Do Degrees Still Pay Off? Changing Returns to Education in Canada from 2005 to 2020.

Shaghayegh Azizmohammadi

A thesis in the Department of Economics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Economics) at
Concordia University
Montreal, Quebec, Canada

April 2025

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the Thesis prepared

By: Shaghayegh Azizmohammadi

Entitled: Do Degrees Still Pay Off? Changing Returns to Education in Canada from 2005 to 2020.

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Economics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

________________________________ Examiner
Dr. Gabrielle Vasey

________________________________ Supervisor
Dr. Christian Sigouin

Approved by: ________________________________
Dr. Christian Sigouin
Graduate Program Director

Date: ________________ ________________________________
Dr. Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Do Degrees Still Pay Off? Changing Returns to Education in Canada from 2005 to 2020.

Shaghayegh Azizmohammadi

This paper studies how returns to educational degrees has evolved in Canada from 2005 to 2020 by combining data from the 2006 and 2021 Canadian Census. The study uses a two-class finite mixture model to capture unobserved heterogeneity that may be correlated with individual characteristics, in a way that resembles the use of fixed effects in a panel data setting. The mixture model reveals two distinct groups and highlights unobserved differences in earnings that standard models could miss. The findings show that while higher education still leads to higher earnings, the return to education has empirically declined over time for individuals with graduate degrees. In contrast, returns to college and trade-level education have increased, suggesting that practical skills has become more valuable in the Canadian job market. These findings reveal the changing nature of the Canadian economy and help policymakers and educators improve education and training programs to better meet the needs of today's job market.

Keywords: University degree, Education level, Mixture Models, Canadian Census, Unobserved heterogeneity, Earnings, Mincer equations.

Acknowledgments

First and foremost, I want to express my deepest appreciation to Professor Christian Sigouin, my supervisor, whose patience, attention to detail, and constant support guided me through every step of this thesis. His thoughtful feedback and steady presence gave me confidence in my work and I felt supported at every step. I also want to thank Dr. McIntosh for his initial contribution to my thesis.

I also want to thank my husband for his support, encouragement, and belief in me. He is the only one who can truly understand the challenges and feelings I went through during creating this thesis.

Contents

List of Tables

1 Introduction

Changes in the earnings distribution by educational level and the many factors that affect it are among the most crucial aspects of labor economics. Understanding how education affects income helps us better discuss inequality and opportunity. It also plays an important role in shaping public policy, predicting job market trends, and deciding how to spend on education. Returns to education have been the subject of earlier Canadian studies (Bar-Or et al. (1995), Boudarbat, Lemieux, and Riddell (2010)), and the increased value of a university degree has been one of the most debated topics, especially when considering the changes in the job market caused by globalization, technological innovation, and changing economic systems.

This study contributes to the existing literature by using latent class mixture models to examine the changes in the economic returns to education, specifically university degrees, for Canadian men between 2005 and 2020. Higher education is often advertised as a way to create job opportunities by providing students with the qualifications and skills they need to get better jobs and achieve financial security. In recent years, there have been growing concerns that the financial returns to post-secondary education may be declining. For example, according to Royal Bank of Canada in 2023, undergraduate tuition fees in Canada increased by 12% between 2012 and 2017, however the average income for graduates increased by just 4% in the five years after graduation, showing a decreasing return on investment.[1] Similarly, the Conference Board of Canada rates Canada 10th out of 15 for the financial advantages that males get from having higher education, showing that the demands of the labor market and education may not be properly matched.[2] At the same time, academic studies like Boudarbat, Lemieux, and Riddell (2010) showed that the value of education for Canadian men increased from 1980 to 2005, especially for those with university degrees. They disputed previous studies using Census data that said returns were declining or remaining the same.

Using data from the 2006 and 2021 Canadian Censuses, this study provides empirical evidence showing the various ways in which a university education affects income. Although earlier research has looked at the overall advantages of education, not many have particularly looked at Canadian male population or how the worth of a university degree may have changed over time, especially with the changing structure of the labor market and the economy. The focus on Canadian men decreases variations related to gender-based income gaps and maintains consistency with previous research. Therefore, this study focuses exclusively on males, as early results suggest that women's income patterns should be studied separately. It provides a deeper understanding of the relationship between income, academic achievement and economic mobility. In order to explain how differences in income have changed and how higher education, in particular, plays a role in

1. Royal Bank of Canada (2023)
2. Conference Board of Canada (n.d.)

this change, this study compares 2005 to 2020 using mixture models to account for unobservable heterogeneity.

The methodology used here models unobservable features using a finite mixture model, where both the mean and variance of residual income vary with class membership, and the probability of class membership depends on individual characteristics—akin to fixed effects in panel data analysis. Although mixture modeling is increasingly used in sample survey data, it is uncommon to apply it to Mincer type equations to control for unobservable factors. Addressing unobserved heterogeneity is particularly important in a labour market context, as educational outcomes—the focus of the present study—are likely influenced by unobservable characteristics such as ability or skills. Failing to account for this could lead to biased estimates of the effects of education on earnings.

In a general theoretical context, Heckman, Lochner, and Todd (2006) have done a review of earnings functions and rates of return on education, emphasizing the variation in returns across different groups and the resulting policy implications. Their work provides an important theoretical foundation for this investigation, especially when it comes to estimating the returns on education by Mincer earnings function. Thus, their approach is relevant to the current study's goal, which is to explain the effects of education on income and assess the variability in returns to education for different income groups. Furthermore, studies like McLachlan and Peel (2000) and Wedel and Kamakura (2000) highlight that mixture models are useful for capturing variation in data that traditional regression analysis can miss. These studies clarify the advantage of using mixture model to understand earnings clearly, especially when unobserved factors play an important role.

The present paper's findings show that while higher education is still desirable, the returns to most degrees have declined between 2005 and 2020. For example, the return to a bachelor's degree decreased by about 3.6%, and the return to a master's or professional degree fell by around 5.9%, both statistically significant. The only group with an increased return is individuals with postsecondary education below the bachelor's level, who saw a small gain of 1.7%. The mixture model identifies two distinct groups of workers in the income distribution. The larger group, comprising approximately 86% of workers, is characterized by lower but more stable earnings. On average, individuals in this group earn about 18% less, and their earnings have a 40% lower standard deviation compared to the smaller, high-income group. Workers with higher levels of education are significantly less likely to belong to this lower-earning group.

This is a very important study for policymakers, educators, and economists, as it clearly explains how the value of the university degrees have changed over time. Additionally, it provides insight into helping workers navigate the increasingly competitive job market, ensuring their educational investments lead to better financial results. In the end, the research contributes to understanding ways to reduce income inequality and guides effective educational policy in Canada.

The rest of the paper is organized as follow. Section 2 reviews the relevant Canadian empirical work as well as related findings in OECD countries. The statistical model used in the research is presented in Section 3 and the data is described in Section 4. The main results are described in Section 5. Section 6 concludes with a summary of key findings and their policy implications..

2 Background and previous literature

There has been a great deal of research done on the relation between income and education, especially in Canada. The wage premium linked to university degrees and the ways in which labor market changes have impacted educational returns throughout time have been key areas of research. By combining research from several nations, such as the Canada, US, Europe, and other G12 countries, this review provides a global overview of this litreture. This review begins with Canada.

The wage advantage for university graduates in Canada compared to those of high school graduates was first studied by Bar-Or et al. (1995). They discovered that the value of a university degree decreased in the 1970s but started to rise again in the 1980s, especially for younger workers. This study started a basis for understanding the relationship between education and earnings. Further analysis by Boudarbat, et al. (2010) supported these findings by showing that the returns on education have increased over time, especially that of university degrees. Their conclusions emphasised how the value of higher education is increased by changes in the labor market and technology. Their results show that in an economy where demand for advanced abilities has grown due to developments in technology, highly educated workers have gained value. Besides these results, Boothby and Drewes (2006) also looked at the returns for a number of different types of post-secondary education programs, including trades, universities and colleges. Although those with a university degree regularly had the biggest wage advantages, there were also notable income benefits from other post-secondary education programs. This study supports the findings of Finnie and Frenette (2003), who examined differences in salaries by field of study and discovered that graduates of business, engineering, and health sciences programs typically made significantly more money than those from the humanities or social sciences. According to Finnie and Frenette (2003)'s research, the type and level of education has a significant impact on future wages because not all university degrees result in the same returns.

Morissette, Picot, and Lu (2012) also studied wage growth over a 30-year period and discovered that not all educational and age groups benefited equally from wage growth. Their research showed that while younger or less educated workers had significantly smaller improvements in wages, older, highly educated workers saw higher increases. This suggests that salary inequality is growing and supports the idea that education has a more complicated impact on earnings. These results show long-term changes in the labor market and emphasize the significance of assessing inequality in wages among different categories, a concept that this thesis explores using latent class analysis.

The relationship between cognitive and unobservable skills in predicting earnings was studied by Green and Riddell (2003). Their study found that cognitive skills, especially literacy are one of the most important factors in determining income, even after considering formal levels of education. Their findings suggest that education does not entirely explain the factors that influence income differences; individual skills also play a crucial role and are normally not measured in most data sets. Osberg (2000) also highlighted the strong impact of literacy and skills obtained outside of formal education on salaries, and what I discover confirm with his findings. This supports the importance of considering unobservable skills in wage studies.

Looking at studies from the United States, Katz and Murphy (1992) found that an increasing wage gap between skilled and unskilled workers was causing a considerable rise in wage inequality in United States between 1963 and 1987. Similarly, Card and Lemieux (2001) found that during the 1980s and 1990s, younger groups (aged 26 to 30) had stronger returns to education as a result of a decreasing supply and growing demand for workers with college degrees. More recently, Psacharopoulos (2024) studied the evaluation and history of educational returns. He found that returns have remained surprisingly stable over decades, consistently around 10%. This shows education as a reliable economic investment and shows important context for studying its current impact on earnings.

European research offers evidence of different trends across countries. For instance, a recent study by España (2025) studied the wage returns to education in Germany, France, Italy, and Spain. It found that Germany and France had the highest returns for tertiary education (about 20%), while Italy and Spain had lower returns (around 11–12%).[3] Brunello, Comi, and Lucifora (2000) studied how education affects wages in the Italian job market. They found that the rise in returns to education in the early 1990s was mostly due to higher wages in the public sector. This shows how the value of education can be significantly impacted by differences between sectors, especially between public and private employment. Their case study explains how institutional structures influence the relationship between education and earnings.

Research from other G12 countries also shows that returns to education vary depending on each country's economy and structures. According to Organisation for Economic Co-operation and Development in 2019, higher returns on university education are found in nations with creative and fast-growing industries, like South Korea and Japan than in countries with more traditional economies, such as Italy and Spain.[4]

3. Tertiary education refers to any level of education pursued after completing secondary school (high school).

4. Organisation for Economic Co-operation and Development (2019)

These international findings show that some countries have lower or less consistent returns on education, depending on their job markets and systems. However, many have seen rising or stable returns, especially in industries where innovation and technology have increased the demand for skilled workers. For example, Canada and the U.S. saw increases in the 1980s and 1990s, but Italy and Spain still show lower returns in recent years. Psacharopoulos (2024) found that returns to education are generally stable over time, though not always growing. Understanding these national differences is important for creating education and job policies that maximize educational investments. Therefore, the current paper aims to document the returns to education in recent years for Canada.

3 Empirical Model

To investigate whether returns to education have changed from 2005 to 2020, this paper estimates a traditional Mincer model using pooled cross-sectional data and incorporates unobserved heterogeneity through a finite mixture model. The model assumes that individuals can be categorized into distinct groups, and that group membership varies based on individual characteristics such as education, field of study, etc..

The traditional Mincer Equation which models the natural logarithm of individual wages $\ln(w_i)$ as a function of work experience (ex_i), education (educ_i), and other control variables (X_i) can be written as:

$$\ln(w_i) = \mu + \gamma_1 \cdot \text{ex}_i + \gamma_2 \cdot \text{ex}_i^2 + \sum_j \gamma_{3j} \cdot \text{educ}_{ij} + X_i\beta + \epsilon_{ij} \tag{1}$$

where the term educ_{ij} is a dummy variable for education level j, $\text{educ}_{ij} = 1$ if individual i has education level j, and 0 otherwise; one level is omitted as the reference category. Control variables typically explain productivity and income differences, and include factors such as geography, union status, marital status, industry, etc. In general, γ_1>0 and γ_2<0, reflecting a life-cycle productivity profile that initially increases, plateaus, and then declines. In addition, γ_{3j} normally increases with the degree attained, indicating rising returns to education.

The two main economic theories which form the basis of this analysis's basic structure are labor market segmentation and the theory of human capital. According to human capital theory, individuals invest in education to improve their productivity and earn higher incomes, which matches the principles of the Mincer equation used in this study. Labor market segmentation theory further supports the use of a latent class method, suggesting that the labor market is divided into separate groups where returns to education may differ, for example, due to differences in individual abilities. Using a latent class model, potentially captures these unobservables.

In economics, time-invariant unobserved heterogeneity is commonly addressed using fixed or random effects models when panel data is available. Fixed effects models account for individual-specific characteristics by allowing each individual to have their own intercept, capturing traits that may correlate with observed variables such as education. This is achieved by observing the same individuals over multiple time periods, using their within-individual variation to proxy for unobserved factors like ability. In contrast, random effects models assume that these unobserved individual-specific factors are uncorrelated with the explanatory variables and treat them as part of the error term, similarly to standard regression techniques.

With cross-sectional data or pseudo-panels (i.e., pooled cross-sections), neither fixed nor random effects models are applicable, as individuals are observed only once. This poses a challenge in labor market analysis, where unobserved factors, such as innate ability, job-skill match, or evolving skill demand, can significantly influence outcomes. For instance, workers with identical levels of education and observable characteristics (e.g., age or gender) may earn substantially different wages due to these unobserved differences. In standard regression models, such heterogeneity is absorbed into the error term, as in random effects models. However, if unobserved factors like ability are correlated with explanatory variables such as education, this introduces endogeneity, leading to biased estimates of the returns to education. Finite mixture models offer a potential solution by classifying individuals into latent groups that systematically capture unobserved heterogeneity. When the probability of class membership is allowed to depend on time-invariant characteristics, the model can approximate fixed effects, thereby reducing the correlation between residuals and explanatory variables and producing less biased estimates.

The Mincer Equation which is used in this paper is:

$$\ln(w_i) = \mu_k + \gamma_1 \cdot \text{ex}_i + \gamma_2 \cdot \text{ex}_i^2 + \sum_j \gamma_{3j} \cdot \text{educ}_{ij} + \sum_j \gamma_{4j} \cdot (\text{year}_i \cdot \text{educ}_{ij}) + X_i\beta + \epsilon_{ik} \quad (2)$$

where $\epsilon_{ik} \sim \mathcal{N}(0, \sigma_k^2)$

$$\sigma_k = \begin{cases} \sigma_H & \text{if } k = H, \text{ with probability } \pi_i, \\ \sigma_L & \text{if } k = L, \text{ with probability } 1 - \pi_i, \end{cases} \quad (3)$$

$$\mu_k = \begin{cases} \mu_H & \text{if } k = H, \text{ with probability } \pi_i, \\ \mu_L & \text{if } k = L, \text{ with probability } 1 - \pi_i, \end{cases} \quad (4)$$

and $\mu_H > \mu_L$. The probability of class membership π_i is assumed to depend on the set of individual characteristics, z_i, which is a subset of the explanatory variables in equation (2), according to

$$\pi_i = \frac{\exp(z_i^\top \delta)}{1 + \exp(z_i^\top \delta)}. \tag{5}$$

Note that μ_k is the value of class k "fixed effect", which is potentially correlated with individual explanatory variables because π_i depends on z_i.

The variable $\text{year}_i = 1$ if the year is 2020 and $\text{year}_i = 0$ if the year is 2005 and it is introduced to examine if returns to education vary from 2005 to 2020. The effect of education on log wages varies over time if $\gamma_{4j} \neq 0$ for any level j. In 2005 ($\text{year}_i = 0$), the effect of education level j is γ_{3j}; in 2020 ($\text{year}_i = 1$), it becomes $\gamma_{3j} + \gamma_{4j}$, reflecting changes in returns between the two years. These coefficients represent the percentage change in earnings associated with education level j, relative to the reference category (Less than high school), holding experience and other control variables constant. Therefore, this framework tests whether returns to education have changed based on γ_{4j}, and is represented by a finite mixture model defined by equations (2), (3), (4), and (5).

To write the likelihood function, θ_{ik} is the conditional mean of $\ln(w_i)$ for individual i in class k, and let σ_k^2 represents the variance of the error term ϵ_i for class k.

$$\theta_{ik} = \gamma_0 + \gamma_1 \cdot \text{ex}_i + \gamma_2 \cdot \text{ex}_i^2 + \sum_j \gamma_{3j} \cdot \text{educ}_{ij} + \sum_j \gamma_{4j} \cdot (\text{year}_i \cdot \text{educ}_{ij}) + X_i\beta + \mu_k. \tag{6}$$

The likelihood for individual i is:

$$L_i = \pi_i \cdot \frac{1}{\sigma_H} \phi\left(\frac{\ln(w_i) - \theta_{iH}}{\sigma_H}\right) + (1 - \pi_i) \cdot \frac{1}{\sigma_L} \phi\left(\frac{\ln(w_i) - \theta_{iL}}{\sigma_L}\right) \tag{7}$$

where the function $\phi(\cdot)$ is the standard Normal density function, defined as

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

The model's parameters are estimated by maximizing the log likelihood: $\log L = \sum_i \log(L_i)$. Note that π_i is not a fixed parameter but is estimated for each individual based on their observed characteristics z_i, as defined in Equation (5). Maximizing the log-likelihood function helps us find the model parameters that best match the data. In this case, it means identifying different wage patterns for each group and the probabilities of class membership that individuals belong to each group.

4 Data

This study uses individual-level data from the 2006 and 2021 Canadian Census of Population, accessed through Statistics Canada's Public Use Microdata Files (PUMF). These files provide retrospective anonymized information on demographics, education, labor market activity, and income for a representative sample of the Canadian population for 2005 and 2020, respectively.

To minimize unobserved heterogeneity, the analysis restricts the sample to a relatively homogeneous group: males aged 25 to 64 who are not attending school, have worked full-time during the reference year, and meet other criteria detailed below. These restrictions help isolate the relationship between educational attainment and earnings by reducing variation due to differences in labor force attachment, demographic background, and institutional contexts.

- Gender: The sample is restricted to men to reduce gender-related heterogeneity in labor market behavior. Female labor supply is more likely to be influenced by caregiving responsibilities, part-time work, or employment interruptions, which can introduce noise unrelated to education. By focusing on men, the analysis captures a group with more consistent labor force participation and earnings trajectories. A binary variable is used to code a person's gender. Gender was used to replace the 2005 Census label (SEX) to keep consistency between the 2020 and 2005 datasets. In order to keep just men (Gender=2) in the analysis, data with Gender=1 (females) were excluded.

- Age (AGEGRP): Individuals aged 25 to 64 are included to focus on the core working-age population and to exclude those still likely to be in school or nearing retirement. Age is a categorical variable that represents several age groups, each of which includes five years, with a total of 8 categories considered in the analysis (agegrp codes 9 to 16). Agegrp$<$9 and agegrp$>$16, which are outside the 25–64 age range, have been excluded in the analysis.

- School attendance (ATTSCH): Only individuals not attending school at the time of the census are included to avoid capturing temporary earnings disruptions. "Did not attend school" (code 1) was the only category kept; all other detailed attendance categories were dropped.

- Employment Duration (WKSWRK): The analysis is restricted to individuals who worked at least 49 weeks in the reference year, which is the top category (49–52 weeks) reported in 2020. While the Mincer model typically uses hourly or weekly wages as the dependent variable, such information is not available in the 2021 Census PUMF. In 2020, weeks worked is only reported in categories, making it difficult to construct reliable wage rates. Restricting the sample to near full-year workers minimize variability in annual earnings due to part-year work and makes earnings a closer approximation of the wage rate.

- Class of worker (COW): Only wage and salary employees are included. Self-employed individuals and unpaid family workers are excluded to improve comparability in wage structures. This variable has significant coding differences between 2005 and 2020. In 2020, category 4 indicates "Self-employed, with paid help, incorporated", whereas in 2005, it stands for "Paid worker-Working for wages, salary, tips, or commission". Also, category 1 in 2005 denoted "Unpaid family workers", but in 2020 it denoted "Employee". Only observations that were categorized as paid employees (cow=4 in 2005 and cow=1 in 2020) were considered. The analysis excluded all other categories, which represented self-employed people, unpaid family workers, and non-applicable cases (cow>1).

- COVID-19 Benefits (COVID-ERB): In 2020, individuals who reported receiving emergency or recovery payments due to COVID-19 are excluded to ensure comparability with 2005 data.

- Indigenous Identity (ABOID): Individuals identifying as Indigenous are excluded, following Cameron and Heckman (1998), due to limited available information and small cell sizes that reduce the reliability of subgroup analysis.

- Geography (PR): Respondents residing in the three northern territories (Yukon, Northwest Territories, and Nunavut) are excluded due to small population sizes and distinct regional labor markets.

The selection of variables is grounded in human capital theory (Mincer (1974)) and empirical work such as Boudarbat, Lemieux, and Riddell (2010). Educational attainment is the primary variable of interest, as it proxies for both productivity-enhancing skills and labor market signaling. Education is measured both by level and field of study to allow for heterogeneity in returns across different types of qualifications.

Although field of study is closely linked to education level, this paper includes it as a separate factor in the model. A binary variable is used to show whether someone studied in a STEM field or not, helping to compare earnings across different fields. While the impact of a STEM degree might change depending on the level of education (for example, it could be stronger at the graduate level), this study does not look at those differences. Instead, it assumes that the effect of having a STEM degree is the same at all education levels to keep the model clear and easy to understand.

Age serves as a proxy for labor market experience and captures cumulative learning, tenure effects, and productivity growth over the life cycle. Although direct measures of experience are unavailable, categorical age indicators are used to account for non-linear earnings profiles. According to economic theory, the relationship between experience and earnings is expected to be positive but concave.

Additional controls—including immigration status, industry of employment, marital status, visible minority status, province or residence—are included to account for observable differences in earnings potential. These covariates help reduce omitted variable bias and clarify the estimated effect of education.

- Educational Attainment (HDGREE): Grouped into six categories as shown below:

  1. Less than high school (hdgree=1)
  2. High school diploma or equivalent (hdgree=2)
  3. Postsecondary below bachelor's level (hdgree=3, 4, 5, 6, 7, 8)
     - Includes non-apprenticeship trades certificate or diploma, apprenticeship certificate, university certificate or diploma below bachelor's level, and college programs.
  4. Bachelor's degree or some postgraduate study (hdgree=9, 10)
     - Includes bachelor's degree, and university certificate or diploma above bachelor level
  5. Master's or professional degree (hdgree=11, 12)
     - Includes degree in medicine, dentistry, veterinary medicine or optometry and master's degree
  6. Doctorate (PhD) (hdgree=13)

  Time variation in the returns to degrees will be measured by adding a year dummy variable to the model and interacting it with educational attainment.

- Field of Study (CIP): Postsecondary education fields are grouped in a variable indicating if education is in a STEM field as follows:

  1. No postsecondary education:
     - 2005: CIP=13
     - 2020: CIP=11
  2. STEM (Science, Technology, Engineering, Math, IT):
     - 2005: CIP=6 (Physical and life sciences), 7 (Mathematics, computer sciences), 8 (Architecture, engineering)
     - 2020: CIP=1 (Science and science technology), 2 (Engineering and engineering technology), 3 (Mathematics and computer sciences)

3. Non-STEM: All remaining CIP categories.

Category 1 corresponds to those with high school degrees only or none.

- Industry (NAICS): Collapsed into six broad sectors:

1. Public and Social Services
   - 2005: 15 (Educational services), 16 (Health care and social assistance), 20 (Public administration)
   - 2020: 61 (Educational services), 62 (Health care and social assistance), 91 (Public administration)
2. Trade and Hospitality
   - 2005: 7 (Retail trade), 18 (Accommodation and food services)
   - 2020: 44 (Retail trade), 72 (Accommodation and food services)
3. Professional and Business Services
   - 2005: 10 (Finance and insurance), 12 (Professional, scientific and technical services)
   - 2020: 52 (Finance and insurance), 54 (Professional, scientific and technical services)
4. Manufacturing and Construction
   - 2005: 4 (Construction), 5 (Manufacturing)
   - 2020: 23 (Construction), 31 (Manufacturing)
5. Resources and Utilities
   - 2005: 1 (Agriculture, forestry, fishing and hunting), 2 (Mining and oil and gas extraction), 3 (Utilities)
   - 2020: 11 (Agriculture, forestry, fishing and hunting), 21 (Mining, quarrying, and oil and gas extraction), 22 (Utilities)
6. Other
   - 2005: 6 (Wholesale trade), 8 (Transportation and warehousing), 9 (Information and cultural industries), 11 (Real estate and rental and leasing), 14 (Administrative and support, waste management and remediation), 17 (Arts, entertainment and recreation), 19 (Other services)
   - 2020: 41 (Wholesale trade), 48 (Transportation and warehousing), 51 (Information and cultural industries), 53 (Real estate and rental and leasing), 56 (Administrative

and support, waste management and remediation), 71 (Arts, entertainment and recreation) 81 (Other services)

- Marital Status (MARSTH): A binary variable called "everm" is coded as 0 if never married and 1 otherwise. This includes those who are currently or previously married or in a common-law relationship. Specifically, codes 2 (Married or living common-law), 4 (Separated), 5 (Divorced), and 6 (Widowed) are recoded as 1. Those who were never married (MARSTH=1 in 2020, MARSTH=4 in 2005) are coded as 0. This binary classification simplifies the study by indicating if an individual has ever been in a common-law or marital relationship, which might impact factors like family roles and income. It is easier to compare people who have never been married to those who have by putting all of those who have ever been in such a relationship together.

- Visible Minority (VISMIN): A binary variable coded 1 for individuals identifying as a visible minority, 0 otherwise. When analyzing differences in employment outcomes, the binary variable makes it easier to compare visible minority and non-minority persons.

- Immigration status (AGEIMM): Immigrants are coded as those who are not likely to have completed high school in Canada, which typically corresponds to individuals who immigrated at or after age 14. This classification is based on the assumption that individuals arriving at a later age are less likely to have completed their secondary education within the Canadian system. Those who were likely to have completed high school in Canada (i.e., those who immigrated before age 14) are coded as non-immigrants. This distinction helps isolate the effect of educational attainment within the Canadian context, as late-arriving immigrants may have different educational experiences and challenges that impact labor market outcomes differently than individuals who completed their education in Canada.

- The model includes dummy variables for provinces to account for provincial differences in cost of living, which may be reflected in differences in earnings, types of industries, etc.

- Wages, salaries and commissions (WAGES): Individuals reporting earnings equal to 1 in 2020 and zero in 2005 are excluded from the analysis to reduce potential biases. The sample is further restricted to remove earnings values that appear unrealistic. Specifically, in 2005, individuals with annual earnings below \$9,750, calculated as 52 weeks × 30 hours × \$6.25 (the lowest provincial wage in Canada at the time) are excluded. Similarly, in 2020, earnings below \$17,862, based on the lowest provincial minimum wage of \$11.45, are dropped. Because Statistics Canada considers 30 hours per week to be full-time employment, this limit is applied.

In order to compare 2005 earnings to 2020 earnings, the Consumer Price Index (CPI) was used to convert all 2005 earnings data to 2020 dollars. In 2005, Canada's CPI was 107.0, and in 2020, it was 137.0 (using 2002 as base year). This adjustment makes sure that differences in earnings are real changes over time, not changes in prices.

In addition to the exclusions mentioned earlier, this study did not include the observations that were top-coded or labeled as missing (not applicable or available) in the dataset. This step is taken to ensure data consistency and accuracy across all observations, and to maintain the reliability of the comparisons between 2005 and 2020.

Along with the variable definitions mentioned above, Tables 1 and 2 provide further information for comparing the Canadian male workforce between 2005 and 2020. Table 1 presents the summary statistics for real earnings and shows that incomes increased at all levels, but the changes were not the same for everyone. People with very low incomes (1st percentile) saw a big increase of about 63%, while the top earners (99th percentile) saw a smaller increase of about 22%. The average and middle incomes also went up by about 20%. This corresponds to an annualized growth rate of approximately 1.25% per year. The increases in real earnings are in line with the general salary trends that have been seen in Canada throughout this time. During these years, the real minimum wage grew by approximately 43%, reflecting improvements in the labor market in general and wage growth across various sectors. Government actions like raising the minimum wage or providing social support programs may help to explain why lower percentiles have had more growth. The standard deviation, an indicator of income inequality, remained almost the same. This shows that although everyone's income increased, the general income gap did not grow significantly, however it changed more at the low and high levels.

Table 2 shows big changes in education. There are far fewer men without any degree in 2020 (5.8%) compared to 2005 (11.7%), while the percentage of people with a bachelor's degree or master's/professional certificate has grown significantly (from 17.7% to 26.9% for a bachelor's degree, and from 5.1% to 9.3% for MA/professional). This shows that Canadian men are becoming more educated overall. We also see some changes in the population. There are slightly more people in the 55–64 age group in 2020 than there were in 2005, showing that the workforce is getting older. There are more immigrants (18.9% in 2020 vs. 15% in 2005) and more visible minorities (23.4% vs. 12.7%). A smaller percentage of people have ever been married, which could be a sign of changing in family and social patterns. Looking at industries, we see a change in the kinds of jobs people have. Fewer people work in manufacturing and construction (down from 28.3% to 21.9%) while professional/business services and public/social services have both grown.

5 Results

This section shows the changes in returns to education for Canadian men between 2005 and 2020 based on the empirical model defined by equations (2) to (5). The set of covariates X_i in equation (2) includes education, age, industry, marital status, immigrant status, visible minority, province, industry, and field of study (STEM/non-STEM). The subset of individual characteristics z_i, influencing π_i in equation (5), includes education, field of study (STEM/non-STEM), visible minority, and immigrant status. Table 3 and 4 report the estimation results for the mixture model. Table 5 reports the corresponding results using Ordinary Least Squares as a benchmark. Table 6 compares the relative performance of both estimation methods using the Bayesian Information Criterion (BIC) and the log-likelihood. All estimates are obtained using the sampling weights, adjusted so that each year is equally weighted. In addition, all standard errors are clustered at the individual level.

The results in Table 3 are in line with expectations from a Mincer regression and, also show that the type of education matters. Table 3 shows that, compared to workers with a STEM background, those with a non-STEM education earn about 6.9% less. This finding supports the earlier results of Finnie and Frenette (2003), who found that graduates in business, engineering, and health sciences earned significantly more than those with degrees in the humanities or social sciences. This difference suggests that STEM-related education continues to offer better financial returns in the labor market. The sample is divided into two groups by the mixture model: a high-wage group (Class 2) and a low-wage group (Class 1) that makes up most of the sample (about 86%). Also according to Table 3, where we can see the class-specific intercepts, the average log earnings in the high-income class (μ_H=10.763) are statistically significantly higher than in the low-income class (μ_L=10.581), which means average earnings differ by about 18% ($\mu_L < \mu_H$). In addition, the standard deviation of log earnings is much greater in the high-income group (σ_H=0.926) than in the low-income group (σ_L=0.376) indicating more unobserved heterogeneity and greater earnings inequality among top earners. ($\sigma_L < \sigma_H$).

Table 4 shows that, people with more education are more likely to be in the high wage group. For example, having a bachelor's degree increases the chance of being in that group by 0.537 and this rises to 0.856 for those with a master's or professional degree. In contrast, visible minority status reduces the probability by 0.413, and being an immigrant increases it by 0.318, their higher probability of being in the top class may reflect unobserved factors such as motivation or ability or because of their foreign education, which can affect both their chances of earning more and their job stability. This shows how both visible factors (like education and being an immigrant) and unobserved ones (like how their education is valued or their communication abilities) play a role in which group they belong to (class membership) and how much they earn. We can also see that, people with non-STEM degrees are also more likely to be in the high-income group (coefficient = 0.524), which might be unexpected. This could mean that even though non-STEM workers usually earn less, some still get high-paying jobs, maybe in areas like law, business, or management.

Factors such as age and marital status, we can see a clear effect on income. Table 3 shows that, income increases over time until the age range of 50–54 and then falls off. People who are married or were married earn more than those who were never married. Being a visible minority or an immigrant is linked to lower income, both characteristics are associated with lower wages. This shows there are still gaps and unfair outcomes in the job market. Where someone lives and what kind of job they have also affect income. According to Table 3, employees in Quebec make less money (-0.134) compared to workers in Ontario (the reference group), but workers in Alberta earn more (+0.070).

In terms of industry, people working in resource or utility jobs earn the most, while trade and hospitality (like restaurants and shops) have the lowest returns (-0.240) in comparison to public and social services. Jobs in professional and business services, as well as manufacturing, also pay more than average. This shows that returns to education depend not just on degrees, but also on job type and location.

People with higher education generally earn more. In Table 3, compared to less than high school, those with a high school diploma earn about 13% more, while those who have attended some college or trade school earn approximately 28% more. People holding a bachelor's degree earn around 51% more, and those with a master's or professional degree see earnings about 64% higher. At the top of the scale, individuals with a PhD earn about 76% more than those with less than high school education. But between 2005 and 2020, the effect of education on earnings has decreased for most degree holders. The interaction terms between year and education levels in Table 3 show statistically significant declines in the returns to bachelor's, master's, and PhD degrees between 2005 and 2020. For example, the return to a master's or professional degree declined by about 5.9%, and for a PhD by 6.8%. On the other hand, the only education group that shows a positive and significant increase over time is those with postsecondary education below the bachelor's level, with an increase of 1.7%. This could suggest that practical and technical skills, such as plumbing, electrical work, or culinary arts, have become more valuable in the Canadian job market.

According to Table 5's outcomes, the estimated returns to education for both the OLS and mixture models are mostly comparable. It means that when general patterns are the main focus, OLS can still be helpful for calculating average returns. However, the mixture model offers more accurate estimates of educational returns. The decrease in return to a PhD from 2005 to 2020, for example, is significant in the mixture model (-0.068) but not in the OLS results (-0.062), even though the point estimates are almost the same. This highlights the mixture model's ability to identify variations that traditional models would miss. Furthermore, Table 6 demonstrates that the mixture model offers a better fit to the data than the OLS model, with the higher log-likelihood (better fit) and a lower BIC (better for comparing models). This empirical evidence confirms the presence of latent heterogeneity in earnings, highlights the importance of accounting for unobserved differences when studying income gaps and education returns, and shows that the mixture model does a better job of capturing these unobserved factors and providing a more detailed view of earnings patterns.

6 Conclusion

This paper studied how return to education for Canadian men have changed between 2005 and 2020, using data from the 2006 and 2021 Canadian Censuses. By using a two-class mixture model, this study was able to show both average effects and unobserved differences within the income distribution. The mixture model, specified with two latent classes, identified distinct wage groups and provided a better understanding of how education affects earnings for different types of workers. The results show that education still has a significant impact on income. People with more education, especially bachelor's and graduate degrees, earn much more than those with less than high school. However, the extra earnings from these degrees have become smaller over time. For example, the return to a master's degree declined by about 5.9%, and for a PhD by 6.8% across both income groups, since the model assumes a common change in returns over time. On the other hand, people with college or trade school education saw an increase in their earnings, which may reflect a growing need for practical skills. About other factors, like visible minorities and immigrants, they typically make less money. This means unfair inequalities still exist in the Canadian labor market.

Future studies should look at these results in more detail by studying how they differ for men and women, and in industries affected by new technology and new trends like remote work. By understanding these differences, policymakers and educators can improve higher education to match the needs of today's job market. This would help all graduates have fair chances and can succeed in valuable and fulfilling careers.

7 Use of Generative AI and AI-assisted tools

During the preparation of my thesis, I used TeXstudio and chatGPT for fixing the LaTeX errors that I cannot recognize by myself, finding mathemathical symbols and specific codes for drawing some of the figures.

After using this tool/service, I reviewed and edited the content as needed and take full responsibility for the content of my thesis.

References

Bar-Or, Yuval, John Burbidge, Lonnie Magee, and A Leslie Robb. 1995. "The wage premium to a university education in Canada, 1971-1991." *Journal of Labor Economics* 13 (4): 762–794.

Boothby, Daniel, and Torben Drewes. 2006. "Postsecondary education in Canada: Returns to university, college and trades education." *Canadian Public Policy/Analyse de Politiques,* 1–21.

Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell. 2010. *The Evolution of the Returns to Human Capital in Canada, 1980–2005.* Technical report Working Paper No. 53. Accessed March 2025. Canadian Labour Market and Skills Researcher Network (CLSRN). https://ideas.repec.org/p/clm/wpaper/53.html.

Brunello, Giorgio, Simona Comi, and Claudio Lucifora. 2000. *The Returns to Education in Italy: A New Look at the Evidence.* IZA Discussion Paper 130. Institute of Labor Economics (IZA). https://docs.iza.org/dp130.pdf.

Cameron, Stephen V, and James J Heckman. 1998. "Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males." *Journal of Political economy* 106 (2): 262–333.

Card, David, and Thomas Lemieux. 2001. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis." *The quarterly journal of economics* 116 (2): 705–746.

Conference Board of Canada. n.d. "Education and Skills - How Canada Performs." https://www.conferenceboard.ca/hcp/education-aspx/.

España, Banco de. 2025. "Wage returns to education in the four largest European economies." *Boletín Económico,* https://www.bde.es/f/webbe/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/25/T1/Files/be2501-art03e.pdf.

Finnie, Ross, and Marc Frenette. 2003. "Earning differences by major field of study: evidence from three cohorts of recent Canadian graduates." *Economics of education review* 22 (2): 179–192.

Green, David A, and W Craig Riddell. 2003. "Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation." *Labour Economics* 10 (2): 165–184.

Heckman, James J, Lance J Lochner, and Petra E Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." *Handbook of the Economics of Education* 1:307–458.

Katz, Lawrence F, and Kevin M Murphy. 1992. "Changes in relative wages, 1963–1987: supply and demand factors." *The quarterly journal of economics* 107 (1): 35–78.

McLachlan, Geoffrey J., and David Peel. 2000. *Finite Mixture Models.* Wiley.

Mincer, Jacob. 1974. "Schooling, experience, and earnings. Human behavior & social institutions no. 2."

Morissette, René, Garnett Picot, and Yuqian Lu. 2012. "Wage growth over the past 30 years: Changing wages by age and education." *Economic insights,* no. 8.

Organisation for Economic Co-operation and Development. 2019. *Education at a Glance 2019: OECD Indicators.* Paris: OECD Publishing. https://doi.org/10.1787/f8d7880d-en. https://www.oecd-ilibrary.org/education/education-at-a-glance-2019_f8d7880d-en.

Osberg, L. 2000. "Schooling, literacy, and individual earnings (No. 89-553-MIE, no. 7)." *Ottawa, ON: Statistics Canada, Human Resource Development Canada.*

Psacharopoulos, George. 2024. "Returns to Education: A Brief History and an Assessment." *Education Economics* 32 (5): 561–565. https://doi.org/10.1080/09645292.2024.2370119. https://www.tandfonline.com/doi/full/10.1080/09645292.2024.2370119.

Royal Bank of Canada. 2023. "Proof Point: Financial Returns After a Post-Secondary Education Have Diminished." RBC Thought Leadership. https://thoughtleadership.rbc.com/proof-point-financial-returns-after-a-post-secondary-education-have-diminished/.

Wedel, Michel, and Wagner A. Kamakura. 2000. *Market Segmentation: Conceptual and Methodological Foundations.* Springer.

8 Appendices

Table 1: Summary Statistics of Real Earnings By Year

Statistic	**2005**	**2020**	**Growth Rate (%)**
1st Percentile	\$15,364	\$25,000	62.7
5th Percentile	\$25,607	\$36,000	40.5
25th Percentile	\$46,093	\$58,000	25.8
Median (50th)	\$66,579	\$80,000	20.2
75th Percentile	\$93,467	\$110,000	17.7
95th Percentile	\$153,644	\$200,000	30.2
99th Percentile	\$409,719	\$499,245	21.9
Mean	\$81,681.48	\$98,309.04	20.3
Standard Deviation	\$88,268.00	\$87,914.00	–0.4
Number of Observations	98,989	89,364	

Note: Percentiles calculated using sampling weights.

Table 2: Proportions of Categorical Variables By Year

Variable	**Category**	**2005**	**2020**
Education	Less than high school	0.117	0.058
	High school or equivalent	0.237	0.207
	Post-secondary $<$ BA	0.403	0.354
	BA or some postgraduate	0.177	0.269
	MA or professional certificate	0.051	0.093
	PhD	0.011	0.016
Field of Study	No post-secondary	0.355	0.265
	STEM	0.339	0.364
	Non-STEM	0.304	0.369
Immigrant Status	No (0)	0.849	0.810
	Yes (1)	0.150	0.189
Ever Married	No (0)	0.271	0.378
	Yes (1)	0.728	0.621
Visible Minority	No (0)	0.872	0.765
	Yes (1)	0.127	0.234
Age Group	25–29	0.105	0.096
	30–34	0.126	0.131
	35–39	0.137	0.144
	40–44	0.167	0.140
	45–49	0.164	0.135
	50–54	0.144	0.133
	55–59	0.104	0.132
	60–64	0.049	0.085
Province	Newfoundland and Labrador	0.011	0.010
	Prince Edward Island	0.003	0.003
	Nova Scotia	0.027	0.024
	New Brunswick	0.022	0.020
	Quebec	0.235	0.220
	Ontario	0.408	0.406
	Manitoba	0.034	0.033
	Saskatchewan	0.025	0.027
	Alberta	0.112	0.119
	British Columbia	0.117	0.133
Industry	Public/Social Services	0.174	0.219
	Trade/Hospitality	0.111	0.092
	Professional/Business Services	0.100	0.167
	Manufacturing/Construction	0.283	0.219
	Resources/Utilities	0.054	0.057
	Other	0.275	0.243

Table 3: Mixture Model Regression Results (Two-Class Model)

Variable	**Coefficient**	**Standard Error**
Education (ref: Less than high school)		
High school or equivalent	0.128*	0.0051
Post-secondary < BA	0.275*	0.0048
BA or some postgrad	0.514*	0.0058
MA or professional cert.	0.642*	0.0082
PhD	0.760*	0.0138
Year x Education Interaction		
2020 x High school	–0.006	0.0081
2020 x Post-secondary < BA	0.017*	0.0075
2020 x BA or some postgrad	–0.036*	0.0083
2020 x MA or prof. cert.	–0.059*	0.0109
2020 x PhD	–0.068*	0.0193
Field of Study (ref: STEM)		
Non-STEM	–0.069*	0.0026
Industry (ref: Public/Social)		
Trade/Hospitality	–0.240*	0.0044
Professional/Business Services	0.125*	0.0036
Manufacturing/Construction	0.030*	0.0030
Resources/Utilities	0.246*	0.0052
Other Industry	–0.043*	0.0031
Age Group (ref: 25–29)		
Age 30–34	0.150*	0.0039
Age 35–39	0.235*	0.0041
Age 40–44	0.278*	0.0040
Age 45–49	0.310*	0.0042
Age 50–54	0.309*	0.0043
Age 55–59	0.273*	0.0046
Age 60–64	0.206*	0.0057
Province (ref: Ontario)		
Newfoundland and Labrador	–0.165*	0.0098
Prince Edward Island	–0.331*	0.0156
Nova Scotia	–0.207*	0.0063
New Brunswick	–0.239*	0.0068
Quebec	–0.134*	0.0027
Manitoba	–0.146*	0.0054
Saskatchewan	–0.074*	0.0063
Alberta	0.070*	0.0035
British Columbia	–0.011*	0.0033
Ever Married	0.145*	0.0024
Visual Minority	–0.139*	0.0031
Immigrant	–0.144*	0.0035
Year = 2020	0.172*	0.0080
Constant(μ_L)	10.581*	0.0059
σ_L (Class 1)	0.376	0.0013
Constant (μ_H, Class 2)	10.763*	0.0128
σ_H (Class 2)	0.926	0.0060

*Significant at the 5% level.

Table 4: Class Membership Model and Posterior Probabilities

Variable	**Coefficient**	**Standard Error**
Education (ref: Less than high school)		
High school or equivalent	0.210*	0.0644
Post-secondary < BA	–0.114	0.0629
BA or some postgrad	0.537*	0.0693
MA or professional cert.	0.856*	0.0763
PhD	0.772*	0.1115
Field of Study (ref: STEM)		
Non-STEM	0.524*	0.0329
Visual Minority	–0.413*	0.0504
Immigrant	0.318*	0.0483
Constant	–2.277*	0.0722
Average Posterior Probabilities:		
Class 1: 0.863 Class 2: 0.136		

*Significant at the 5% level.

Table 5: OLS Regression Results

Variable	**Coefficient**	**Standard Error**
Education (ref: Less than high school)		
High school or equivalent	0.139*	0.0054
Post-secondary < BA	0.273*	0.0050
BA or some postgrad	0.549*	0.0065
MA or professional certificate	0.688*	0.0101
PhD	0.787*	0.0174
Year x Education Interaction		
2020 # High school or equivalent	–0.011	0.0087
2020 # Post-secondary < BA	0.013	0.0081
2020 # BA or some postgrad	–0.043*	0.0091
2020 # MA or professional certificate	–0.056*	0.0131
2020 # PhD	–0.062	0.0229
Field of Study (ref: STEM)		
Non-STEM	–0.045*	0.0029
Industry (ref: Public/Social)		
Trade/Hospitality	–0.190*	0.0046
Professional/Business Services	0.190*	0.0044
Manufacturing/Construction	0.064*	0.0033
Resources/Utilities	0.279*	0.0057
Other	–0.002	0.0033
Age Group (ref: 25–29)		
30–34	0.152*	0.0042
35–39	0.244*	0.0043
40–44	0.297*	0.0044
45–49	0.335*	0.0045
50–54	0.335*	0.0047
55–59	0.299*	0.0050
60–64	0.239*	0.0062
Province (ref: Ontario)		
Newfoundland and Labrador	–0.168*	0.0104
Prince Edward Island	–0.330*	0.0168
Nova Scotia	–0.217*	0.0067
New Brunswick	–0.243*	0.0072
Quebec	–0.138*	0.0030
Manitoba	–0.149*	0.0062
Saskatchewan	–0.091*	0.0069
Alberta	0.067*	0.0040
British Columbia	–0.013*	0.0038
Ever Married (Yes)	0.160*	0.0027
Visual Minority (Yes)	–0.159*	0.0034
Immigrant (Yes)	–0.156*	0.0040
Year = 2020	0.193*	0.0080
Constant	10.522*	0.0063
σ (Std. dev. of residuals)	0.491	0.0013

*Significant at the 5% level.

Table 6: Comparison of Mixture Model vs. OLS

Statistic	**Mixture Model**	**OLS Model**
Log-likelihood	–4,662,456	–4,958,879
BIC	9,125,507	9,918,219
Number of Parameters (N)	49	38