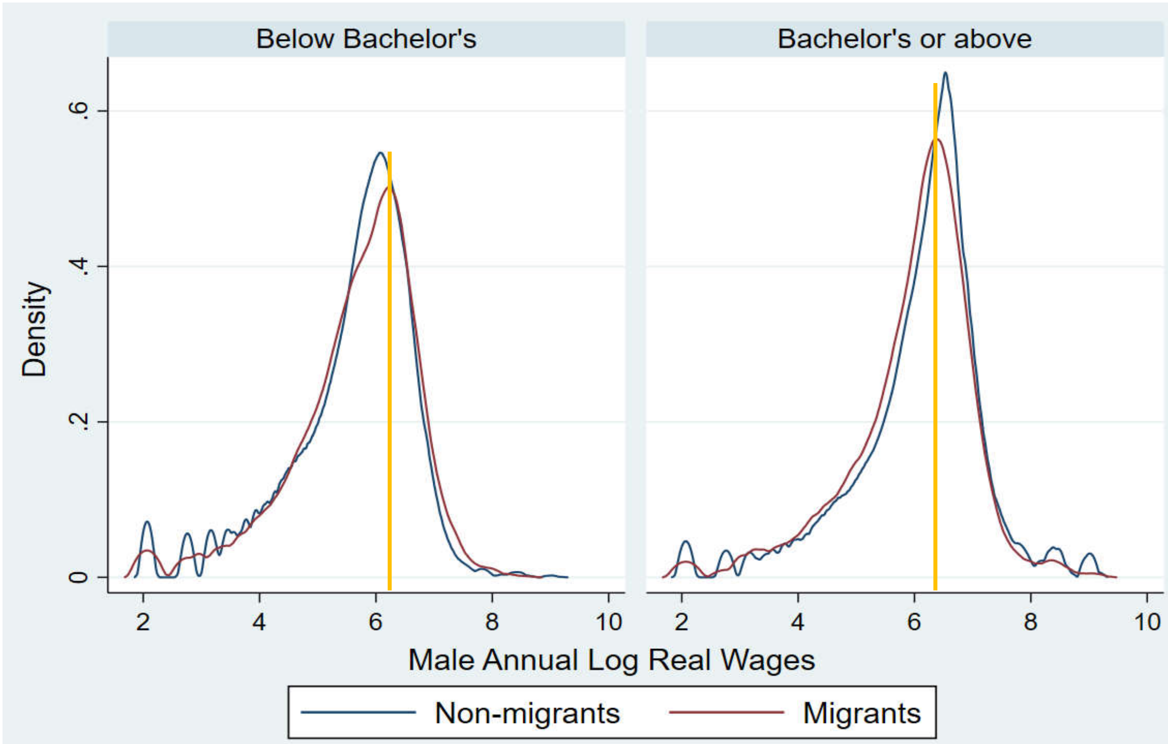

Figure 5-2. Male wage structure by migration and education

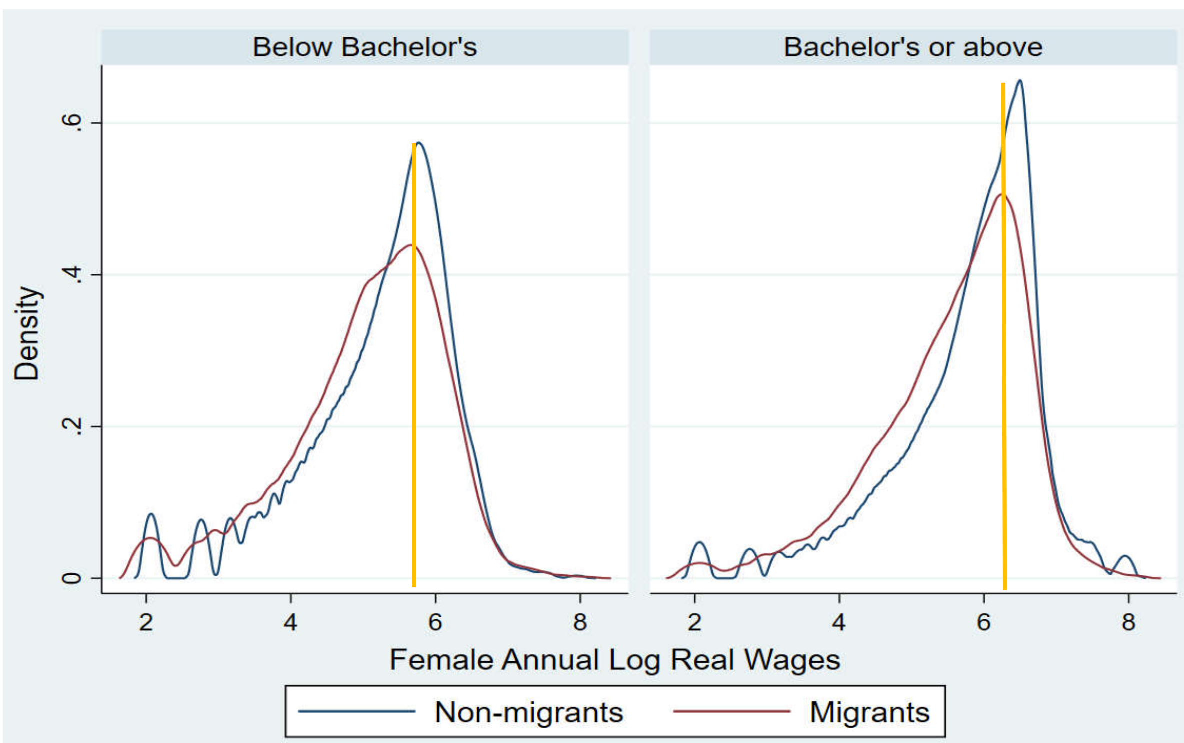

Figure 5-3. Female wage structure by migration and education

5.3 Empirical Analysis

This section proceeds to the counterfactual analysis of wage dispersion in Canada. The aim is to evaluate the effect of interprovincial migration on overall wage structure, as well as that for skilled and unskilled workers. The concept behind the counterfactual analysis is simple. What wage distribution would have prevailed if there was no interprovincial migration? In fact, the current wage dispersion observed in the sample is a combination of migrants' and non-migrants' wage structures (i.e. wage distribution with migration), whereas in a counterfactual case, a wage distribution is estimated in the absence of migration. Therefore, the counterfactual wage structure is not observable and needs to be estimated.

To obtain estimation results, a methodology introduced by DiNardo et al. (1996)[28] is considered. In this context, the counterfactual wage structure is going to be a weighted average of the conditional wage dispersion of migrants. More accurately, the estimated wage dispersion attempts to measure the wage distribution of non-migrants if there was no migration and if these non-migrants had been paid according to the wage structure of migrants.

Applying the wage structure of migrants to their stayer counterparts may sound problematic at first glance, in particular, because these two classes of workers are not identical. Fortunately, DiNardo, Fortin, and Lemieux's approach (DFL from now on) use a re-weighting measure in which each non-migrant will be assigned a probability as weight

---

[28] Available in STATA under "DFL" syntax, standing for Dinardo, Fortin and Lemieux model.

before applying migrants' wages to their non-migrant counterparts. What follows presents more details of the empirical model.

5.3.1 Econometric Model

Despite a simple concept, counterfactual computations are complex. Skipping algebra (to be found in DFL), the following model needs to be considered. Equation 5-1 illustrates a formula that is intended to reflect the counterfactual wage dispersion;

$$CF(W) = \int h(W^{migrant} \mid X)\, \emptyset\, g\left(X^{non-migrants}\right) dX \qquad (5-1)$$

$$such\ that \qquad \emptyset = \frac{1 - p(i = migrant \mid X)}{p(i = migrant \mid X)}$$

where $CF(W)$ represents the counterfactual wage distribution, $h\left(W^{migrant}\right)$ captures the observed wage structure of inter-provincial migrants conditional on observed characteristics X, $\emptyset$ is the weight variable that is going to be estimated using a logit model, $g\left(X^{non-migrant}\right)$ is the distribution of observed characteristics of non-migrant workers, and $p(i = migrant \mid X)$ is the probability of being a migrant for an individual i conditional on the observable characteristics X.

To obtain the counterfactual estimate of the wage distribution, it is necessary to estimate the probability values of the weight variable $\emptyset$ in the first place. As a result, a simple logit regression of migration decisions on a number of factors will be conducted to extract the estimated odds ratios of being a migrant. The logit model[29] is presented below using the independent variables that have been found important in the literature (Brown et al., 2012; Coulombe, 2006; Robinson and Tomes, 1980).

$$m_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 hdgree_i + \beta_4 sex_i + \beta_5 married_i + \beta_6 unemp_i + \beta_7 kid_i + \beta_8 prairies_i + e_i$$

where m_i is a migration dummy variable for individual i, age is the age of individual, $hdgree$ is the highest education acquired that has 13 categories, sex is a gender dummy variable, $married$ is a dummy variable that equals one if the individual is married or living common-law, $unemp$ reflects the unemployment rate of the province that individual i lives in, kid is a dummy variable that equals one if individual i has at least one child of age 14 years old or below[30], $prairies$ is a dummy variable if the original province of the migrant is Manitoba or Saskatchewan, and finally e_i is the error term.

Lastly, the following formula is applied to compute the counterfactual Theil's values using the estimated wage densities (DiNardo et al., 1996):

$$T = \int [\ln \hat{f}(v)]\hat{f}(v)dv \quad ; \quad \hat{f}(v) = \frac{\hat{f}(w)}{v} \quad and \quad v = \exp(w)$$

where $\hat{f}(w)$ is the counterfactual wage distribution.

---

[29] Estimated results for logit model are available in Appendix IV

[30] PUMF provides information on the number of children by age category, i.e. 0 to 1 years, 2 to 5 years, 6 to 14 years, 15 to 24 years, etc. In this study, only the first 3 categories will be considered to proxy the presence of children.

5.3.2 Empirical Results

This section provides a visual and numerical presentation of empirical findings for the observed wage structure of migrants and non-migrants, as well as the counterfactual wage distribution in the absence of interprovincial migration[31]. All the findings are presented in one single diagram, Figure 5-4, to make it easier for wage comparisons across groups of workers.

The first graph in Figure 5-4 illustrates the density of real wage structures for observed non-migrants (dashed blue), observed migrants (in solid red) as well as the counterfactual wage distribution (i.e. the wage distribution in the absence of interprovincial migration (the solid green line labelled as CF_Non_migrant)) with corresponding Theil's values attached. The second graph illustrates the density differences between non-migrant factual (dashed blue curve) and counterfactual wage densities (solid green curve)[32]. Technically, this second graph maps the effect of migration exerted on different segments of the wage distribution.

[31] Please note that the counterfactual wage densities of non-migrants' are interpreted as overall wage structure in the absence of migration.
[32] Differences are calculated according to factual minus counterfactual of non-migrants.

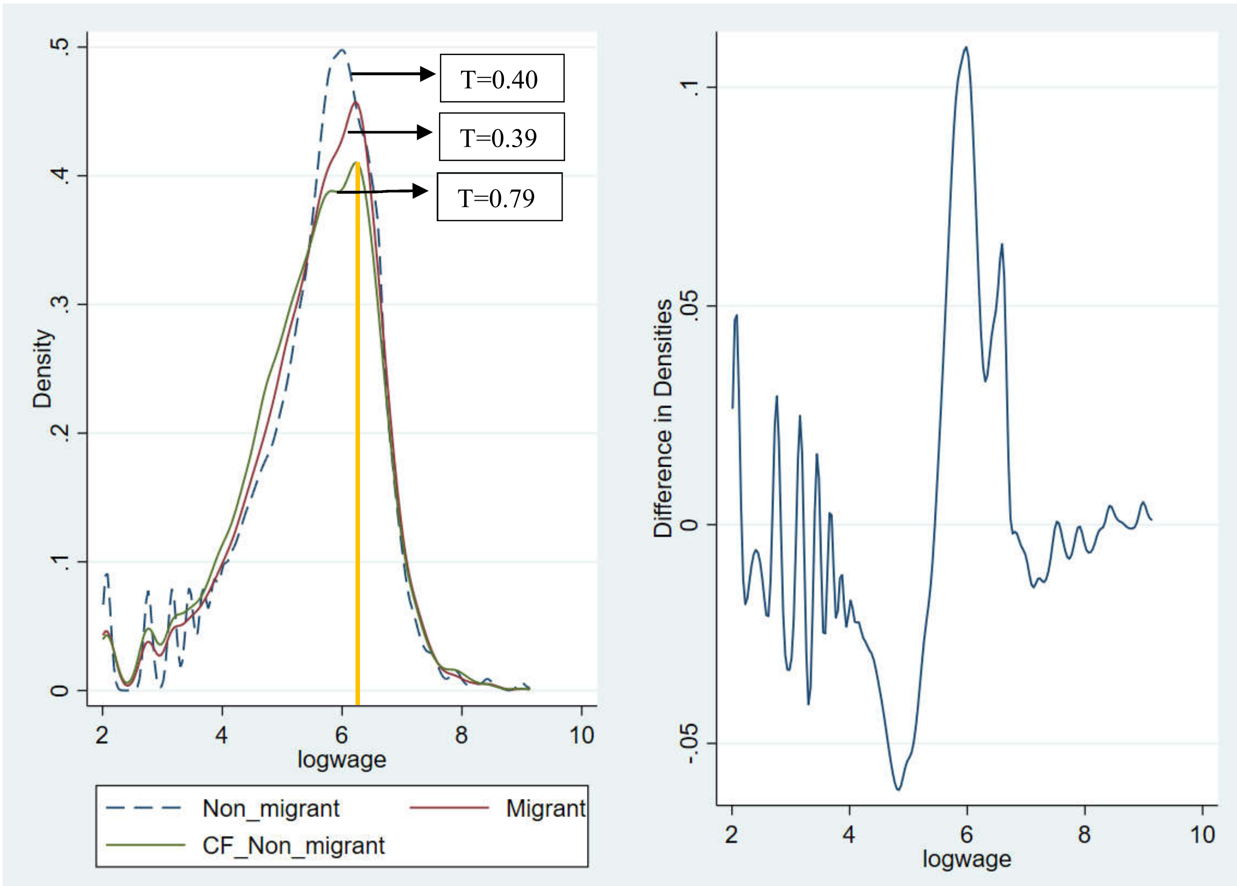

Figure 5-4. Factual and counterfactual wage densities with corresponding Theil's values (all observations)

Starting with Theil's findings, the results show that almost all wage inequality in Canada can be attributed to within-group wage differences (i.e. within migrants, and within non-migrants). Between-group component accounts for less than 1% of wage inequality. That is, although the population of interprovincial migrants makes up only 3% of the sample, wage inequality within this group (T=0.39) is almost as large as that for non-migrants (T=0.4). More importantly, counterfactual wage inequality (i.e. T=0.79) is substantially larger than the other Theil's values. In other words, wage inequality would have been twice as large in the absence of internal migration.

Although migration reduces wage inequality (as shown by Theil's values), it is worthwhile to add that the effect is not uniform across the wage distribution. This finding can be best presented with the use of wage density curves. In particular, the second graph maps clearly the differing impacts of interprovincial migration exerted on different segments of the wage distribution.

As evident, the migration effect is mixed and present in almost the entire wage distribution. Beginning from the lower end of the distribution, it shows that migration has a volatile impact on the lower end with both increasing and decreasing densities. In the mid-lower portion of the distribution, density sharply declines, whereas it largely increases in the middle, as well as in the low-upper portion. The effect, however, gradually fades away in the upper portion, in particular in the higher end.

It is also worthwhile to discuss the possible explanations for such differing impacts, although further research is required to provide a more accurate answer. First of all, with self-selective migration, it is reasonable to observe that density increases in the middle of wage distribution because migrants have moved more likely for wage opportunities. However, this is not a full explanation since different segments of wage density are affected unequally. In this case, other indirect effects of migration need to be considered for clarification. For instance, migration increases the demand for local goods and services (e.g. physicians, hair-dressers, kindergarten professions and so on) in the destination province, which has a positive impact on wage structure in that province. At the same time, offsetting impact may be expected if migrants displace existing workers or compress their wages because of an increase in labour supply in destination province.

In the context of this study, it is equally important to also think about the source provinces because the empirical analysis (conducted in section 5.3) does not distinguish between out-migration and in-migration. In this case, any worker who is leaving a province would not only impact the wage structure in the destination province but also in the source province. A first mechanism that can be relevant is a decrease in labour supply in the source province causing wage increases in that province. Another mechanism could be long-run job losses in the source provinces due to dis-agglomeration or human capital losses of persistently negative net migration of workers, particularly skilled workers, as can be understood from the literature above.

5.3.3 Migration Effects Across Skill Groups

It is also interesting to further discuss the implications that internal migration has for the wage structure of specific groups of individuals. Such groups of interest could be defined by education to proxy for skilled and unskilled workers (Card and DiNardo, 2002). Therefore skilled workers are defined as those with a bachelor's degree or above, whereas the rest of the workers are considered unskilled. Findings are presented in Figure 5-5 and Figure 5-6, which illustrate wage densities with corresponding T-values for skilled workers, and unskilled workers, respectively.

The effect of internal migration on skilled workers can be seen in almost the entire wage distribution by looking at the second graph of Figure 5-5. Migration increases the mass of skilled workers in the lower end of the wage distribution, although the magnitude of the effect is notably small relative to other segments. As we move to higher levels, migration reduces the density up to the middle of the distribution. In the middle, however, migration has a positive impact with the largest effect in magnitude compared to other segments of the distribution. The effect sharply declines (in magnitude) at the higher portion of the wage distribution, and gradually disappears as we move towards the higher end. And lastly, it seems that wage inequality among skilled workers would have been larger by almost 40% (0.61 vs 0.41) if there was no internal migration.

As for unskilled workers, a similar pattern of migration impact is observed in Figure 5-6 (second graph), although the magnitude of the effects is quite different from what was observed for skilled workers. To make a comparison between the impacts on skilled and unskilled workers, the migration effect seems to be larger for unskilled workers at the lower end of the wage distribution. In the middle of the unskilled wage distribution, however, the magnitude of the effect is smaller. It is because the mass of skilled workers increases by 0.2 in the middle of their wage dispersion, whereas the mass of unskilled workers increases by only 0.16. It is also interesting to realize that Theil's index would have been substantially larger for unskilled workers if there was no migration (0.88 in the counterfactual case vs 0.34 in the factual case).

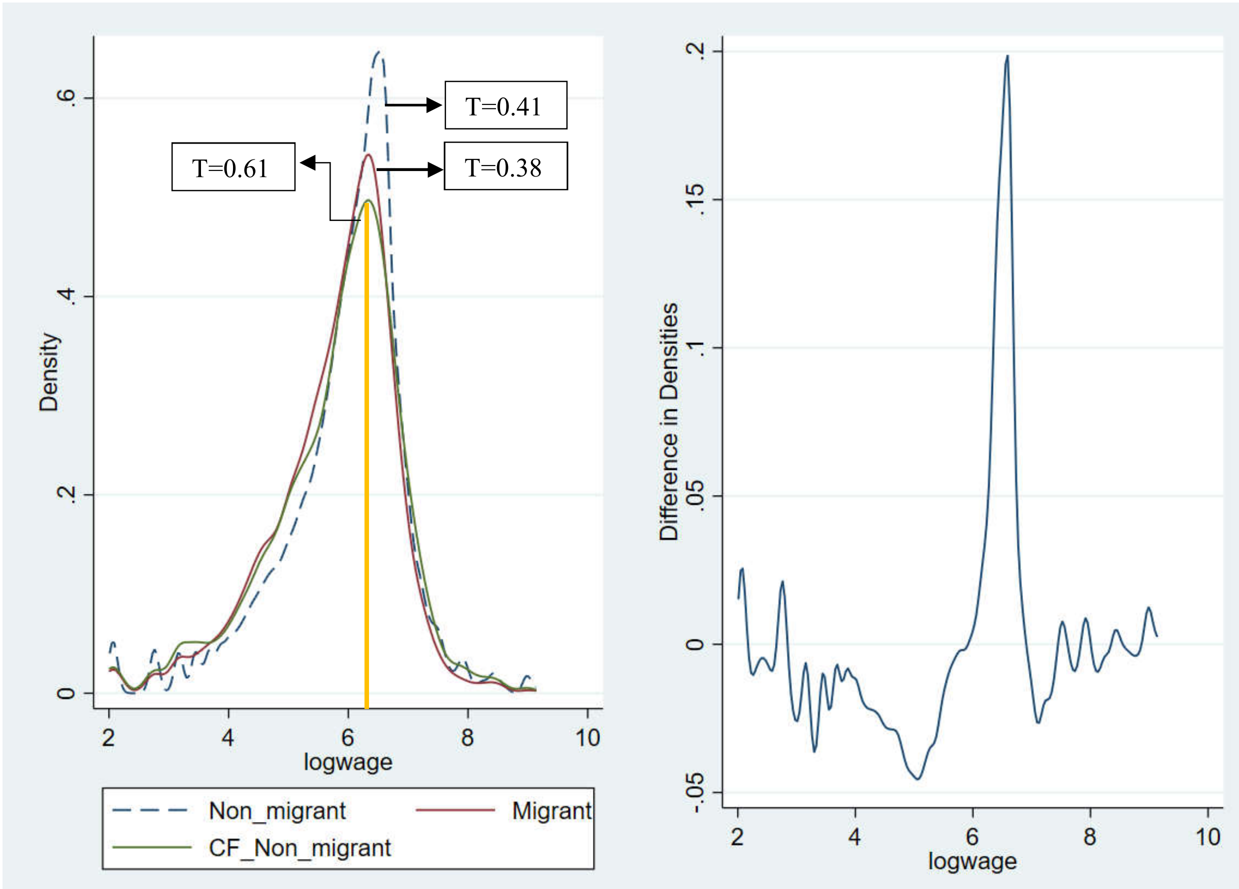

Figure 5-5. Skilled workers wage densities with corresponding Theil's values

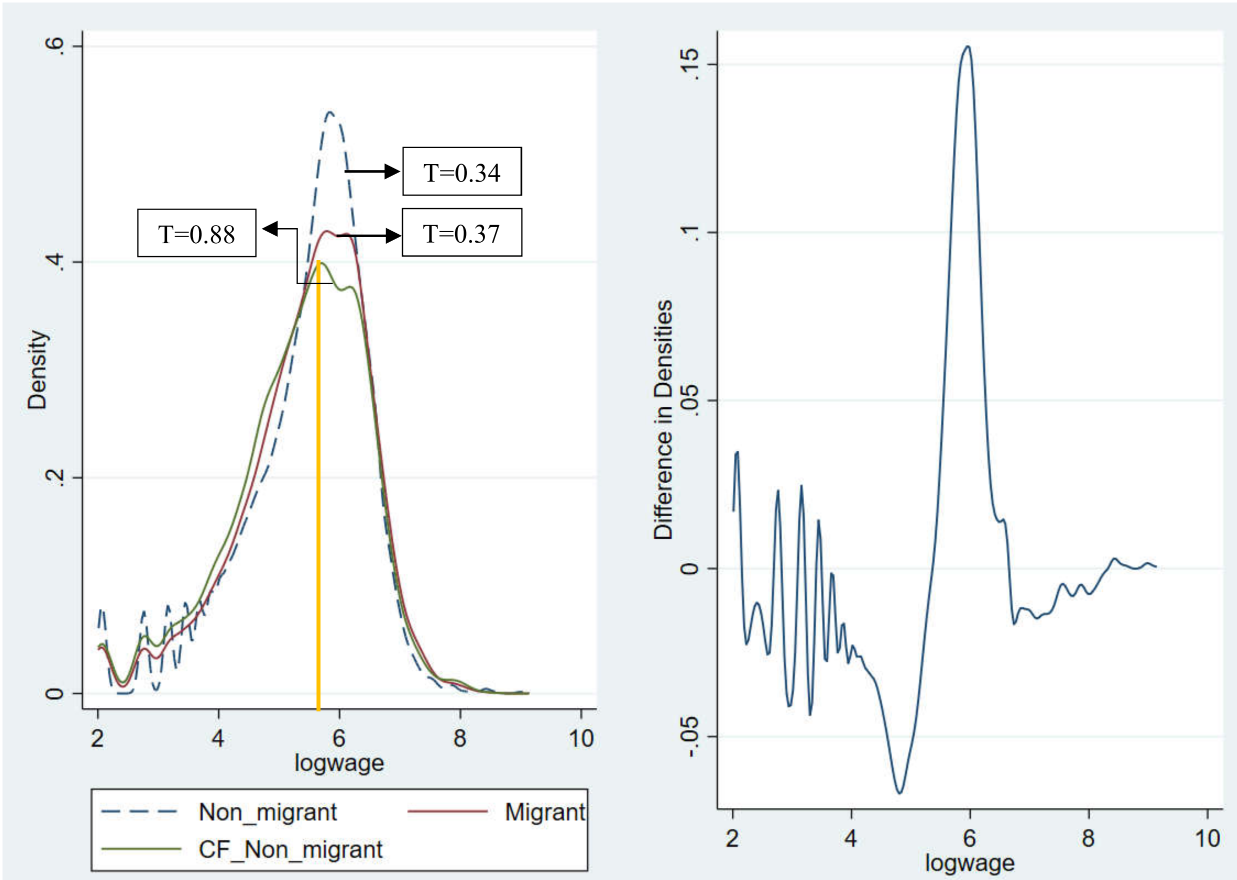

Figure 5-6. Unskilled workers wage structures with corresponding Theil's values

5.4 Summary and Conclusion

This is the first study, to the best of my knowledge, that attempts to assess the impact of interprovincial migration on wage distribution in Canada. Using the 2016 Census PUMF, a counterfactual model is adopted that enables the researcher to learn about wage structure in the absence of internal migration.

In summary, the findings show that internal migration cuts wage inequality (measured by Theil's index) by almost half. It also changes the appearance of wage density (represented by Kernel wage density). Looking at the lower end of the distribution for all observations reveals that internal migration has volatile impacts (i.e. both increasing and decreasing effects), although overall, it seems that these effects are somewhat offsetting each other as the absolute value of changes in the densities are more or less equal. Moving from the lower end toward higher wages, mass sharply decreases in the mid-lower portion and substantially rises in the middle (i.e. some degree of polarizing effect is detected). Internal migration, however, seems to have less effect on the upper range of wages such that it almost fades away in the very high end of the distribution.

To present the results in numeric fashion, corresponding Theil's values are also computed for all the wage densities (factual wage density, and counterfactual). The inequality effect of interprovincial migration is substantially large. The results show that the magnitude of Theil's value would have been twice as large in the absence of migration (i.e. 0.79 instead of 0.4, as shown in Figure 5-4).

Furthermore, it is found that interprovincial migration is an important determinant of wage distribution across skill groups (measured by education). In the absence of migration, wider wage density curves should be expected for both skilled and unskilled workers. According to estimated counterfactual Theil's value, wage inequality would have been 1.5 and 2.5 times larger for skilled and unskilled workers, respectively.

A limitation of this study, however, is that it cannot provide a clear explanation as to why such volatility is observed in the lower end of the wage distribution. Possible explanations are provided in sub-section 5.3.2, but further research is required to address this challenge.

Chapter 6

Overall Conclusion

Essay Summaries:

Dealing with a challenging and controversial field of study, this project is an empirical exercise to discover and test various potential determinants of wage inequality for Canada at the industry level in Chapter 3 and CMA level in Chapter 4. Chapter 5 examines the effect of interprovincial migration on wage inequality.

Considering two-digit industries as the unit of study in Chapter 3 (i.e. twenty industrial NAICS 2007 groups), a two-round empirical analysis is conducted in which wage inequalities are measured by Theil's values for industries in the first step. The decomposition results demonstrate that industry differentials account for about 10% of Canadian wage inequality in each census year (1996, 2001, 2006, 2011, and 2016), while the rest of inequality is attributed to variation within industries.

In the second round, a pooled OLS regression analysis is conducted to generate estimates for wage inequality within the industries to explore the determinants of inequality variations within industry groups. The dependent variable is the log of the within-industry Theil's index from the first round.

In this regression model, Lerner's index (to represent industry monopoly power) is a variable of interest examined for the first time in the Canadian literature. Lerner's index is found to be an important and significant variable that affects wage distribution within industries. The estimated coefficient is also robust to model specification. It turns out that the effect of this index is negative (one-point increase in the Lerner's index is associated with a 0.19% decrease in within-industry inequality). In this case, an industry with larger monopoly power is predicted to realize a lower level of inequality, a finding that is consistent with the rent-sharing theory in the monopoly literature.

All the other variables in the regression model (such as union share, export and import exposures, innovation index, the growth rate of the industry, job status as well as demographic variables) are also included as control variables, following the literature. Their corresponding estimated coefficients are more or less consistent with what the literature generally expects and finds. More details are available in the empirical section of Chapter 3.

The second empirical essay (Chapter 4) is also a two-round analysis, but from a different angle, i.e. within and between wage inequality for CMAs rather than for industries. In the first round, Theil's decomposition results show that almost all wage inequality observed across CMAs is attributable to inequality within CMAs (accounting for over 97% of overall wage inequality across all CMAs) rather than between them (contributing by 3% or less).

The generated Theil's values for CMAs in the first round are used as the dependent variable in the second round to investigate potential determinants of within-CMA inequalities in an

OLS framework. Of all the independent variables, two are tested for the first time in the literature, namely average city rent (ACR), and regional growth clusters (RGCs).

ACR (a measure of the industrial composition of CMAs in terms of high-pay industries and low-pay industries) negatively impacts the within-CMA inequality. That is, CMAs with a 1% larger composition of high-pay jobs would experience a lower level of inequality (by 0.14%), assuming other factors constant.

Regional Growth Clusters (RGCs), a variable to capture the effect of the growth rate of major industrial clusters of a CMA, is also found to be an important determinant of within-CMA inequality. It is shown that a faster rate of growth of industrial clusters tends to increase the earnings inequality by 0.51%. Several other control variables are also included in the regression equation, and the estimated coefficients are consistent with previous studies. More details on estimation results are available in Table 4-2.

Turning attention to the effect of interprovincial (internal) labour mobility on wage inequality in Canada, the third essay (Chapter 5) attempts to discover any potential impact that it may have on the Canadian wage structure. To do so, a semiparametric counterfactual model is considered in which Kernel wage densities, and the Theil's index are estimated in the absence of migration. Using the 2016 PUMF, the results show that wage inequality would have been almost twice larger in the absence of migration (counterfactual Theil=0.79 vs factual Theil=0.4), a finding consistent with self-selection theory that predicts a reduction in within-wage inequalities due to internal migration.

In addition, the estimated wage density shows that internal migration imposes a larger impact in the lower portion of the wage distribution than in the upper portion because the effect begins to disappear at the higher end of the distribution. In the lower portion, some degree of polarization is observed.

Additional numerical results are found to be interesting when exploring the effect of interprovincial migration on skill groups (measured by education). In the absence of internal migration, both skilled and unskilled workers would have experienced larger wage inequality, i.e. 48% and 158% larger, respectively. Looking at wage densities, the migration effect seems to be larger for unskilled workers in the lower end of the wage distribution than it is for skilled workers. In the middle of the unskilled wage distribution, however, the magnitude of the effect is relatively smaller. In the upper portion of wage distribution, the effect fades away for both groups.

Contributions and Implications:

As a fairly comprehensive examination of wage inequality, this project attempts to make a contribution to the literature by testing new determinants of wage inequality in Canada. Although the goal was not to offer specific policy suggestions, this examination finds that a variety of policies will influence wage inequality.

At the industry level (e.g. aviation sector or oil sector), easing merger and acquisition regulations can increase the relative monopoly power of the industry and, therefore, somewhat reduce the wage inequality within that industry. At the CMA level, targeted policies could be considered for regional growth clusters (RGCs). As described in Chapter 4, the growth of each CMA's economy heavily relies on these RGCs. An effective policy toward reducing inequality could be regulations that could provide incentives for these RGCs to spend more resources on R&D and computer expenditures. And finally, the results of the third essay imply that policies could be considered for facilitating interprovincial migration opportunities as it has a large mitigating impact on wage inequality.

Future research can also benefit from this study. One interesting research question concerns how wage inequality responds to economic booms and busts, including the recent economic shocks due to the COVID-19 epidemic. Although it requires further research, such shock impacts inequality through various sources.

On the one hand, with recently massive layoffs, wage inequality is expected to rise because the effect of unemployment on within-CMA inequality is positive, as presented in Table 4-2. The estimated coefficient for part-time share is also positive, meaning that an increase in the share of part-time workers would result in an increase in wage inequality. On the other hand, a slower growth rate in CMA's RGCs is expected to mitigate within-CMA inequality because the estimated coefficient for RGC is positive and fairly large. Therefore, the final effect could go either way in the coming years, depending on the magnitude of changes in the determinants of wage inequality.

Pitfalls:

Having said that, it is equally important to mention that this study faces a number of limitations and pitfalls. The first limitation is that the Lerner's index of Chapter 3 is estimated under a lack of readily available data. Lerner's values are computed using aggregate data (multifactor productivity data) at two-digit industrial groups, whereas a more accurate way would use plant-level data, which is hardly available.

In Chapters 3 and 4, several data sources have been combined to conduct the regression analysis. There can be found some degree of inconsistency across data sources, or even within a specific data source (e.g. definition of R&D has changed over time, or industrial groups from input-output tables (for constructing trade variables and R&D) differ from that in the Census). Such inconsistency would cause noise in the estimates.

The estimated coefficient for innovation is not robust, as found in Chapters 3 and 4, whereas previous studies have found significant evidence for it. This limitation arises due to the lack of readily available data. The R&D definition of innovation has been considered in this study using input-output tables. This definition is suitable for the industrial aspect of inequality, whereas the output definition of R&D using a patent dataset might be more appropriate when the urban aspect of inequality is under investigation.

And lastly, it is found in Chapter 5 that internal migration is an important determinant of wage inequality with differing impacts on certain ranges of the wage density. In particular, in the lower end, the volatile impact is observed. However, a limitation of this study is that it is not able to offer a clear explanation as to why such volatility is observed in the lower end of the wage distribution. Possible explanations are provided in sub-section 5.3.2, but further research is required to address this challenge.

Bibliography

Acemoglu, D. (1999). Changes in Unemployment and Wage Inequality-An Alternative Theory and Some Evidence. The American Economic Review *89*, 1259–1278.

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market (National Bureau of Economic Research).

Akita, T. (2003). Decomposing Regional Income Inequality in China and Indonesia Using Two-Stage nested Theil Decomposition Method. The Annals of Regional Science *37*, 55–77.

Andersson, M., Klaesson, J., and Larsson, J.P. (2014). The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies?: The sources of the urban wage premium by worker skills. Papers in Regional Science *93*, 727–747.

Baum-Snow, N., and Pavan, R. (2013). Inequality and city size. Review of Economics and Statistics *95*, 1535–1548.

Beach, C.M., Finnie, R., Gray, D., Statistics Canada, and Business and Labour Market Analysis Division (2006). Summary of: The impact of macroeconomic conditions on the instability and long-run inequality of worker's earnings in Canada (Ottawa, Ont.: Statistics Canada).

Beaudry, P., Green, D.A., and Sand, B.M. (2012). Does Industrial Composition Matter for Wages? A Test of Search and Matching Theory. Econometrica *80*, 1063–1104.

Belman, D. (1988). Concentration, Unionism, and Labor Earnings: A Sample Selection Approach. The Review of Economics and Statistics *70*, 391.

Berman, E., Bound, J., and Griliches, Z. (1994). Changes in the Demand for Skilled Labor in US Manufacturing: Evidence from the Annual Survey of Manufacturers. The Quarterly Journal of Economics *109*, 367–397.

Blau, F.D., and Kahn, L.M. (1994). Rising Wage Inequality and the U.S. Gender Gap. The American Economic Review *84*, 23–28.

Blau, F.D., and Kahn, L.M. (2000). Gender differences in pay (National bureau of economic research).

Block, S., Galabuzi, G.-E., and Tranjan, R. (2019). Canada's Colour Coded Income Inequality. Canadian Center for Public Alternatives.

Bolton, K., and Breau, S. (2012). Growing Unequal? Changes in the Distribution of Earnings across Canadian Cities. Urban Studies *49*, 1377–1396.

Borjas, G.J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. The Quarterly Journal of Economics *118*, 1335–1374.

Borjas, G.J., and Ramey, V.A. (1995). Foreign Competition, Market Power, and Wage Inequality. The Quarterly Journal of Economics *110*, 1075–1110.

Borjas, G.J., Freeman, R.B., Katz, L.F., DiNardo, J., and Abowd, J.M. (1997). How Much Do Immigration and Trade Affect Labor Market Outcomes? Brookings Papers on Economic Activity *1997*, 1–90.

Boudarbat, B., Lemieux, T., and Riddell, W.C. (2014). The evolution of the Returns to Human Capital in Canada, 1980-2006. Canadian Labour Market and Skills Researcher Network

Bound, J., and Johnson, G. (1992). Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations. The American Economic Review *82*, 371–392.

Bourguignon, F. (1981). Pareto Superiority of Unegalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function. Econometrica *49*, 1469.

Breau, S., and Rigby, D.L. (2010). International Trade and Wage Inequality in Canada. Journal of Economic Geography *10*, 55–86.

Breau, S., Kogler, D.F., and Bolton, K.C. (2014). On the Relationship between Innovation and Wage Inequality: New Evidence from Canadian Cities: Innovation and Wage Inequality. Economic Geography *90*, 351–373.

Brown, W.M., Newbold, K.B., Statistics Canada, and Economic Analysis Division (2012). Cities and growth: moving to Toronto : income gains associated with large metropolitan labour markets (Ottawa, Ont.: Statistics Canada, Economic Analysis Division).

Card, D. (2009). Immigration and Inequality. NBER Working Papper Series 14683.

Card, D., and DiNardo, J.E. (2002). Skill biased technological change and rising wage inequality: some problems and puzzles (National Bureau of Economic Research).

Card, D.E., and Freeman, R.B. (1993). Small Differences That Matter: Labor Markets and Income Maintenance in Canada and the United States (Chicago: University of Chicago Press).

Card, D., Lemieux, T., and Riddell, W.C. (2004). Unions and wage inequality. Journal of Labor Research *25*, 519–559.

Card, D., Cardoso, A.R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. Journal of Labor Economics *36*, S13–S70.

Chakravorty, S. (1996). Urban inequality revisited: the determinants of income distribution. Urban Affairs Review *31*, 759–777.

Checchi, D., García-Peñalosa, C., and Vivian, L. (2016). Are changes in the dispersion of hours worked a cause of increased earnings inequality? IZA Journal of European Labor Studies *5*, 15.

Cingano, F. (2014). Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment, and Migration Working Papers. No. 163. OECD Publishing.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! Journal of Urban Economics *63*, 723–742.

Coulombe, S. (2006). Internal Migration, Asymmetric Shocks, and Interprovincial Economic Adjustments in Canada. International Regional Science Review *29*, 199–223.

DiNardo, J., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages 1973-1992: A Semiparametric Approach.

DiNardo, J., Fortin, N.M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. Econometrica *64*, 1001.

Dunne, T., Foster, L., Haltiwanger, J., and Troske, K.R. (2004). Wage and Productivity Dispersion in US Manufacturing: The Role of Computer Investment. Journal of Labor Economics *22*, 397–429.

Esther Mirjam, G., Romuald, M., and Hillel, R. (2019). Regional Migration and Wage Inequality in the West African Economic and Monetary Union.

Farber, H.S., and Western, B. (2001). Accounting for the Decline of Unions in the Private Sector, 1973–1998. Journal of Labor Research *22*, 459–485.

Finnie, R. (2001). Unequal Inequality: the Distribution of Individuals' Earnings by Province. Canadian Journal of Regional Science *24*.

Florida, R., and Mellander, C. (2016). The Geography of Inequality: Difference and Determinants of Wage and Income Inequality across US Metros. Regional Studies *50*, 79–92.

Foley, K., and Green, D.A. (2016). Why more education will not solve rising inequality (and may make it worse). Income Inequality: The Canadian Story. Montreal: Institute for Research on Public Policy.

Fong, F. (2017). Income Inequality in Canada - The Urban Gap. Chartered Professional Accountants Canada.

Fortin, N.M., and Lemieux, T. (2014). Changes in Wage Inequality in Canada: An Interprovincial Perspective. Working Paper.

Furman, J., and Orszag, P. (2015). A Firm-Level Perspective on the Role of Rents in the Rise in Inequality. Presentation at "A Just Society" Centennial Event in Honor of Joseph Stiglitz Columbia University.

Galbraith, J.K., and Hale, J.T. (2014). The Evolution of Economic Inequality in the United States, 1969–2012: Evidence from Data on Inter-Industrial Earnings and Inter-Regional Incomes. World Economic Review *3*, 1–19.

Galor, O. (2009). Inequality and economic development: An overview (Working Paper).

Gera, S., and Grenier, G. (1994). Interindustry Wage Differentials and Efficiency Wages: Some Canadian Evidence. The Canadian Journal of Economics *27*, 81.

Glaeser, E.L., and Maré, D.C. (2001). Cities and Skills. Journal of Labor Economics *19*, 316–342.

Gray, D., Mills, J.A., and Zandvakili, S. (2003). Statistical analysis of inequality with decompositions: the Canadian experience. Empirical Economics *28*, 291–302.

Green, D.A., and Sand, B.M. (2015). Has the Canadian labour market polarized? Canadian Journal of Economics/Revue Canadienne D'économique *48*, 612–646.

Hall, B. (2014). R&D, Innovation, Productivity,(and Growth). In Presentation at International Conference on Total Factor Productivity at the National Bank of Belgium, Brussels.

Hall, R.E. (2018). Using Empirical Marginal Cost to Measure Market Power in The US Economy (National Bureau of Economic Research).

Harrison, A., McLaren, J., and McMillan, M. (2011). Recent Perspectives on Trade and Inequality. Annual Review of Economics *3*, 261–289.

Haworth, C.T., Long, J.E., and Rasmussen, D.W. (1978). Income distribution, city size, and urban growth. Urban Studies *15*, 1–7.

Heckman, J.J., and Honore, B.E. (1990). The empirical content of the Roy model. Econometrica: Journal of the Econometric Society 1121–1149.

Heisz, A. (2016). Trends in income inequality in Canada and elsewhere. Income Inequality: The Canadian Story. Montreal: Institute for Research on Public Policy.

Johnson, S., and Kuhn, P. (2004). Increasing Male Earnings Inequality in Canada and the United States, 1981-1997: The Role of Hours Changes versus Wage Changes. Canadian Public Policy / Analyse de Politiques *30*, 155.

Juhn, C., Murphy, K.M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. Journal of Political Economy 410–442.

Kaldor, N. (1955). Alternative Theories of Distribution. The Review of Economic Studies *23*, 83.

Kanbur, R., and Rapoport, H. (2005). Migration selectivity and the evolution of spatial. Journal of Economic Geography *5*.

Katz, L.F., and Murphy, K.M. (1992). Changes in Relative Wages, 1963-1987 Supply and Demand Factors. The Quarterly Journal of Economics *107*, 35–78.

Korpi, M. (2007). Does size of local labour markets affect wage inequality? a rank-size rule of income distribution. Journal of Economic Geography *8*, 211–237.

Kremer, M., and Maskin, E. (1996). Wage Inequality and Segregation by Skill (National Bureau of Economic Research).

Krugman, P.R. (2008). Trade and Wages, Reconsidered.

Lazear, E.P., and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. Journal of Political Economy *89*, 841–864.

Lee, D.S. (1999). Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage? The Quarterly Journal of Economics *114*, 977–1023.

Legree, S., Schirle, T., and Skuterud, M. (2016). Can Labour Relations Reform Reduce Wage Inequality? Income Inequality: The Canadian Story. Montreal: Institute for Research on Public Policy *V*, 325–360.

Lemieux, T. (2006). Post-secondary education and increasing wage inequality (National Bureau of Economic Research).

Marchand, J. (2012). Local labor market impacts of energy boom-bust-boom in Western Canada. Journal of Urban Economics *71*, 165–174.

Marchand, J. (2015). The distributional impacts of an energy boom in Western Canada. Canadian Journal of Economics/Revue Canadienne D'économique *48*, 714–735.

Matano, A., and Naticchioni, P. (2011). Matano 2011 What drives the urban wage premium.pdf.

Mion, G., and Naticchioni, P. (2009). Spatial Sorting and Matching of Skills and Firms. Canadian Journal of Economics/Revue Canadienne D'économique *42*.

Mirrlees, J.A. (1971). An Exploration in the Theory of Optimum Income Taxation. The Review of Economic Studies *38*, 175.

Mocan, H.N. (1999). Structural unemployment, cyclical unemployment, and income inequality. Review of Economics and Statistics *81*, 122–134.

Moore, E.G., and Pacey, M.A. (2003). Changing Income Inequality and Immigration in Canada, 1980-1995. Canadian Public Policy / Analyse de Politiques *29*, 33.

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of Econometrics *121*, 175–212.

Moretti, E. (2013). Real Wage Inequality. American Economic Journal: Applied Economics *5*, 65–103.

Moretti, E., and Thulin, P. (2012). Local multipliers and human capital in the US and Sweden. Research Institute of Industrial Economics.

Moyser, M. (2017). Women in Canada: A Gender-based Statistical Report. Statistics Canada. Catalogue no. 89-503-X

Mueller, R.E. (1998). Public–Private Sector Wage Differentials in Canada: Evidence from Quantile Regressions. Economics Letters *60*, 229–235.

OECD (2011). An overview of growing income inequalities in OECD countries: main findings.

Osberg, L. (2018). The age of increasing inequality: the astonishing rise of Canada's 1% (Toronto : James Lorimer & Company Ltd., Publishers).

Palacios, M., Lammam, C., Ren, F., and Clemens, J. (2016). Comparing Government and Private Sector Compensation in Canada. Journal of Self-Governance & Management Economics *4*.

Pavcnik, N. (2011). Globalization and Within-Country Income Inequality. Making Globalization Socially Sustainable 233.

Peters, J. (2014). Neoliberalism, Inequality and Austerity in the Rich World Democracies. Baines, Donna and Stephen Mcbride (Eds).

Phan, D., and Coxhead, I. (2010). Inter-provincial migration and inequality during Vietnam's transition. Journal of Development Economics *91*, 100–112.

Picot, G., and Hou, F. (2016). Immigration Poverty and Income Inequality in Canada. Income Inequality: The Canadian Story. Montreal: Institute for Research on Public Policy *5*, 249–286.

Robinson, C., and Tomes, N. (1980). Self-selection and Interprovincial Migration in Canada. Department of Economics Research Reports, 8019. London, ON: Department of Economics, University of Western Ontario

Roy, A.D. (1951). Some Thoughts On the Distribution of Earnings. Oxford Economics Papers *3*, 135–146.

Saez, E., and Veall, M.R. (2005). The Evolution of High Incomes in Northern America_Lessons from Canadian Evidence. The American Economic Review *95*, 831–849.

Saunders, D. (2018). Internal Migration: Overview, 2015/2016. Statistics Canada. Catalogue no. 91-209-X

Sharpe, A., Arsenault, J.-F., and Ershov, D. (2007). The Impact of Interprovincial Migration on Aggregate Output and Labour Productivity in Canada, 1987-2006 (Center for the Study of Living Standards (CSLS)).

Shorrocks, A., and Wan, G. (2005). Spatial Decomposition of Inequality. Journal of Economic Geography *5*, 59–81.

Simpson, W. (1986). Analysis of Part-Time Pay in Canada. The Canadian Journal of Economics *19*, 798.

Slaper, T.F., Harmon, K.M., and Rubin, B.M. (2018). Industry clusters and regional economic performance: A study across US metropolitan statistical areas. Economic Development Quarterly *32*, 44–59.

Song, J., and Price, D.J. (2015). Firning Up Inequality. Centre for Economic Performance.

Walks, A., Twigge-Molecey, A., University of Toronto, and Cities Centre (2014). Income inequality and polarization in Canada's cities: an examination and new form of measurement.

Warman, C., and Worswick, C. (2016). Technological Change and Declining Immigrants' Earnings Outcomes: Implications for Income Inequality in Canada. Income Inequality: The Canadian Story. Montreal: Institute for Research on Public Policy *5*, 287–297.

Wolfson, M.C., and Murphy, B.B. (1998). New Views on Inequality Trends in Canada and the United States. Monthly Lab. Rev. *121*, 3.

Yalnizyan, A. (2011). The rise of Canada's richest 1% (Ottawa, Ont.: Canadian Centre for Policy Alternatives).

(2008). Growing unequal? income distribution and poverty in OECD countries (Paris: OECD, Organisation for Economic Co-operation and Development).

Appendix

Appendix I

Scatter plots of within-industry T values (in the log) in the vertical axis, with other control variables in the horizontal axis:

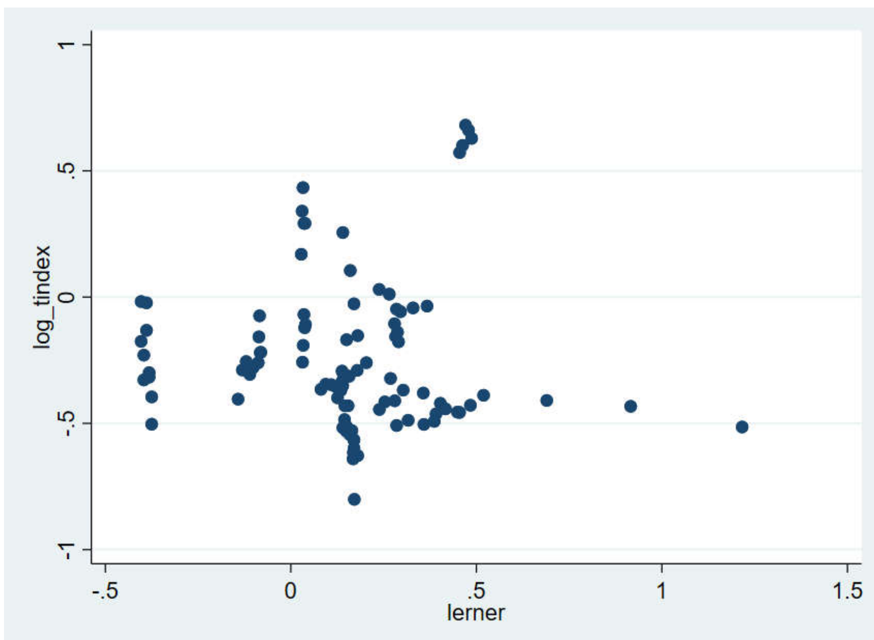

Figure A-1: Industry T index vs sectoral Lerner's index

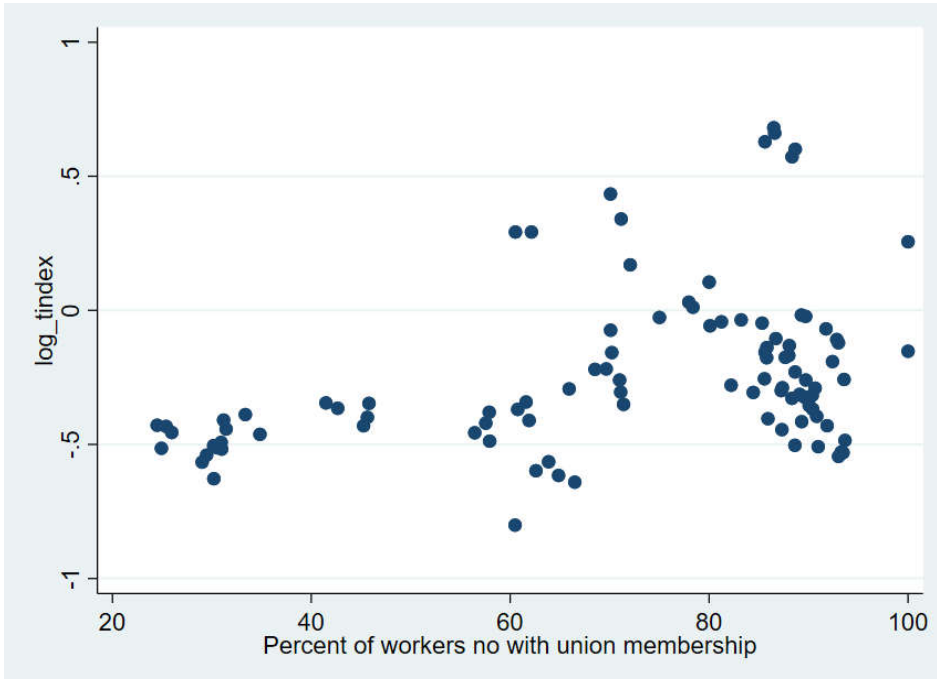

Figure A-2: Industry T index vs industry union rates

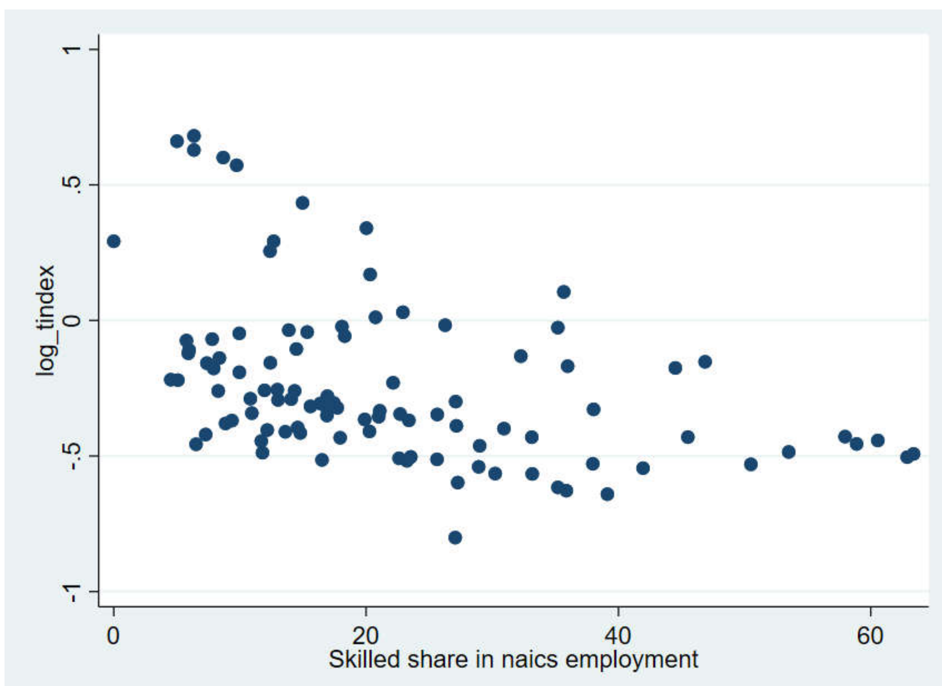

Figure A-3: Industry T index vs share of workers with bachelor's degree or above in industry employment

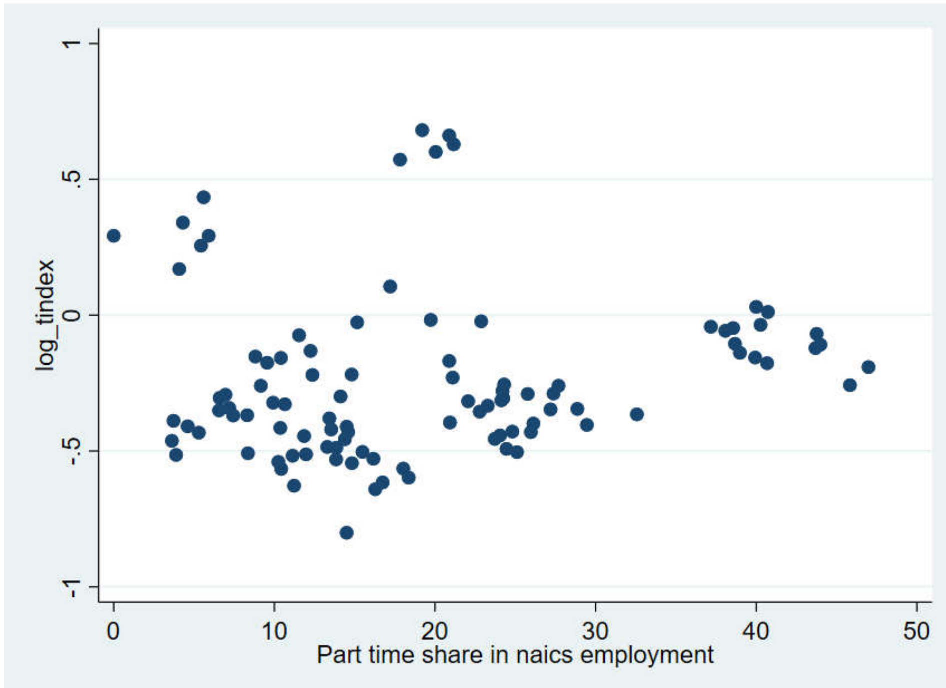

Figure A-4: Industry T index vs share of part-time workers in industry employment

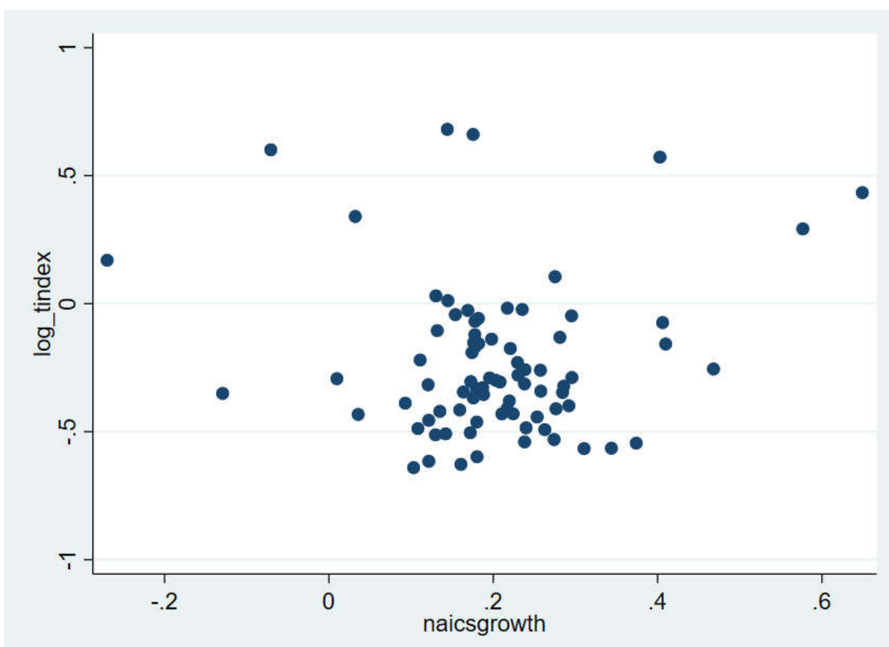

Figure A-5: Industry T index vs industry rate of growth

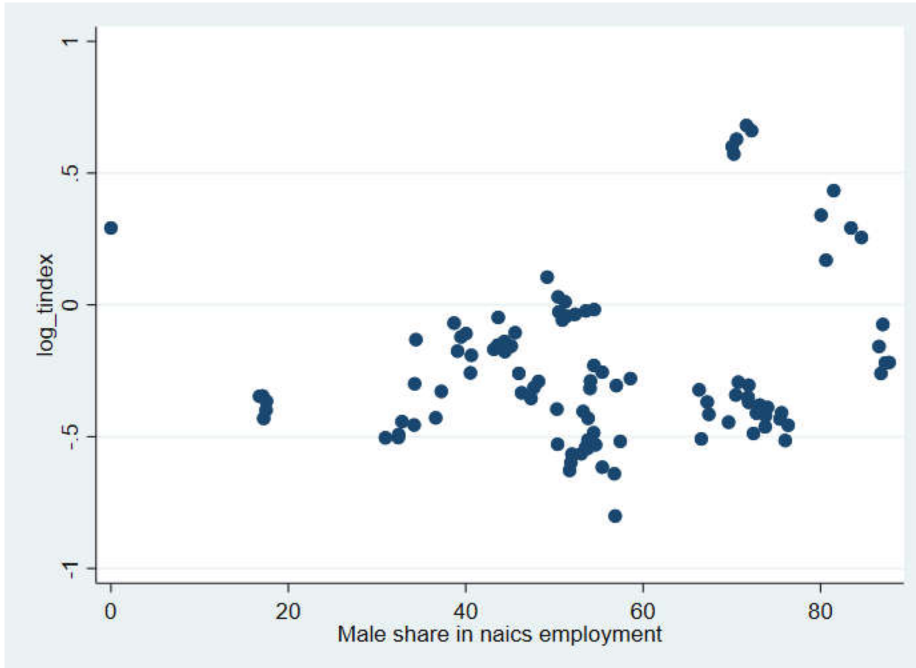

Figure A-6: Industry T index vs share of male workers in industry employment

Appendix II

Theil's values for individual CMAs (Within-CMA wage inequality):

CMA	1995	2000	2005	2010	2015	CMA average over time
St. John's	0.36897	0.36415	0.40276	0.35099	0.45962	0.389298
Halifax	0.33576	0.3821	0.38469	0.34727	0.40037	0.370038
Moncton	0.363	0.35971	0.33427	0.34278	0.32698	0.345348
Saint John	0.36599	0.51305	0.38354	0.35619	0.33026	0.389806
Saguenay	0.34057	0.3139	0.32065	0.3266	0.31699	0.323742
Quebec	0.31756	0.31486	0.30883	0.29188	0.30647	0.30792
Sherbrook	0.33645	0.31863	0.32082	0.32902	0.32624	0.326232
Trois-Riviera's	0.34392	0.33933	0.32809	0.3134	0.31021	0.32699
Montreal	0.35835	0.37368	0.41041	0.40595	0.45056	0.39979
Ottawa	0.32462	0.36873	0.38263	0.34536	0.36577	0.357422
Kingston	0.35487	0.42027	0.39765	0.37113	0.38363	0.38551
Peterborough	0.38569	0.38442	0.39158	0.37307	0.39027	0.385006
Oshawa	0.30791	0.31339	0.35462	0.36472	0.36981	0.34209
Toronto	0.39755	0.48972	0.58536	0.51913	0.5905	0.516452
Hamilton	0.35353	0.38495	0.44139	0.42359	0.45646	0.411984
St.Catharines	0.36235	0.36373	0.3981	0.37415	0.38096	0.375858
Kitchener	0.35932	0.37121	0.43927	0.41354	0.39642	0.395952
Brantford	0.33504	0.31034	0.35053	0.36937	0.31614	0.336284
Guelph	0.33938	0.35131	0.39367	0.3903	0.42487	0.379906
London	0.35505	0.3852	0.37714	0.38055	0.38081	0.37575
Windsor	0.35058	0.37763	0.39417	0.39116	0.41697	0.386102
Barrie	0.33726	0.33409	0.3588	0.36471	0.41779	0.36253
Sudbury	0.35264	0.35781	0.37705	0.38774	0.34649	0.364346
Thunder Bay	0.32692	0.34445	0.34001	0.37808	0.34148	0.346188
Winnipeg	0.35659	0.38146	0.39914	0.35931	0.43826	0.386952
Regina	0.34694	0.34683	0.34739	0.35661	0.34633	0.34882
Saskatoon	0.38	0.39535	0.51101	0.38908	0.39335	0.413758
Calgary	0.42367	0.46647	0.71362	0.60123	0.59009	0.559016
Edmonton	0.38753	0.38942	0.48085	0.40227	0.43178	0.41837
Kelowna	0.36302	0.37514	0.42113	0.39253	0.43305	0.396974
Abbotsford	0.35853	0.37042	0.39351	0.35737	0.35978	0.367922
Vancouver	0.37925	0.40456	0.49165	0.47172	0.45724	0.440884
Victoria	0.32723	0.33013	0.41293	0.37994	0.39625	0.369296
Average across CMAs	0.354425	0.372619	0.404462	0.382447	0.395521	0.381895

Appendix III

Scatter plots of within-CMA T values (in the log) in the vertical axis, with other control variables in the horizontal axis.

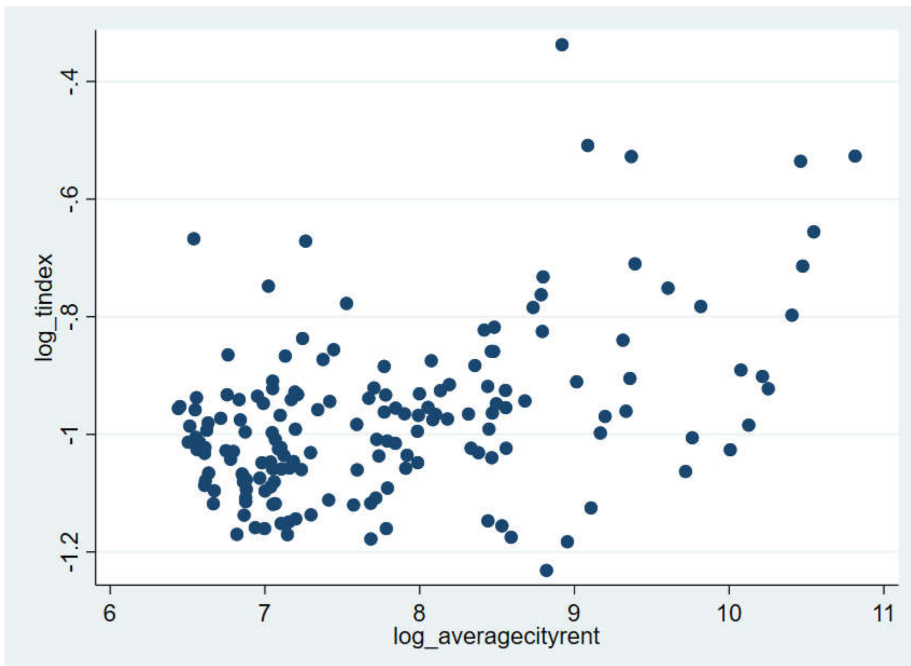

Figure A-7: CMA T index vs log of ACR index

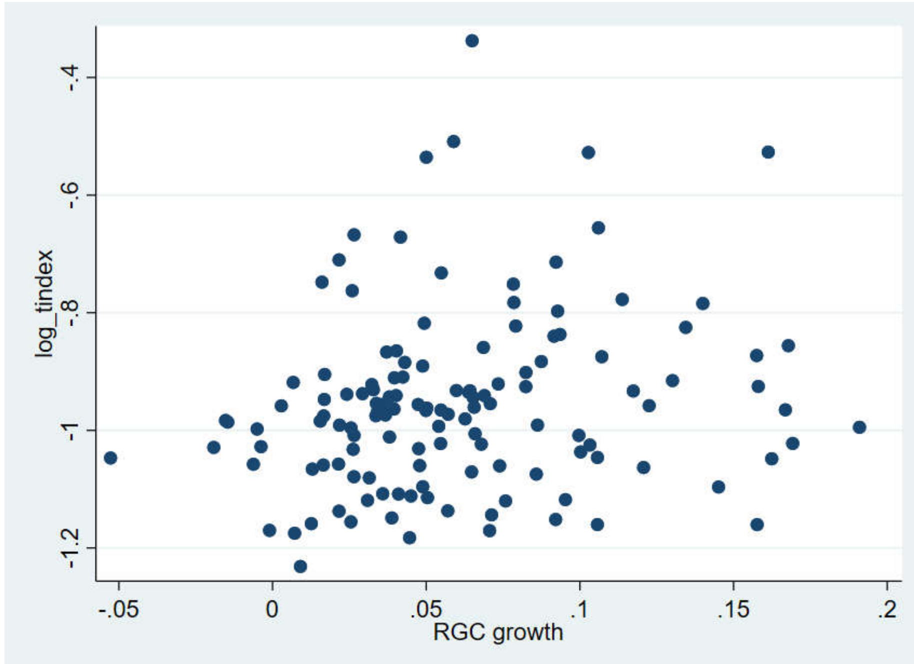

Figure A-8: CMA T index vs employment growth rate of RGCs

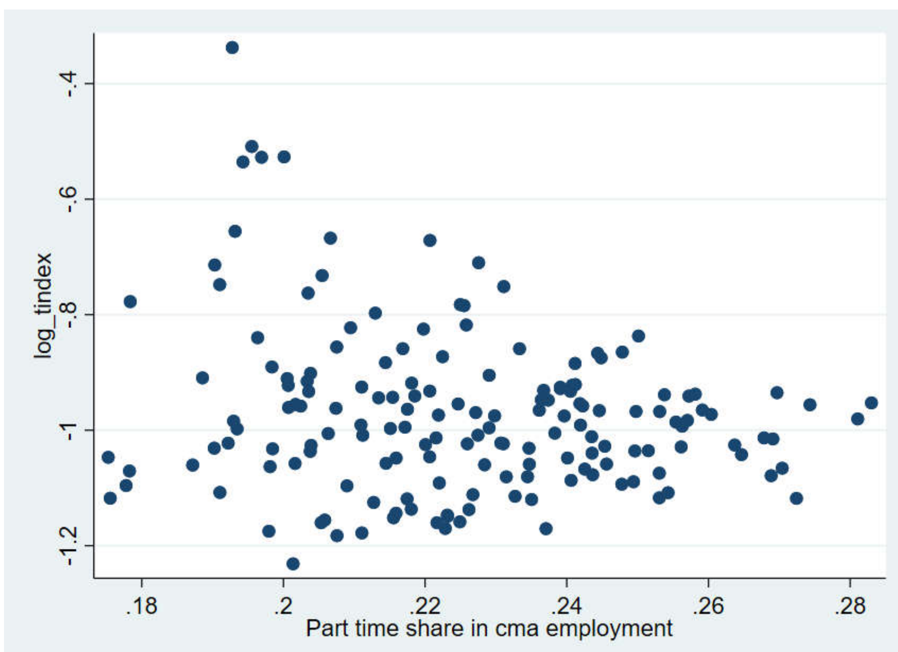

Figure A-9: CMA T index vs share of part-time workers in CMA total employment

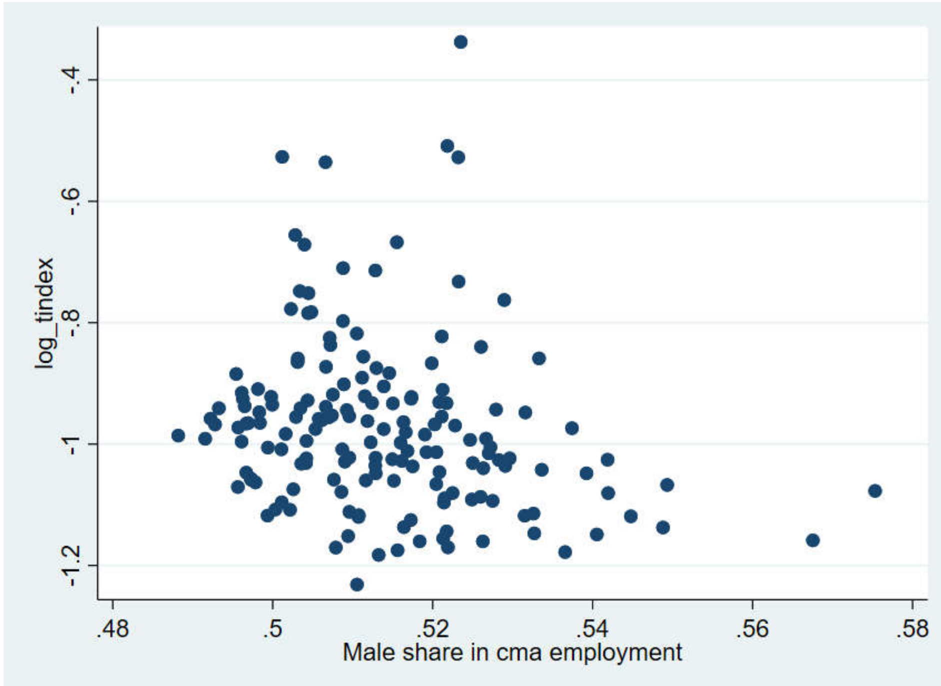

Figure A-10: CMA T index vs share of male workers in CMA total employment

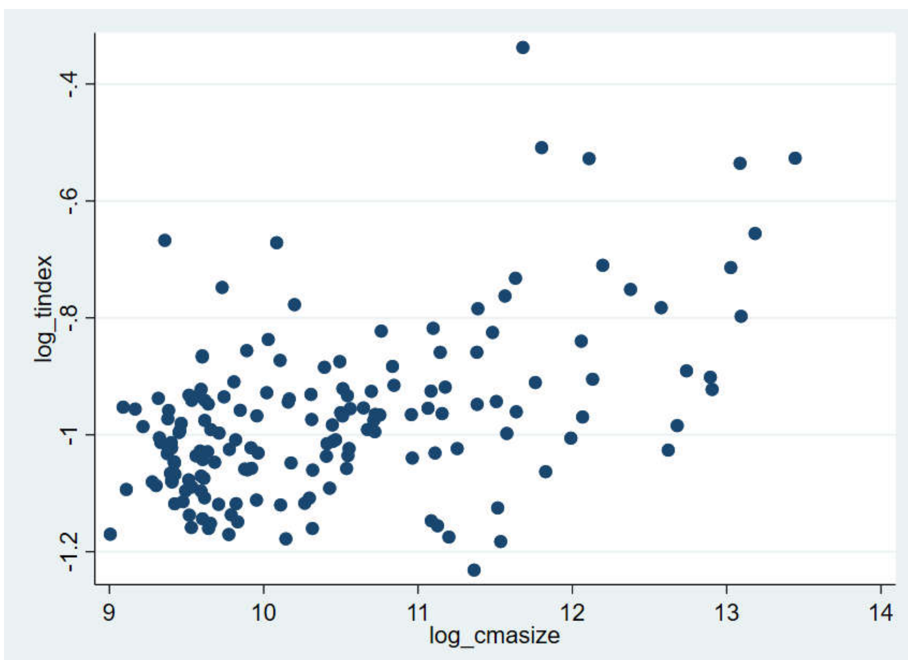

Figure A-11: CMA T index vs CMA size (population size of age 15 or over)

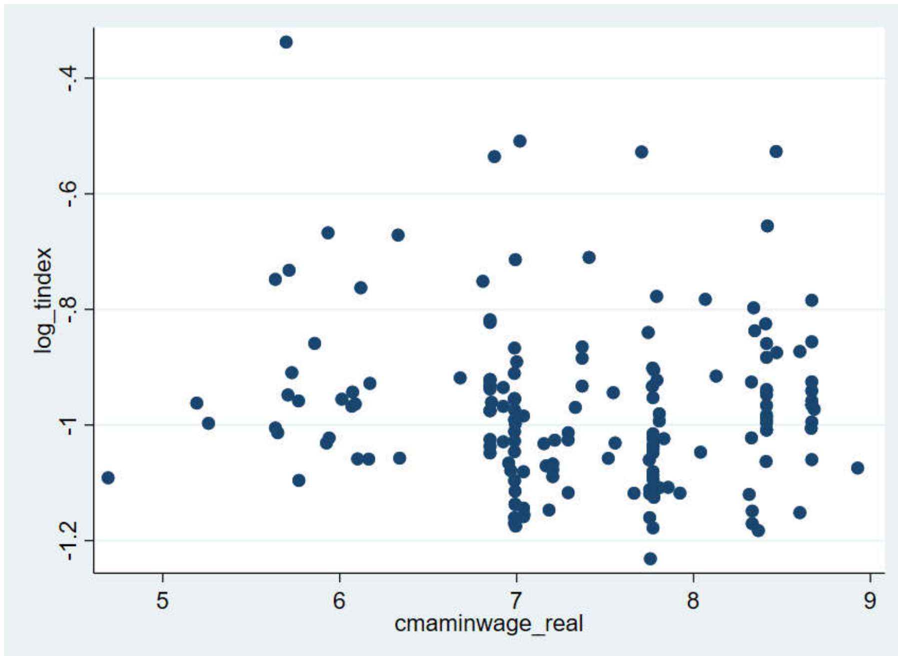

Figure A-12: CMA T index vs minimum wages (deflated by provincial CPI)

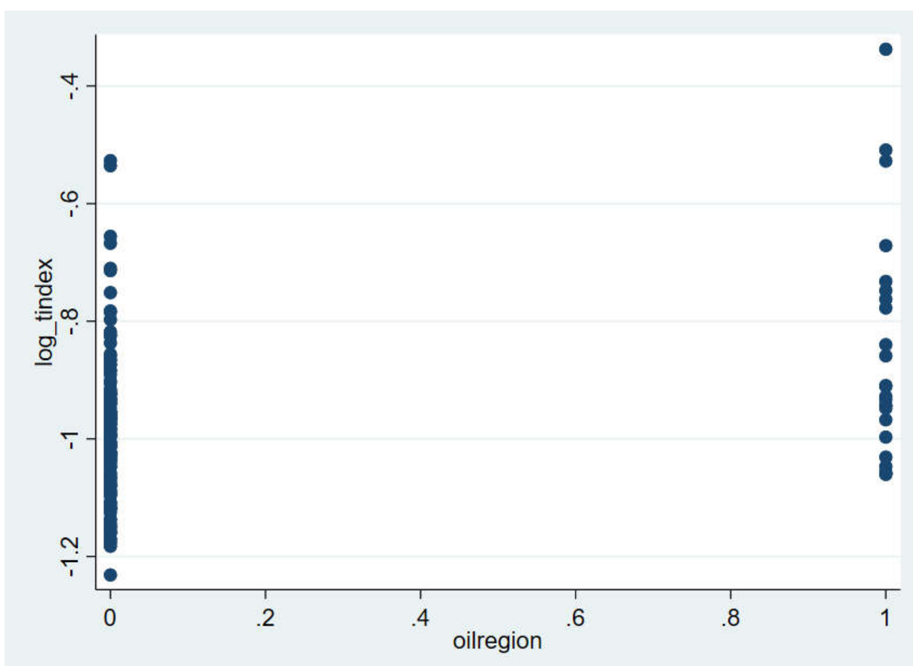

Figure A-13: CM T index vs oil-region dummy variable

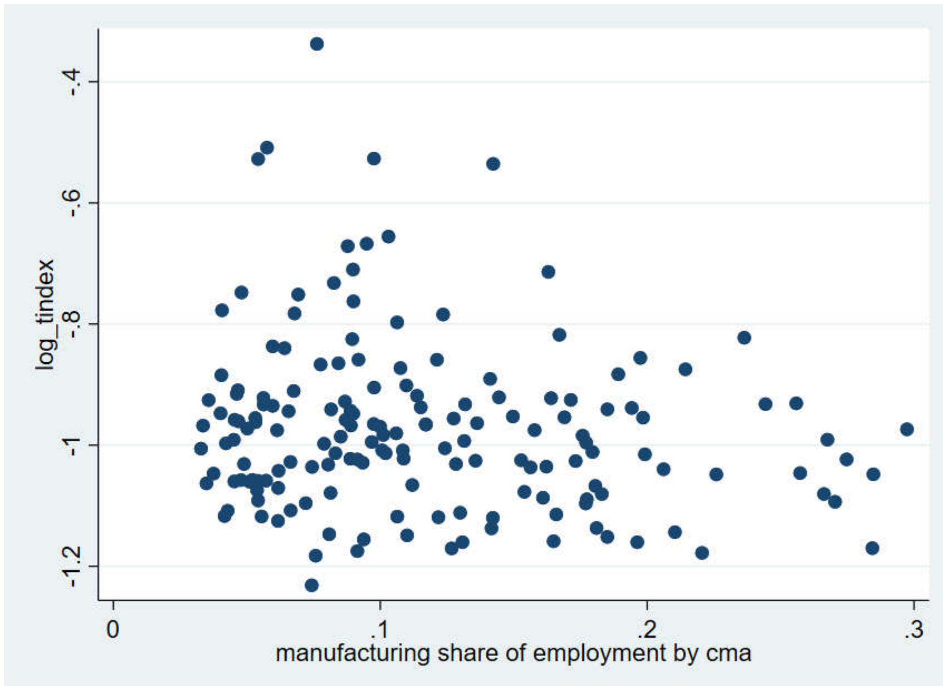

Figure A-14: CMA T index vs manufacturing share in total employment of CMA

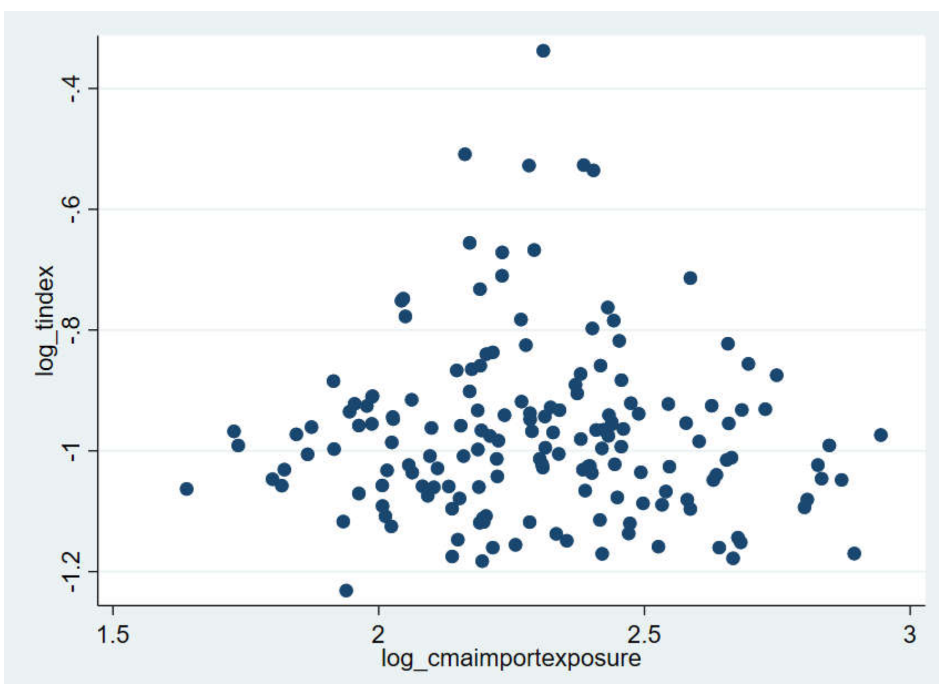

Figure A-15: CMA T index vs CMA import exposure

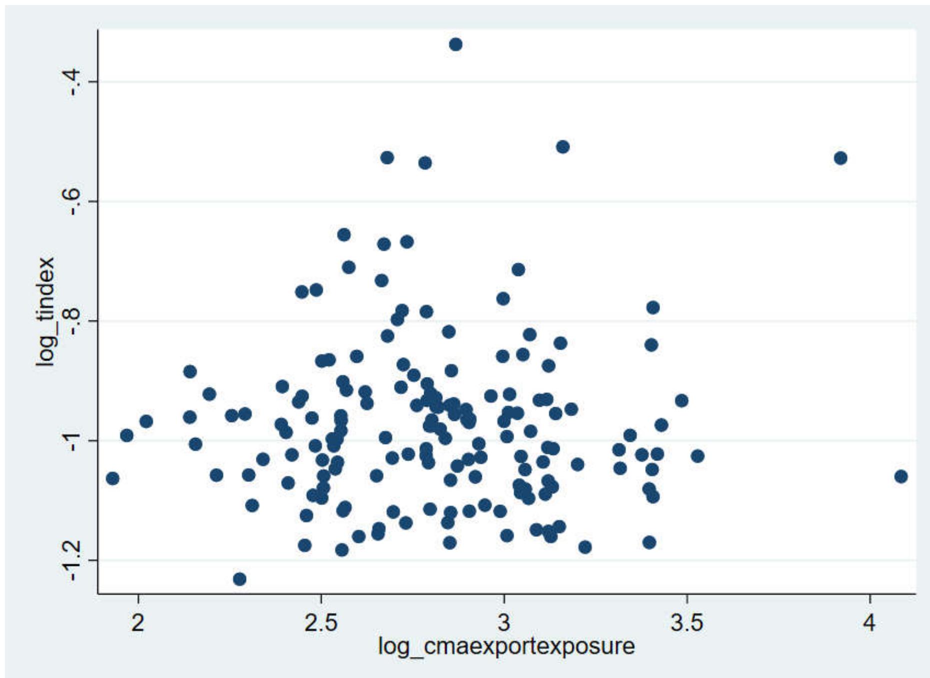

Figure A-16: CMA T index vs CMA export exposure

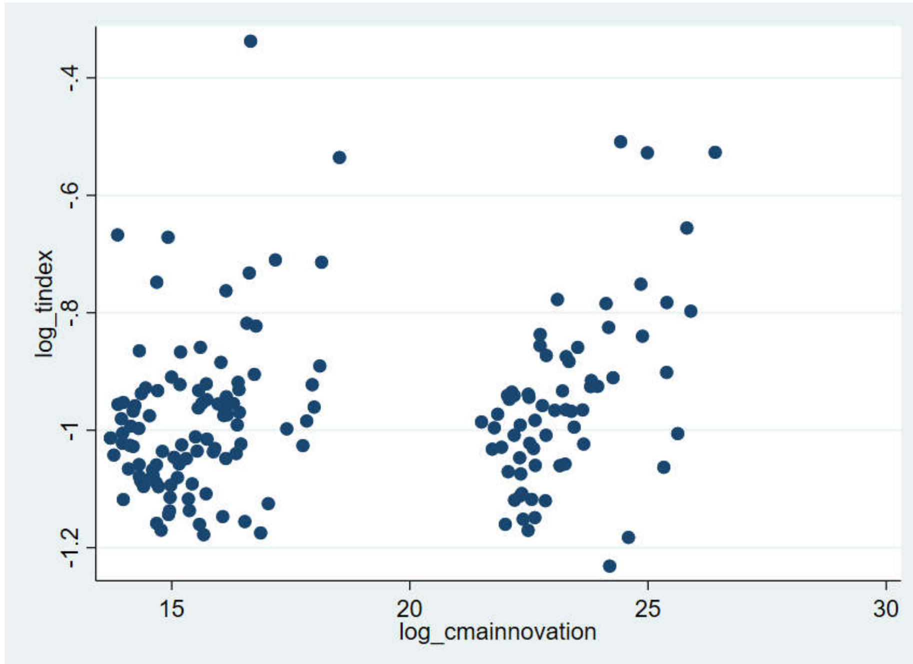

Figure A-17: CMA T index vs CMA innovation index

Appendix IV

The reweighting procedure in the DFL method is based on a random assignment of probability values for each observation. However, the method is designed to allow for more flexible assignment of weights to observations by using the following logit model of migration in which the probability of being a migrant is estimated and applied as weight. Estimation results are provided.

$$m_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 hdgree_i + \beta_4 sex_i + \beta_5 married_i + \beta_6 unemp_i + \beta_7 kid_i + \beta_8 prairies_i + e_i$$

Estimation results (migration* is the dependent variable), with standard errors reported in parentheses

Independent Var	**For all observations**	**For skilled workers**	**For unskilled workers**
agegrp	0.41 (0.04)	0.09 (0.07)	0.52 (0.04)
agegrp^2	-0.03 (0.00)	-0.02 (0.00)	-0.03 (0.00)
hdgree	0.06 (0.00)	0.14 (0.01)	0.02 (0.01)
sex	0.13 (0.02)	0.18 (0.03)	0.09 (0.02)
married	0.21 (0.02)	0.15 (0.03)	0.24 (0.03)
kid	-0.92 (0.02)	-0.87 (0.03)	-0.91 (0.02)
prairies	1.05 (0.03)	1.13 (0.05)	1.01 (0.04)
unemp	-0.05 (0.01)	-0.04 (0.02)	-0.05 (0.01)
cons	-4.44 (0.20)	-3.28 (0.43)	-4.85 (0.24)

*migration = 1 if the individual is a migrant, and =0 if non-migrant