University of British Columbia

Department of Economics

ECON 398

Introduction to Applied Economics

Term Paper Cover Sheet and Declaration

Student Name: Suhani Jain

Student Number: 72907694

Date: 5th December 2023

Title: How does Education Qualification affect Wage Level?

Declaration: I understand the University's policy on academic integrity. (See www.students.ubc.ca/calendar/index.cfm?tree=3,54,111,959). The ideas and written work submitted for this assignment are my own except where properly cited. This work has not been presented by myself or another person for grading in another course. I understand that the assignments will not be graded unless accompanied by this cover sheet and signed declaration.

Signed

Suhani Jain

How does Education Qualification affect Wage Level

Introduction:

The relationship between education and income has long been a topic of significant interest and study in economics. As societies evolve and industries progress, the connection between an individual's level of education and their earning potential has become increasingly evident. This paper aims to explore and analyze the intricate correlation between education attainment and wage levels. By delving into various socioeconomic factors and statistical data, this study seeks to elucidate how different levels of education impact an individual's earning capacity in today's dynamic and competitive workforce. Understanding this relationship is crucial not only for individuals navigating their educational paths but also for policymakers and institutions aiming to foster economic growth and reduce income disparities. Through empirical evidence and thorough analysis, this paper endeavors to shed light on the multifaceted interplay between education and wages, offering valuable insights into the significance of education as a determinant of economic success and social mobility.

This study adopts a mixed-methods approach to examine the relationship between education and wage levels. Quantitative analysis involves the utilization of large-scale dataset from reputable sources, such as census data. Statistical methods, including regression analysis, are employed to scrutinize the correlation between different levels of education and corresponding wage outcomes while controlling for relevant variables such as age, gender, immigration status, and visible minority.

Furthermore, qualitative assessments are conducted through literature reviews to provide a deeper understanding of the contextual nuances shaping the education-wage nexus. This multifaceted methodology aims to offer a comprehensive and nuanced exploration of the intricate relationship between education attainment and wage levels.

Preliminary findings suggest a clear positive correlation between higher levels of education and increased earning potential. Individuals with advanced degrees tend to command higher wages compared to those with lower educational qualifications. The study examines the link between education and wages using two approaches: quantitative analysis of 2021 Census data from Statistics Canada and qualitative analysis of various research papers.

Quantitative analysis involves cleaning Census data, studying education and income distributions, and running regression models to quantify education's impact on wages while considering other factors. Qualitative analysis reviews academic papers, explores theories, and showcases case studies to illustrate how education influences wages in different contexts.

.

In today's world, young people, especially those from Generation Z, are focused on finding a path that leads not just to a job but to a fulfilling career with good financial prospects. Many students want to know which educational choices like different types of degrees or levels of education could lead to higher starting salaries or better pay overall. This research topic was chosen because there's a strong interest among students in understanding how their educational choices might impact their future earnings. As young individuals embark on their educational

journeys, they're eager to figure out which educational paths could potentially offer them better financial opportunities. This investigation into how education and income levels are connected is timely and relevant, addressing the very real concerns of students who want to make informed decisions about their future careers.

Background:

Numerous studies have extensively investigated the relationship between education qualifications and wage levels, contributing to a rich body of research in economics and social sciences.

Helen et al. (2006) highlighted the substantial impact of education on wages, emphasizing the positive correlation between higher educational attainment and increased earning potential. The study considered various levels of education and their differential effects on wages, shedding light on the importance of advanced degrees in augmenting income.

Additionally, Jan et al. (2021) explored the evolving wage structure, emphasizing the changing dynamics between educational qualifications and wages in response to technological advancements and economic shifts. This research illuminated the role of specific skills acquired through education and their influence on wage differentials in the modern labor market.

Furthermore, Brahim et al. (2013) investigated recent trends in wage inequality, emphasizing disparities in wages based on educational backgrounds. This study underscored the widening earnings gap between individuals with varying levels of education, highlighting the increasing importance of education in determining wages.

Moreover, the article published by the University of Arizona Global Campus (2023) explored the relationship between income inequality and the earnings gap, emphasizing the role of education as a crucial factor in mitigating wage disparities and reducing income inequality.

Additionally, Taylor et al. (2014) delved into the rising costs of not pursuing higher education, elucidating the economic consequences of limited educational qualifications on wage prospects and career advancement.

Lastly, Kapoor et al. (2021)examined the investment in education, highlighting its pivotal role in shaping wage levels and addressing socio-economic challenges, particularly in developing economies.

Suter et al. (2016) discuss the income disparities between men and career women. It likely explores the wage gap and the factors contributing to differences in earnings between men and women in the workforce. The authors might delve into various aspects such as gender discrimination, occupational choices, education levels, and work experience, highlighting how these elements influence income discrepancies. Additionally, the article might propose potential solutions or policy recommendations to address these disparities and promote greater gender equality in terms of earnings in the workplace. It also emphasizes that fewer women than men are able to break away from only average-paying jobs into those with higher incomes. However in the Census data 2021, the data shows that women earn higher than men. This might be due to occupational choices, work hours or part-time work. These studies collectively underscore the significance of education qualifications in shaping wage levels, highlighting the multidimensional impact of education on earnings and its implications for individuals, economies, and societal inequality.

Data:

In this study utilizing the 2021 Census Public Use Microdata File from Statistics Canada, the primary investigation centers on understanding the causal link between education qualifications and wage levels in Canada. The independent variable of interest is the education qualification, encompassing various levels of educational attainment within the dataset. The first level is No certificate, diploma or degree. The second level of education is Secondary education which comprises individuals who have achieved a high (secondary) school diploma or equivalency certificate, non-apprenticeship trades certificate or diploma or an apprenticeship certificate. The third level is Post-secondary education which comprises of Program of 3 months to less than 1 year , Program of 1 to 2 years , Program of more than 2 years , University certificate or diploma below bachelor level , Bachelor's degree, University certificate or diploma above bachelor level and the fourth level is Advanced degrees which comprises of a Degree in medicine, dentistry, veterinary medicine, or optometry, Master's degree or an Earned doctorate

This diverse range forms the treatment group, allowing examination of how different education levels influence wage earnings. To establish a comparative analysis, individuals with the highest education level, such as those with advanced degrees, are considered the reference or baseline group. This comparison aids in discerning differences in wage outcomes based on varying education qualifications.The control variables—age, gender, visible minority status, and immigration status—serve the specific purpose of isolating and evaluating their influence solely on the dependent variable, wage levels.

Age: Age may directly impact an individual's wage levels due to varying career stages and experience levels. Younger individuals might be in the process of education completion, while older individuals might have established careers influencing their wages differently.

Gender: Gender disparities in wages are well-documented. Utilizing gender as a control variable allows for the examination of wage differences beyond the influence of education qualifications.

Visible Minority Status: This variable may impact an individual's access to education and job opportunities, thus influencing wage disparities. Controlling for visible minority status helps in assessing the impact of education on wages while minimizing the effects of these disparities.

Immigration Status: Immigration status may affect education opportunities and labor market outcomes. Considering immigration status as a control variable aids in isolating the influence of education on wages among immigrants and non-immigrants. These control variables are included to meticulously isolate the impact of education qualifications on wage levels, ensuring that any observed effects can be attributed primarily to variations in education attainment while accounting for potential influences of these demographic factors on wage outcomes.

Methodology:

The analysis conducted using the 2021 Census data from Statistics Canada aims to delve into the relationship between education qualifications and wage levels among individuals in Canada. The methodology followed a systematic approach involving data management, visualization, and regression analysis.

Initially, the data was imported into R utilizing libraries like tidyverse, haven, dplyr, ggplot2, and stargazer for efficient data handling and analysis. Relevant variables, such as gender, visible minority status, wages, education level, and age group, were selected from the dataset using specific functions to isolate crucial information for the analysis. Data cleaning was a crucial step to ensure the accuracy and reliability of the analysis. Operations such as filtering

out incomplete or erroneous entries and creating new variables, like 'degree_modified' to categorize education levels into distinct groups, were performed using functions like mutate() and filter().

The transformation of wage data into logarithmic values using the log() function was employed to normalize skewed wage distributions and establish a more linear relationship between education and wages. This log transformation aids in addressing potential data irregularities and enhances the statistical analysis. Descriptive statistics, including mean and maximum wages across various education levels, were calculated using summarize() to gain insights into the wage distribution. Visualization techniques such as bar plots with ggplot() were utilized to present average log wages against educational attainment levels.

The analysis proceeded to regression modeling to assess the direct impact of education on log wages without initially incorporating control variables. Later stages of the analysis planned to introduce control variables like age, gender, religion, and visible minority status into the regression model. This step-by-step approach aimed to progressively capture a more comprehensive understanding of how multiple factors collectively influence wage outcomes, while considering potential confounding effects.

The methodology adopted in this analysis revolves around employing multiple linear regression models to scrutinize the connection between log wages and a spectrum of demographic and socioeconomic factors. These models are constructed iteratively, each probing

distinct combinations of independent variables like education level, age, immigration status, gender, and visible minority status vis-à-vis their influence on log wages. Rigorous variable selection is pivotal, guided by the research question and desired insights into their individual and collective effects on earnings. Using statistical software, regression coefficients are estimated and assessed for significance through tests, gauging the strength and statistical relevance of these relationships. Metrics such as adjusted R-squared aid in comprehending the proportion of variance in log wages explained by the amalgamation of independent variables. The interpretation of coefficients involves discerning the impact of each variable on log wages while holding other factors constant. The findings are then distilled into insights delineating the determinants of wage disparities and the nuanced influence of demographic attributes on income levels, culminating in a comprehensive understanding of the analyzed relationships.

For regression coefficients to be interpreted causally, adherence to several crucial assumptions is imperative. First, ensuring the absence of omitted variable bias demands that all pertinent variables influencing both the independent variables (such as education level, age) and the dependent variable (log wages) are encompassed within the model. Secondly, exogeneity dictates that independent variables remain free from correlation with the error term, signifying their immunity to unobserved factors influencing the dependent variable. Endogeneity, if present, suggests a correlation between independent variables and the error term, jeopardizing the premise of a causal relationship. Additionally, it's essential to avoid multicollinearity among independent variables, ensuring they aren't highly correlated, while also upholding linearity and additivity in the relationship between independent and dependent variables. These assumptions are crucial for validly interpreting regression coefficients as causal indicators.

Results:

The analysis of mean wages revealed distinct variations among different educational categories. Individuals with Advanced Degrees exhibited the highest mean wages, whereas those without formal education displayed comparatively lower mean earnings. Simultaneously, standard deviation in wages within each education level highlighted varying income disparities. Lower educational qualifications corresponded to greater wage differences, indicating increased income inequality among less-educated groups. This is indicated in Figure 1 and 2.

```
# A tibble: 4 × 3
  degree_modified                  max_wages mean_wages
  <chr>                                <dbl>      <dbl>
1 Advanced_Degrees                    967998     87217.
2 No_certificate_diploma_or_degree    967998     27693.
3 Post_Secondary_Education            967998     59585.
4 Secondary_Education                 967998     40587.
```

Fig 1. Mean Wages

```
# A tibble: 4 × 4
  degree_modified                min_wages max_wages sd_wages
  <chr>                              <dbl>     <dbl>    <dbl>
1 Advanced_Degrees                       1    967998   98007.
2 No_certificate_diploma_or_degree       1    967998   34499.
3 Post_Secondary_Education               1    967998   65102.
4 Secondary_Education                    1    967998   44268.
```

Fig 2. Min, Max and Standard deviation

Moreover, the bar graph visualization depicts a clear positive correlation between higher education levels and average log wages. Each educational category exhibited a progressive

increase in log wages from individuals without formal education to those with Advanced Degrees. This visual representation underscored the significance of educational qualifications in shaping wage disparities. This is shown in figure 3.

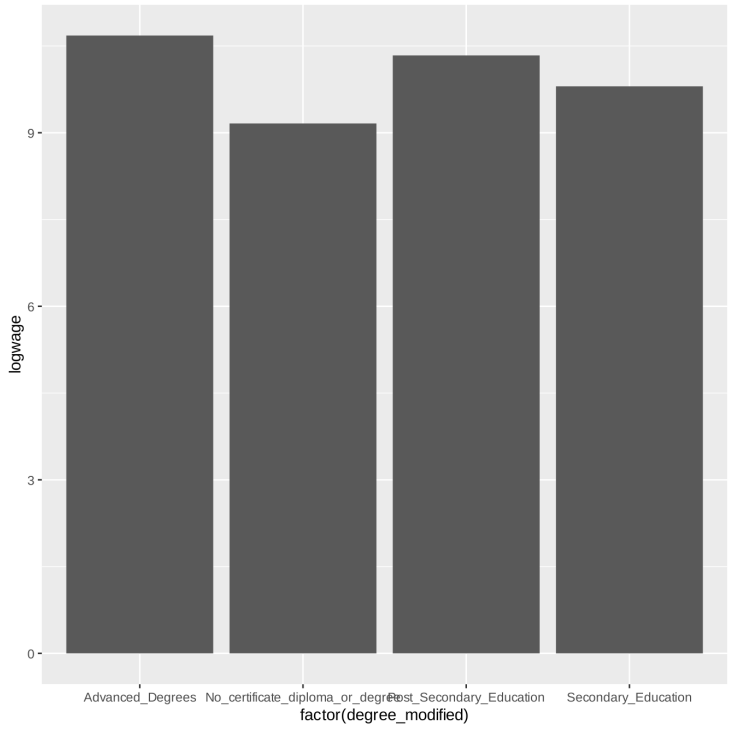

Fig 3. Bar Graph depicting trend in wages

Additionally, the regression analyses, spanning models regression1 through regression5, provided quantitative insights into log wage determinants. Initially focusing on education alone (regression1), subsequent models incorporated control variables like age group, immigration status, gender, and visible minority status.

Comparison of Regression Results

	Dependent variable:				
	logwage				
	(1)	(2)	(3)	(4)	(5)
degree_modifiedNo_certificate_diploma_or_degree	−1.520***	−1.522***	−1.524***	−1.562***	−1.565***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
degree_modifiedPost_Secondary_Education	−0.344***	−0.347***	−0.347***	−0.331***	−0.333***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
degree_modifiedSecondary_Education	−0.878***	−0.881***	−0.882***	−0.915***	−0.918***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
age_group		−0.009***	−0.009***	−0.009***	−0.009***
		(0.0005)	(0.0005)	(0.0005)	(0.0005)
Immigration_status			−0.004***	−0.004***	−0.002***
			(0.001)	(0.001)	(0.001)
gender				0.344***	0.344***
				(0.006)	(0.006)
visible_minority					−0.002***
					(0.0002)
Constant	10.681***	10.789***	10.795***	10.287***	10.294***
	(0.010)	(0.011)	(0.011)	(0.014)	(0.014)
Observations	518,729	518,729	518,729	518,729	518,729
R2	0.042	0.042	0.042	0.049	0.049
Note:					*p<0.1; **p<0.05; ***p<0.01

Fig 4. Summary of Regression 1-5

The table presents the results of regression analysis with the dependent variable being the natural logarithm of wages (logwage) and several independent variables related to education, demographic factors, and other attributes. Summary of the key finding were that individuals without certificates, diplomas, or degrees have significantly lower logwages compared to those with Advanced degrees (Reference category). As the level of education decreases from Post-Secondary Education to Secondary Education, there's a consistent decrease in logwages. Older age is associated with slightly lower logwages, as indicated by the negative coefficient for age_group. Being an immigrant is associated with lower logwages. Women tend to have higher logwages compared to men, as indicated by the positive coefficient for gender. Being part of a visible minority might be associated with slightly lower logwages. The constant term represents the estimated logwage for the reference categories of all other variables.The overall model

explains a small percentage of the variance in logwages, with R-squared values ranging from 0.042 to 0.049 across different model specifications.

In summary, the results suggest that education level, age, immigration status, gender, and visible minority status are associated with variations in logwages. However, the model explains only a small portion of the total variance in logwages, indicating that there might be other unaccounted factors influencing wage differences among individuals. In essence, these combined results underscored the crucial role of education in influencing wage disparities while also highlighting the impact of various demographic and socioeconomic factors. The findings collectively emphasize the multifaceted nature of income differentials, indicating the necessity for comprehensive policies addressing diverse determinants of wage inequalities.

Discussion + extension:

The regression analysis suggests a connection between higher education and increased wages, indicating that individuals with more education tend to earn more. However, establishing a direct cause-and-effect relationship between education levels and wages is challenging due to several important factors.

Firstly, while the data shows a correlation, it doesn't confirm causation. Factors like innate abilities, motivation, or other unmeasured variables might influence both education attainment and wages separately, making it hard to attribute wage differences solely to education.

Secondly, there's the issue of endogeneity, where unobservable factors related to wages may influence education levels. For instance, individuals from wealthier backgrounds might have

better access to education, impacting both their education levels and eventual earnings. This can lead to a biased estimation of the impact of education on wages.

Another challenge is reverse causality, where higher wages might enable individuals to afford better education, leading to higher educational qualifications. This suggests that higher income levels could be the cause of higher education rather than the other way around. Additionally, the analysis might overlook important variables affecting wages, such as specific skills acquired during education or job experience. This omission of relevant factors could lead to a bias in estimating the true relationship between education and wages.

Including falsification tests provides a means to check the validity of the observed relationship between education qualifications and wages. For instance, one falsification test could involve examining whether education levels influence an outcome that logically shouldn't be affected by education.

For instance, if education genuinely affects wages, one could test whether education levels influence other unrelated outcomes (e.g., favorite food choices) using the same dataset. A lack of correlation between education and these unrelated outcomes would strengthen the argument that the observed relationship between education and wages isn't merely coincidental.

Additionally, employing placebo tests, where education is randomly assigned and checking for impacts on outcomes not related to wages could serve as a falsification test. If the

impact of education is only observed on wages but not on these placebo outcomes, it would further support the causal interpretation of the education-wage relationship.

By integrating falsification tests into the analysis, the paper not only strengthens the causal inference of the relationship between education qualifications and wage levels but also helps identify potential biases or limitations within the dataset or model, enhancing the overall robustness and validity of the findings.

Conclusion:

The research findings presented through regression analyses have elucidated significant correlations between various demographic factors and logwages. Education levels, age, immigration status, gender, and visible minority status have demonstrated noteworthy associations with earnings. Specifically, lower logwages were observed for individuals lacking certificate/diploma, post-secondary, or secondary education compared to those with degrees. The inclusion of age alongside education levels marginally enhanced the model's predictive capability. Furthermore, immigrant status showed a correlation with decreased logwages, while gender exhibited a strong association with increased earnings for females within the model's scope. Conversely, being part of a visible minority was associated with lower logwages.

To extend this preliminary exploration, future research endeavors should aim to establish causative relationships between these demographic variables and logwages, employing longitudinal or experimental designs for a more robust investigation. Moreover, inclusion of

additional nuanced variables such as work experience, industry sectors, geographic location, or job positions would enrich the analysis. Exploring potential interaction effects between variables and integrating qualitative research methods could provide a deeper understanding of individual experiences and perceptions related to wage disparities. The implications of these findings for policy-making, labor market interventions, and strategies aimed at reducing economic inequalities should also be thoroughly examined.

In conclusion, this research paper underscores the intricate interplay between demographic factors and logwages, highlighting avenues for further investigation to ascertain causation, explore additional variables, and delve deeper into the lived experiences underlying these associations. Such an expanded inquiry holds the potential to offer substantial insights for policymakers and stakeholders in addressing economic disparities and fostering equal opportunities in the workforce.

References:

Bol, T., & Heisig, J. P. (2021). Explaining wage differentials by field of study among higher education graduates: Evidence from a large-scale survey of adult skills. Social Science Research, 99, 102594-102594. https://doi.org/10.1016/j.ssresearch.2021.102594

Boudarbat, B., Lemieux, T., & Riddell, W. C. (2006). In Green D. A., Kesselman J. R.,eds (Eds.), Recent trends in wage inequality and the wage structure in canada Vancouver: University of British Columbia Press. Retrieved from https://www.proquest.com/books/recent-trends-wage-inequality-structure-canada/docview/56822673/se-2

Connolly, H. C., & Gottschalk, P. (2006, October 15). Differences in wage growth by education level: Do less-educated workers gain less from work experience?. by Helen C. Connolly, Peter Gottschalk :: SSRN. https://dx.doi.org/10.2139/ssrn.937356

https://www.uagc.edu/blog/income-inequality-and-the-earnings-gap-between-educated-and-non-educated. (n.d.).

Suter, L. E., & Miller, H. P. (1973). Income Differences Between Men and Career Women. *American Journal of Sociology*, *78*(4), 962–974. http://www.jstor.org/stable/2776614

Taylor, P., Fry, R., & Oates, R. (2014). The rising cost of not going to college

The curious case of investment in Education. The Economic Times. (n.d.). https://economictimes.indiatimes.com/news/economy/policy/the-curious-case-of-investment-in-education/articleshow/80306498.cms