Nonparametric Approach Statistical Test for Lorenz Dominance in Income Distribution: A Case Study on Gender Income Inequality

Miang-Hong NGERNG (ngerng.mh@help.edu.my)
HELP University, Malaysia

Abstract

Income inequality measurement is pivotal in economic studies, aiding researchers and policymakers in understanding and addressing disparities. The Lorenz Curve, depicting income distribution, shows deviations from perfect equality, while Lorenz Dominance, introduced by Atkinson, compares the welfare implications of different income distributions. Shorrocks' Generalized Lorenz Curve (GLC) extends this concept, allowing comparisons across distributions with different means, integrating descriptive inequality aspects and fairness considerations.

This paper builds on the approach of Beach and Davidson (1983) by using non-parametric methods to estimate Lorenz ordinates. It develops t-tests and goodness-of-fit tests to assess Lorenz dominance between two income distributions. An empirical study of 1991 Canadian data examines income disparities between male and female workers. Findings indicate significant gender income inequality, influenced by factors like education and age, with the Lorenz and Generalized Lorenz Curves providing a comprehensive view of the distributional structures. Statistical inference confirms the observed disparities, highlighting the continued relevance of gender-based income inequality studies in shaping equitable economic policies.

***Keywords*:** Gender Income Inequality, Generalized Lorenz Curves, Nonparametric Methods

JEL Classification: J16 - Economics of Gender; Non- Discrimination,
C14 - Semiparametric and Nonparametric Methods
D31 - Personal Income, Wealth, and Their Distributions

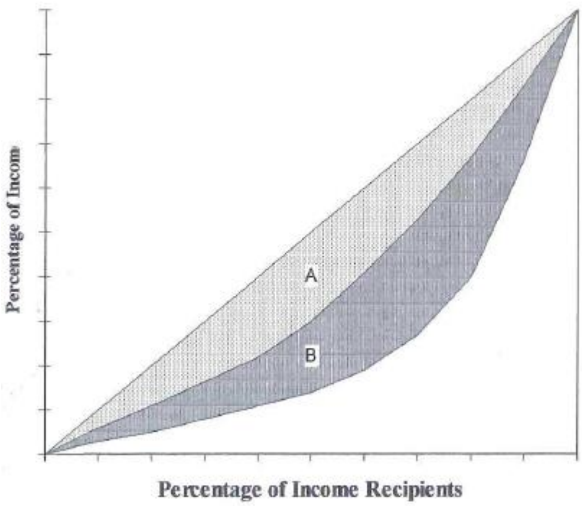
1.0 Introduction

Income inequality measures an important in Economic studies as by employing these measures, researchers and policymakers can gain insights into the nature, extent, and implications of inequality, thereby informing decisions and actions to promote more equitable societies.

Income inequality can be measured by two approaches, Positive measures employ metrics like range, standard deviation, Gini coefficient, and Theil index, statistical tools to quantify income variation without making value judgments. Normative measures incorporate ethical reasoning with a social welfare function that exhibits diminishing returns with inequality, indices such as Dalton and Atkinson are such examples. (Amartya, 1980) differentiates inequality analysis between descriptive, look at factual income differences, and equity aspects that involve ethical judgments about these differences.

The Lorenz Curve shows income distribution by plotting the proportion of total income held by the bottom *x*% of the population. A perfectly equal distribution forms a 45-degree line, with deviations indicating greater inequality. (Atkinson, 1970) introduced Lorenz Dominance to compare welfare implications of income distributions of equal mean. Distribution A dominates B if A's Lorenz Curve is above B's, assuming conditions like diminishing marginal utility. (Shorrocks, 1983) extended this concept with the Generalized Lorenz Curve (GLC), which multiplies the Lorenz Curve by mean income. This allows comparisons between distributions with different means, providing a more comprehensive view of income inequality and social welfare. The GLC integrates descriptive aspects of inequality and some aspects of fairness without additional assumptions.

Figure 1: Dominance of Lorenz curves



The Lorenz curve is typically used as a descriptive tool to illustrate income inequality rather than as an analytical tool for statistical inference. This can be seen in examples like the (Dagum, 1985) analysis of income inequality by education and gender in Canada. The complexity of the mathematics involved in measuring inequality, as documented in the works edited by (Biewen & Flachaire, 2018), and the computational difficulties in applying theoretical statistical expressions to empirical problems, may contribute to this limited use of the Lorenz curve in statistical inference.

Studies have shifted to non-parametric approaches for estimating Lorenz ordinates, restating the problem in a more general form and utilizing adequate statistical procedures. The term "non-parametric," originating from Wolfowitz in 1942 (Noether, 1967), suggests that the underlying population should not be specified by a finite number of parameters. Non-parametric methods have desirable properties, including requiring only a few assumptions about the population and not assuming any specific distribution for the underlying population.

This paper builds on the nonparametric estimation of Lorenz coordinates by (Beach & Davidson, 1983) to derive t-tests and goodness-of-fit tests for assessing Lorenz dominance between two income distributions. The method was applied to the study of income disparity between male and female workers, an issue that has garnered significant attention in both international organizations and the academic arena. Income distribution differences by sex are present in every country, regardless of its stage of development. Differential earnings between men and women have long been a subject of interest, with many empirical studies attributing income inequality to differences in productivity-related factors brought by each individual to the labour market.

Over the past forty years, one of the most significant increases in labour supply has been the rise in female labour participation. Since the early 20th century, the number of female workers has grown dramatically, particularly during the 1960s and 1970s. This shift can be attributed to both economic and social factors. Traditionally, women stayed at home, but the financial benefits of joining the workforce have outweighed this norm. As the cost of staying home has risen with increasing living standards, and declining fertility rates have reduced the benefits of home production, more women have entered the labour force. Social changes in the 1960s also facilitated this influx, leading to greater acceptance of women in the workforce.

The entrance of women into the labour market has altered income distribution in several ways. Firstly, the increase of low-paid women in the labour market initially raised overall income inequality. Secondly, more dual-earning families emerged, reducing inequality compared to single-earning families. Thirdly, the rise in female-headed households, often with lower incomes, increased inequality. Lastly, as occupational segregation decreased, women began competing with men for wages. This study focuses on the last aspect, examining income inequality between men and women.

In their 1973 paper, Malkiel and Malkiel addressed two questions: whether salary structures could be explained by variables like education and experience, and whether women with similar qualifications to men received the same pay. They found that education, experience, and productivity proxies explained about 75% of salary variance for both genders. While men and women at the same job levels received equal pay, women with the same qualifications as men were often assigned lower job levels. Doiron and Barrett (1996) decomposed annual earnings into hours worked and hourly pay to examine gender earnings differentials. Their results indicated that greater female inequality in earnings was due to a more unequal distribution of hours worked among women. Over time, both men and women saw comparable decreases in inequality in hours worked and hourly earnings.

Our study on the gender income gap in Canada in 1991 complements the (Pelletier & Patterson, 2018) study by providing historical context and data prior to their recent analysis. The proposed approach not only allows for a detailed examination of distributional structures but also permits formal statistical inference to assess the statistical significance of Lorenz dominance. The process entails estimating Lorenz curve ordinates and involves computing sample quantiles and conditional means based on order statistics. This provides sample estimates of Lorenz curve ordinates and enables statistical inference through the establishment of asymptotic variance-covariance matrices. The method determines Lorenz curve ordinates for specific deciles, typically using statistical software like SAS for computation and visualization tools like Microsoft Excel for plotting. Additionally, this framework extends to the Generalized Lorenz Curve, offering insights into income distributions beyond mere descriptive purposes.

The methodology will be explained in the next section, followed by an explanation of the findings in Section 3 and concluding remarks in Section 4.

2.0 Methodology

The basic strategy here is to perform a statistical investigation for the degree of inequality captured by the Lorenz curve derived from the distribution. The methodology of Beach and Davidson (1983) is being employed, as introduced in the previous chapter. For the empirical study done in this academic exercise, the data was obtained from the 1991 survey done in Canada.

2.1 Data Description

The data was collected from a random survey of approximately fifteen percent of the population. To ensure manageability, the sample size was reduced to around 20,000, resulting in 24,290 total observations. The dataset includes information on income, gender, age, and education. The ages range from 1 to 85, but only those aged 15 and above are considered for income analysis. Consequently, the minimum age included is 15. The ages are categorized into four groups:

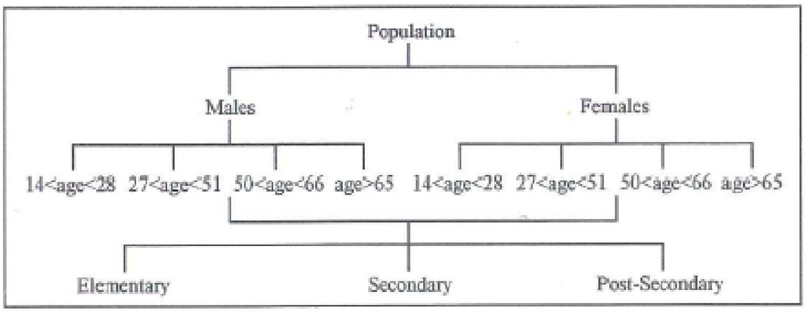
- 15 to 27
- 28 to 50
- 51 to 65
- 66 and above

Educational levels are scaled from one to fourteen and divided into three categories:
- 1 to 6 for elementary schooling
- 7 to 10 for secondary schooling
- 11 to 14 for post-secondary schooling

Gender is coded as:
- 1 for females
- 2 for males

Initially, there were 12,179 females and 12,111 males. After excluding those under 15, the number of females is reduced to 8,543 and the number of males to 8,929.

Figure: Division of Population



2.2. Method of Analysis

Following the estimation method proposed by (Beach and Davidson, 1983), we calculate the Lorenz ordinates, Generalized Lorenz ordinates, and their standard deviations, assuming independence between the compared distributions. We develop a t-test to assess the individual gaps in Generalized Lorenz ordinates between the two gender groups and a chi-square test for the differences of the nine quartiles Generalized Lorenz ordinates collectively.

2.2.1 Estimation of Lorenz Curve and Generalized Lorenz Curve

In the paper by Beach and Davidson (1983), an approach was developed such that it enabled an examination of not only the detailed structure of a given distribution but also extended formal statistical inference to the Lorenz curves.

The Lorenz curves by Beach and Davidson's definition is characterized by a set of ordinates Φ_1, Φ_2, Φ_K corresponding to the quantile proportions p_1 , p_2 , ... , p_K.. Therefore, for decile proportions, $K = 9$ and $P_1 = 0.1$, $P_2 = 0.2$, ..., $P_K = 0.9$.

The sample quantile ξ_{Pi} is defined as the r^{th} order statistic in the random sample Y_1, Y_2 , ... , Y_N , where, arranging in ascending order, Y_1 is the smallest size income, Y_N is the largest size income, and $r_i = [N p_i]$ represents the largest integer less than or equal to $N p_i$. Hence, conforming to the abscissae is a set of K population quantiles $\xi_{P1} < \xi_{P2} <....< \xi_{PK}$ and the Lorenz curve ordinates $\Phi(\xi_{P1}) < \Phi(\xi_{P2}) < .... < \Phi(\xi_{PK})$.

Then $\Phi(\xi_{Pi}) = p_i \, \gamma_i / \mu$ is the corresponding sample estimate of the Lorenz curves ordinates where, μ is the estimated sample mean E(Y), that is $\mu = \sum_1^N y_j/N$ and γ_i is the conditional mean of cumulated income less than or equal to ξ_{Pi} , $E(Y|Y\leq \xi_{Pi})$; that is $\gamma_i = \sum_1^{r_i} y_j/r_i$.

The sample estimate of Lorenz curve ordinate is computed as:

$$\Phi(\xi_{Pi}) = \frac{\sum_{j=1}^{r_i} y_j}{\sum_{j=1}^{N} y_j} = \frac{r_i}{N}\frac{\sum_{j=1}^{r_i} y_j/r_i}{\sum_{j=1}^{r_i} y_j/N} = p_i\frac{\gamma_i}{\mu}.$$

Based on the known results of the sampling distribution of on order statistics, the Lorenz ordinates is multivariate normal with asymptotic variance-covariance matrix given as below.

In the case of deciles, the elements of the variance-covariance matrix denoted as v_{ij} are

$$v_{ij} = \frac{1}{\mu^2}[\Psi_{ij} + \phi_i\phi_j\sigma^2 - \phi_i\Psi_{j,10} - \phi_j\Psi_{i,10}] \quad i \leq j = 1, ..., 9.$$

where

$$\Psi_{ij} = p_i[\lambda_i^2 + (1-p_i)(\xi_{Pi} - \gamma_j) + (\xi_{Pi} - \gamma_i)(\gamma_j - \gamma_i)] \quad i \leq j = 1, ..., 9,$$

$$\Psi_{i,10} = p_i[\lambda_i^2 + (\xi_{Pi} - \gamma_i)(\mu - \gamma_i)] \quad i = 1, ..., 9,$$

σ^2 is the weighted sample variance, and

λ_i^2 is the variance of earnings in deciles 1 to i.

With the results, the Lorenz ordinate is calculated for each individual decile the Lorenz curve can be plotted by Microsoft Excel with ordinate vs decile. The Generalized Lorenz curve, is defined as $GL(\xi_{Pi}) = \mu \, \Phi(\xi_{Pi})$. With the Lorenz curve ordinates already complied, the Generalized Lorenz ordinates can also be computed for each individual decile as well. Hence the Generalized Lorenz curve is plotted in the same way.

The values for Lorenz and Generalized Lorenz curve ordinates, for example for males, ages from 15 to 27, are calculated and tabulated in table 1.

Table 1 Lorenz and Generalized Lorenz curve ordinates: (Males, Age from 15 to 27)

i	P_i	$E(Y\|Y\leq \xi_{Pi})$	$E(Y)$	$\Phi(\xi_{Pi})$	$GL(\xi_{Pi})$
0	0.0	0.00	14,970.86	0.0000	0.00
1	0.1	705.42	14,970.86	0.0047	70.54
2	0.2	1,431.36	14,970.86	0.0191	286.27
3	0.3	2,310.81	14,970.86	0.0463	693.24
4	0.4	3,346.56	14,970.86	0.0894	1,338.62
5	0.5	4,599.83	14,970.86	0.1536	2,299.92
6	0.6	6,064.42	14,970.86	0.2430	3,638 65
7	0.7	7,750.71	14,970.86	0.3624	5,425.50
8	0.8	9,622.98	14,970.86	0.5142	7,698.38
9	0.9	11,734.09	14,970.86	0.7054	10,560.70
10	1.0	14,970.86	14,970.86	1.0000	14,970.86

Here, $\Phi(\xi_{Pi}) = p_i \frac{\gamma_i}{\mu}$ and $GL(\xi_{Pi}) = E(Y) \times \Phi(\xi_{Pi}) = p_i\gamma_i$.

The standard deviation for the whole sample and the conditional standard deviation for which is less than or equal to ξ_{Pi} can be computed using method in Kakwani, N.C. (1990) as:

$$\sigma_p^2 = \frac{p}{\mu^2}[\lambda_p^2 + (1-p)(y_p - \frac{\mu\Phi(\xi_P)}{p})^2] + (\frac{\Phi(\xi_P)}{\mu})^2\sigma^2 - 2\frac{p\Phi(\xi_P)}{\mu^2}[\lambda_p^2 + \frac{\mu\phi(p)}{p}(y_p - \frac{\mu\Phi(\xi_P)}{p})]$$

for $0.1 < p < 0.9$ and $\phi(p) = p - L(p)$.

The values for the standard deviations, for example males, ages 15 to 27 are calculated and tabulated as table 2 below.

Table 2 Standard Deviations: (Males, Age from 15 to 27)

i	P_i	$\sigma^2 (Y\|Y\leq \xi_{Pi})$	σ_{pi}^2
1	0.1	388.46	8.5459E-05
2	0.2	845.21	0.00092
3	0.3	1,473.15	0.00445
4	0.4	2,247.32	0.01416
5	0.5	3,237.25	0.03614
6	0.6	4,435.71	0.07699
7	0.7	5,858.90	0.14050
8	0.8	7,403.92	0.21791
9	0.9	9,224.90	0.28017

2.2.2 Test for Lorenz Dominance Between Lorenz Curve and Generalized Lorenz Curve of two gender groups

The hypothesis is $H_0 : L_M = L_F$ vs $H_1 : L_M - L_F > 0$

where L_M is the Lorenz ordinates for males' group and L_F is the Lorenz ordinates for females' group. Using the computed values of the Lorenz ordinates, Generalized Lorenz ordinates, and their standard deviations, we can perform t-tests on the individual Lorenz ordinates for the two gender groups as independent samples.

$$t_i = \frac{\hat{\Phi}_M(\xi_{Pi}) - \hat{\Phi}_F(\xi_{Pi})}{\sqrt{\frac{\hat{\sigma}_M^2}{N_M} + \frac{\hat{\sigma}_F^2}{N_F}}} \quad \text{for } i = 1, 2, ..., 9,$$

where $\hat{\Phi}_{M,F}(\xi_{Pi})$ is the Lorenz ordinate estimates for the decile p_i for $i = 1, 2, ..., 9$; and

M = Male or F= female

$\hat{\sigma}_{M,F}^2$ is the standard deviation of Lorenz ordinates estimates for males or females, and

$N_{M,F}$ is the total number of males or females group in the respective decile.

To compare the Lorenz Curve of the two gender group as a whole for Lorenz dominance, we use the chi-square statistics.

$\chi^2 = \sum_{i=1}^{9} t_i^2 \sim \chi_9^2$ (chi-square distribution with 9 degree of freedom).

The level of significance used is five percent, which will be tested against the t-statistics as described. If the t-test is significant and H_0 is rejected, this implies that an inequality exists between men and women, suggesting that the income for men may be greater than that for women, ceteris paribus.

The Lorenz curve is plotted as a descriptive tool, providing a visual representation of the wage differential between males and females. Statistical inference using the t-statistic or the chi-square statistic is then performed to validate the observations from the graph.

3.0 Findings

3.1 Descriptive Statistics of Various Groups

Tables 3, 4 and 5 provide descriptive statistics for males and females in the sample data collected in Canada for the year 1991. The mean income is also included in the three tables.

Table 3 : Descriptive Statistic for the Sample

	Nos.	Percentage	Mean Income
Males	8929	51.10%	29,840.57
Females	8543	48.90%	17,571.34
Total	17472	100.00%	23,841.49

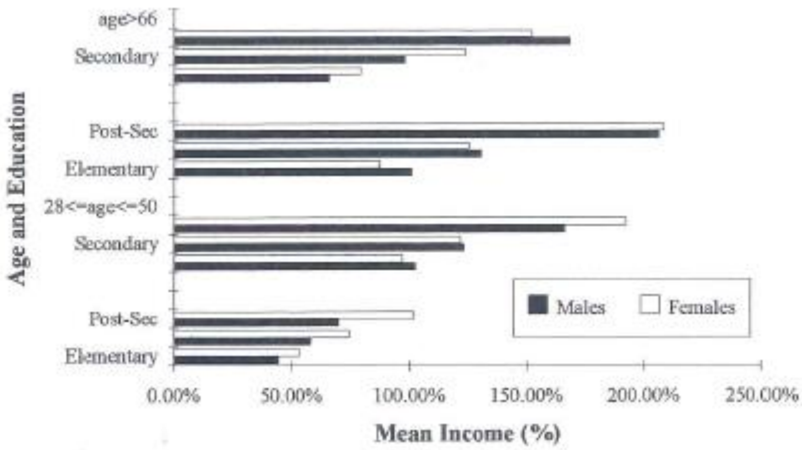
Table 4: Descriptive Statistic for Males

Age	Education	Nos	Percentage	Mean Income
15 to 27	Elementary	1282	66.56%	13,331.42
	Secondary	505	26.22%	17,484.53
	Post-Sec	139	7.22%	20,959.10
	Total	1926	100.00%	14,970.86
28 to 50	Elementary	2213	50.72%	30,604.99
	Secondary	1330	30.48%	36,875.23
	Post-Sec	820	18.79%	49,582.42
	Total	4363	100.00%	36,083.08
51 to 65	Elementary	1099	69.60%	30,219.53
	Secondary	304	19.25%	39,094.37
	Post-Sec	176	11.15%	61,637.88
	Total	1579	100.00%	35,430.16
≥ 66	Elementary	854	80.49%	19,700.65
	Secondary	143	13.48%	29,349.85
	Post-Sec	64	6.03%	50,257.31
	Total	1061	100.00%	22,844.35

Table 5: Descriptive Statistic for Females

Age	Education	Nos	Percentage	Mean Income
15 to 27	Elementary	1110	57.99%	9,417.11
	Secondary	616	32.18%	13,098.78
	Post-Sec	188	9.82%	17,947.81
	Total	1914	100 00%	11,439.93
28 to 50	Elementary	2066	52.76%	17,097.08
	Secondary	1247	31.84%	21,444.26
	Post-Sec	603	15.40%	33,858.30
	Total	3916	100.00%	21,062.34
51 to 65	Elementary	960	72.18%	15,405.94
	Secondary	275	20.68%	22,061.47
	Post-Sec	95	7.14%	36,674.81
	Total	1330	100.00%	18,301.29
≥ 66	Elementary	1151	83.22%	14,001.75
	Secondary	187	13.52%	21,784.95
	Post-Sec	45	3.25%	26,784.93
	Total	1383	100.00%	15,470.08

Figure 3: Income Spread: Age and Education



Regarding the life-cycle effect, mean income is lowest for those aged 15 to 27, likely due to a lack of job experience. There is a significant increase in mean income from the first to the second age group. For those aged 51 to 65, incomes either remain stable or increase. Finally, for those aged 66 and above, mean income decreases as they begin to retire. This observed pattern aligns well with the life-cycle theory, which posits that individuals enter the workforce with a low income, see their income rise and peak as they age, and then experience a decline as they retire.

There is a strong correlation between income and education level, with the highest income typically associated with the highest levels of education and skill. Individuals who invest in their education, incurring opportunity costs in terms of wages or time, are expected to earn higher salaries to compensate for their educational investment. Additionally, different fields of education lead to different occupations, resulting in varied salary levels.

The observed pattern supports both the life-cycle theory and the human capital theory of education. Regardless of gender, individuals with higher education levels tend to earn more within each age group. However, However, by judging on the mean income, is it really true that income inequality exists, even though they are of the same age group and educational level? To determine if income inequality truly exists between males and females of the same age group and educational level, we will apply the Lorenz dominance criterion to ascertain whether the disparity in income between genders is statistically significant or due to random chance.

3.2 Lorenz Curves and Generalized Lorenz Curves

The use of the Lorenz curve is one of the most common methods for describing the income distribution graphically. However, if the Lorenz curves cross, it is then difficult to explain the existence of income inequality. Hence the Generalized Lorenz curve is plotted to allow for the possibility of Lorenz curves intersecting.

3.2.1 Lorenz Curve

The Lorenz curve illustrates the quantitative relationship between the percentage of income recipients and the percentage of total income they receive over a given period, in this case, a year. Figures 4 to 7 display the Lorenz curves using the computed decile data from Canada in 1991. Due to page limitations, only selected Lorenz curve figures are presented.

Figure 4: Lorenz curve: 28 ≤ Age ≤ 50

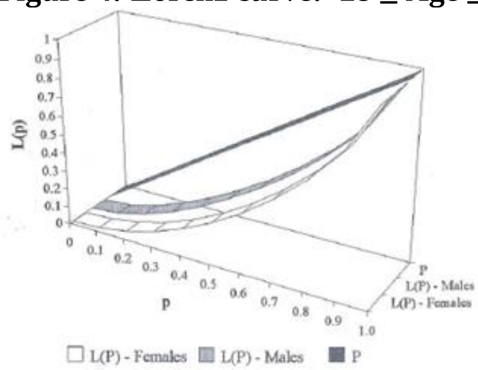


Figure 5: Lorenz curve: Elementary Education

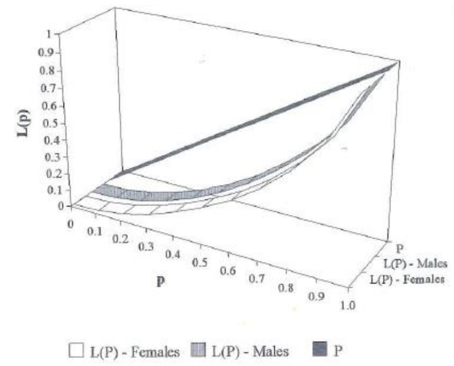


Figure 6: Lorenz curve: 51 ≤ Age ≤ 65 and Elementary Education

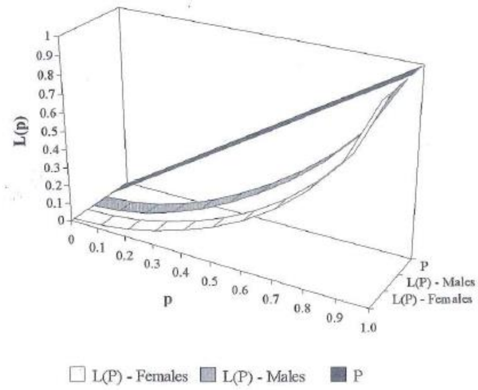
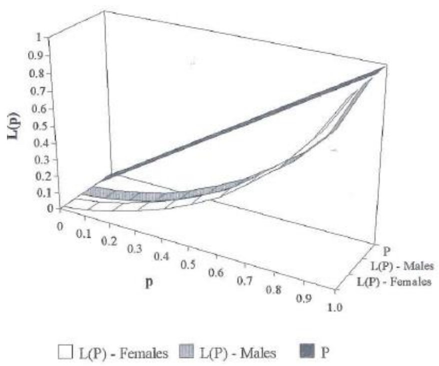


Figure 7: Lorenz curve: Age ≥ 65 and Post-Secondary Education



The previous four figures show the Lorenz curves for various age groups and education levels. In Figures 4 to 6, the Lorenz curves for both females and males appear to converge as they approach the tenth decile. However, determining whether the Lorenz curves actually intersect is challenging. Figure 7 clearly shows the Lorenz curves crossing, starting around the sixth decile. The Lorenz curve provides a useful visual aid for comparing inequality between males and females. From Figures 4 to 7, it is observed that initially, the Lorenz curves for males dominate those for females, indicating that the area between the Lorenz curve for males and the 45-degree line is smaller than that for females. However, as they approach the last decile, the Lorenz curves appear to intersect. This pattern holds true for different combinations of age groups and educational levels.

Due to differences in mean income levels, it may be difficult to ascertain income inequality between males and females based on Lorenz curve dominance alone. Therefore, the

Generalized Lorenz Curve, as introduced by (Shorrocks, 1983), is plotted to provide a more comprehensive comparison.

3.3.2 Generalized Lorenz Curves

If the Lorenz curves intersect, as seen in Figures 4 to 7, can a criterion based on mean income be established to achieve unanimous welfare? The answer is yes, and it lies in the development of the Generalized Lorenz Curve. The Generalized Lorenz Curve is created by scaling up the ordinary Lorenz Curve by the mean of the distribution. Figures 8 to 11 display the Generalized Lorenz Curves plotted using data from the same groups as the Lorenz Curves.

Figure 8: Generalized Lorenz curve: 28 ⩽ Age ⩽ 50

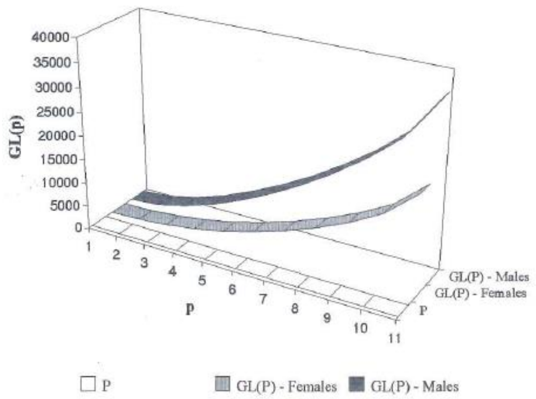


Figure 9: Generalized Lorenz curve: Elementary Education

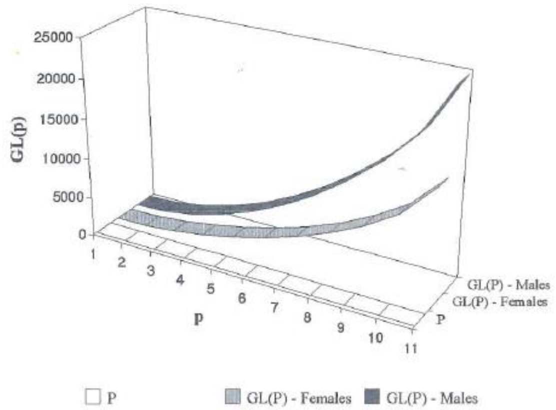


Figure 10: Generalized Lorenz curve: 51 ⩽ Age ⩽ 65 and Elementary Education

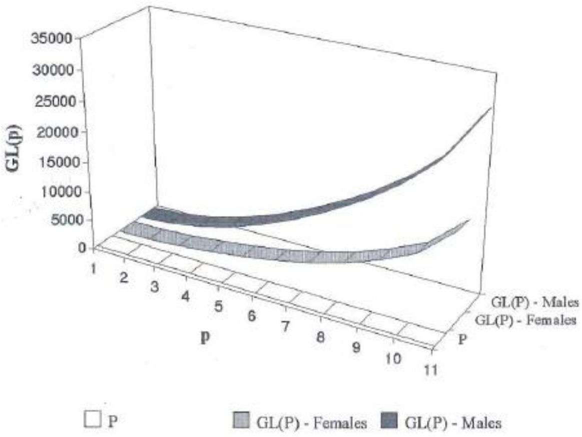
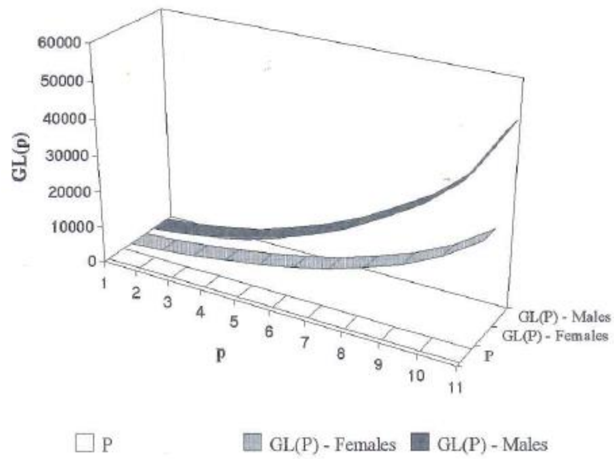


Figure 11: Generalized Lorenz curve : Age > 65 and Post-Secondary Education



The four figures above show no crossing between the Generalized Lorenz Curves for men and women. The income inequality between gender groups is illustrated even more clearly using the Generalized Lorenz Curve criterion, highlighting that females earn less than males. A statistical test can now be used to formally confirm our graphical observation.

3.3.3. Statistical Test

To enhance the analysis with a rigorous approach, three statistical hypothesis tests were developed and applied to the Lorenz ordinates. Tables 6, 7, and 8 present the statistical test results for individual Lorenz ordinates, comparing male and female distributions.

Table 6: Statistical Test results on Lorenz Ordinates: Age

Age	Significant at 5%
15 to 27	Y
28 to 50	N
51 to 65	Y
66 and above	Y

Table 7: Statistical Test results on Lorenz Ordinates: Education

Education	Significant at 5%
Elementary	N
Secondary	Y
Post-secondary	Y

Table 8: Statistical Test results on Lorenz Ordinates: General

Age	Education	Significant at 5%
15 to 27	Elementary	Y
	Second	Y
	Post-secondary	Y
28 to 50	Elementary	Y
	Second	Y
	Post-secondary	Y
51 to 65	Elementary	N
	Second	Y
	Post-secondary	Y
66 and above	Elementary	Y
	Second	Y
	Post-secondary	Y

In summary, the analysis reveals a significant income disparity between male and female workers, as evidenced by the differences in estimated Lorenz ordinates. However, three exceptions emerge.

First, the t-test for the 28-50 age group shows no significant difference, suggesting potential income equality within this group. However, Table 6 indicates significant income gaps within different education levels for this age group. This suggests that the test's power may have been masked by the wide variation in income due to education.

Second, the test statistic is also insignificant for individuals with an elementary education level. This is likely due to the low-income variation within this occupational group.

Finally, the 51-65 age group with an elementary education level again shows no significant difference. This might be for the same reason as the second exception.

These findings might cautiously suggest a general trend of income equality between genders. However, it's important to consider studies like (Pelletier et al., 2018) which show a pattern of narrowing income gaps, albeit not entirely eliminated, even up to 2020. Further research is needed to solidify these observations.

4.0 Summary and Concluding Remarks

4.1 Summary and Findings

The mean income was first used to compare the income distribution of males and females. To provide a graphical illustration, the Lorenz curves and the Generalized Lorenz curves are plotted. Finally, a statistical inference was made. The use of the mean income to

compare the difference in income between males and females is vague. Since the Lorenz curves intersect, it is hard to tell whether an inequality actually exists. The Generalized Lorenz curve is therefore plotted, revealing an inequality in income between the gender groups. The statistical test is then used to obtain more scientific judgment with statistical inference.

The Lorenz curve served as a descriptive tool for income inequality. (Atkinson, 1970)'s work linked Lorenz curves with social welfare functions, enabling normative interpretations. (Shorrocks, 1983)'s extension allowed for Lorenz dominance criterion application even with unequal means. The Lorenz curve ordinates were calculated using Beach and Davidson's (1983) distribution-free approach, offering advantages like avoiding distributional assumptions and facilitating statistical comparisons.

This study investigated gender income disparity based on the assumption of equal pay for equal ability. The study defines ability through qualifications (education level) and experience (approximated by age). The Lorenz dominance criterion was used for comparisons.

The analysis suggests a general income inequality between genders. However, three exceptions emerged:

- Individuals aged 27 to 51
- Individuals with only elementary education
- Individuals aged 50 to 66 with only elementary education

The results suggest that the disparity diminishes for older, less-educated workers. Income inequality between genders may be less pronounced for this demographic.

4.2 Limitations and Recommendations for Future Research

The Lorenz dominance criterion, employed in this study, offers a partial ranking of income distributions, not a complete one. It assumes non-intersecting Lorenz curves, which wasn't always the case here. However, the Generalized Lorenz curves addressed this issue.

A major limitation is the data size, representing only a small portion of the 1991 Canadian population. Limited data access or participant reluctance often restricts sample size. While random sampling minimizes bias, a larger sample could potentially refine the results.

Time and space constraints prevented comparisons of income inequality across time periods. Investigating income distribution changes for males and females over several years would enhance the research.

Canada's diverse geography, encompassing rural and urban areas, was not considered. Studies by (Lecaillon et al., 1984) suggest income differences between genders in these areas. Future research could incorporate geographical variations.

The study's focus on Canada limits generalizability to other countries with varying population sizes, male-female ratios, and development stages. Future research could explore more suitable measures and estimation procedures conducive to straightforward statistical inference. However, no single ideal measure exists for income inequality. Different measures serve different purposes. High-quality data and comprehensive, accurate measures are crucial for obtaining better results.

Acknowledgment: The data for this analysis is provided by Michael Hoy, Department of Economics and Finance, Gordon S. Lang School of Business, MacKinnon Building, University of Guelph, Guelph, N1G2W1, Canada

Reference

Atkinson, Anthony B. (1970) "On the Measurement of Inequality". Journal of Economic Theory, Volume 2, Issue 3, September 1970, Pages 244-263 https://doi.org/10.1016/0022-0531(70)90039-6

Beach, C.M. and Davidson, R. (1983) "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares", Review of Economic Studies 50(4):723-35 https://doi.org/10.2307/2297772

Dagum, C. (1985) "Analyses of Income Distribution and Inequality by Education and Sex in Canada". Advances in Econometrics, Vol. 4: Economic Inequality: Survey Methods and Measurements Editors: R.L. Basmann (Texas A&M U), George F. Rhodes, Jr. (Colorado State U). Published in 1985. JAI PRESS INC. Greenwich, Connecticut ISBN: 0-89232-580-1

Doiron, Denise J & Barrett, Garry F, (1996). "Inequality in Male and Female Earnings: The Role of Hours and Wages," The Review of Economics and Statistics, MIT Press, vol. 78(3), pages 410-420, August. https://doi.org/10.2307/2109788

Kakwani, N. (1990). Large Sample Distribution of Several Inequality Measures: With Application to Côte d'Ivoire. In: Carter, R.A.L., Dutta, J., Ullah, A. (eds) Contributions to Econometric Theory and Application. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-9016-9_3

Lecaillon, J., Paukert, F., Morrisson, C., and Germidis, D. (1984) Income Distribution and Economic Development: An Analytical Survey, Geneva: International Office ISBN: 9789221033660

Malkiel, Burton G. and Judith A. Malkiel (1973) "Male-Female Pay Differentials in Professional Employment" American Economic Review Vol. 63, No. 4 (Sep., 1973), pp. 693-705 https://www.jstor.org/stable/1808858

Noether, G.E. (1967) Elements of Nonparametric Statistics, New York, London, Sydney: John Wiley & Sons Inc. SIAM series in applied mathematics, ISSN 0558-0293

Pelletier, Rachelle and Martha Patterson (2018), “The gender wage gap in Canada: 1998 to 2018”, in Biewen, Martin (Ed.); Flachaire, Emmanuel (Ed.) (2018): Econometrics and income inequality Centre for Market Information, Statistics Canada and Melissa Moyser, Centre for Gender, Diversity, and Inclusion, Statistics Canada, ISBN 978-3-03897-367-6, MDPI, Basel, https://doi.org/10.3390/books978-3-03897-367-6

Sen, Amartya (1972) On Economic Inequality, Oxford: Clarendon Press. ISBN: 9780198281931

Sen, A (1980) Description as Choice, Oxford Economic Papers, Volume 32, Issue 3, November 1980, Pages 353–369, https://doi.org/10.1093/oxfordjournals.oep.a041484

Shorrocks, Anthony F. (1983) "Ranking Income Distributions". Economica, Vol. 50, No. 197 (Feb., 1983), pp. 3-17 (15 pages), John Wiley, https://doi.org/10.2307/2554117