A cover of a report

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Student ID** | **First**  **name** | **Parts**  **contributed** | **Contribution %** | **Signature** |
| s3938354 | Nhu | Part 2 | 100% | Không có mô tả. |
| s4000175 | Ly | Part 1, 4 | 100% | A black signature on a white background  Description automatically generated |
| s3837364 | Quang | Part 3 | 100% |  |
| s3877030 | Tram Anh | Part 5 | 100% | Không có mô tả. |
| s3926760 | Quang Anh | Part 3 | 100% | Không có mô tả. |

[PART 1: INTRODUCTION AND LITERATURE REVIEW 3](#_Toc76931130)

[Introduction 4](#_Toc2024204174)

[Literature review 4](#_Toc1554443412)

[Research objectives 5](#_Toc1508406456)

[PART 2: DATA CLEANING AND DATA DESCRIPTION 5](#_Toc305011152)

[Descriptive Statistics 5](#_Toc1921794487)

[Data Visualization 7](#_Toc4604970)

[Missing Values and Outliers 7](#_Toc1257482298)

[Mathematical Transformations 8](#_Toc1831451747)

[PART 3: MODEL SPECIFICATION AND ANALYSIS 8](#_Toc347408280)

[Model specification 9](#_Toc425343936)

[Estimation and interpretation 10](#_Toc2076762320)

[PART 4: CONCLUSION AND POLICY IMPLICATIONS 15](#_Toc427200071)

[Conclusion and recommendations 15](#_Toc922677831)

[Analysis limitations 16](#_Toc869289387)

[PART 5: INDUSTRIAL TALK 16](#_Toc1368952726)

[REFERENCES 17](#_Toc476073218)

[APPENDIX 22](#_Toc1188432601)

PART 1: INTRODUCTION AND LITERATURE REVIEW

Introduction

Economic globalisation, defined as the interdependence between nations resulted by demolishing trade barriers, increasing capital flow and spreading technologies (Gao 2000), is often seen as the key to global economic prosperity. By adapting to globalisation, developing countries can accelerate wealth creation by receiving technology from developed countries and exporting goods and services, thus closing the global income gap between rich and poor countries. However, their labour force is often exploited by developed countries for wealth hoarding, leading to the widening of disparity. The relationship between globalisation and income inequality remains unambiguous, hence the need to understand the linkages between globalisation and income inequality in both developed and developing countries, and other factors conducive to income inequality to address and mediate the situation, setting a path for global economic prosperity.

## **Literature review**

Income disparity is most measured by the Gini coefficient which represents the dispersion of income held by the bottom x% of the population compared to the standard 45° angle (Gini 1912). It is also known to have increased throughout global economic growth, partially influenced by globalisation (Milanovic 2022; Gozgor and Jha 2015). As such, much literature was devoted to the theorised relationship between globalisation and income inequality, such as the Stolper–Samuelson theorem (1941), stating that welfare would be worse off in a developed nation as opposed to a developing nation in the event of increased international trade. While some empirical research has concluded with evidence that globalisation contributes to the reduction of inequality by transfer of wealth from rich areas to poor areas (Goldberg and Pavcnik 2007; Wei and Wu 2001), others found that globalisation enhances disparity by financially benefiting some but not others, resulting in enlarged income gap (Archaryya and Marjit 2000. Bardhan (2007) has illustrated difficulty in establishing a causal relationship between inequality and globalisation, given diverse market conditions different factors of production combined with countries’ varied approaches to trade reform.

Regarding other factors that may influence income inequality, works of Fellman (1976) and Jakobsson (1976) suggest that inequality may be reduced by progressive income tax. Furthermore, education is found to have a significant inverse influence on income inequality by Yang, Qiu (2016), while unemployment, inflation and GDP growth has been found to have a positive correlation with Gini coefficient (Rubin and Segal 2014; Cysne 2009; Erosa and Ventura 2002).

Research objectives

This paper aims to help establish a clearer understanding of how globalisation might be related to inequality by constructing a multiple regression model to investigate the effect of globalisation, as well as other factors that influence income inequality. Based on Gao’s (2000) definition, globalisation can be speculated as measured by a country’s net exports per GDP or Foreign Direct Investment (FDI) inflow volume per GDP. Based on conducted literature review, a relationship is inconclusive between either of these indicators and income inequality.

Trade tax, hypothesised as a trade barrier in this analysis, can also be used for a two-tailed test regarding its effects on income inequality. Other factors hypothesised to have a positive correlation with inequality include inflation, unemployment, GDP growth. Contrarily, factors hypothesised to have a negative correlation with income inequality include education and income tax. The data of 162 countries is therefore taken from the World Development Indicators Database of World Bank, with all observations recorded in the year 2020.

PART 2: DATA CLEANING AND DATA DESCRIPTION

Descriptive Statistics

There are a total of 162 nations included in the initial analysis with nine significant variables. The variables covered in the list are the Gini coefficient (gini), inflation, unemployment rate (unemployment), GDP growth rate (gdpgrowth), net exports per GDP (netexppergdp), FDI per GDP (fdipergdp), income taxes (incometax), international trade’s taxes (tradetax), and the mean years of schooling (education).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Median | Mean | Standard Deviation | Min | Max |
| gini | 35.6% | 37.0% | 7.5% | 24.6% | 59.1% |
| inflation | 1.9% | 8.3% | 46.8% | -2.6% | 557.2% |
| unemployment | 6.1% | 8.2% | 5.7% | 0.1% | 28.0% |
| gdpgrowth | -3.7% | -4.5% | 8.4% | -54.2% | 43.5% |
| netexppergdp | -4% | -7% | 13% | -56.2% | 37.4% |
| fdipergdp | -1% | -2% | 13% | -74.0% | 111.7% |
| education | 10.50 | 9.91 | 2.93 | 1.7 | 14.1 |
| incometax | 24.1% | 25.0% | 12.3% | 1.4% | 64.5% |
| tradetax | 3.7% | 6.7% | 7.9% | 0.0% | 40.2% |

***Table 2.1:*** Descriptive statistics of significant variables

As for **Gini**, its median and mean values were approximately **36%** across 162 nations in our dataset for 2020. This suggests the presence of a moderate degree of income inequality within global populations.

Regarding **inflation**, this variable **averaged approximately 8.28%** throughout countries worldwide during 2020. The elevated level is due to this variable’s outliers of several countries, specifically Angola, Ethiopia, Haiti, Lebanon, Sudan, and Zimbabwe, where inflation rates exceeded 20%.

In 2020, the global **unemployment rate's median** was approximately **6.11%**. This observed level surpassed the natural rate of 4.5%, representing the minimum level of employment that an economy can maintain without inducing inflation (Amadeo 2021).

Given the elevated unemployment levels, a weak labour market was the inevitable outcome in 2020. During this period, the **average global GDP saw a contraction of 4.52%**. The shrinkage resulting from the COVID-19 crisis has led to the most severe economic slump since the Great Depression (Siegel et al. 2021).

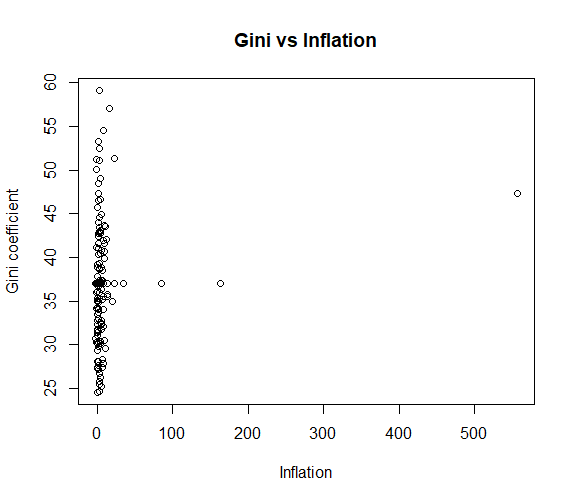
Regarding **the net exports per GDP**, its average value was approximately **–7%** across 162 countries in 2020. This negative figure reflects that nations across the globe saw limitations on fully engaging in international trade and export their goods and services.

The average value of **FDI per GDP** across nations worldwide in 2020 was **–2%,** indicating that the outflow of FDI from a country exceeds the inflow of FDI compared to its GDP. The adverse statistic was mainly due to the substantial decrease of 42% in worldwide FDI (UNCTAD 2021).

The global **income tax** **exhibits substantial variation** among the 162 nations for 2020. The observed variation primarily arises from disparities in the political climate and social welfare policies across different countries (OECD 2020). Similarly, **trade tax** alsovaries from country to country, ranging from negative to positive in 2020.

Data Visualization

***Gini vs. inflation***



***Figure 1:*** Scatter plot of the Gini and inflation

The scatter plot depicted in Figure 1 illustrates a negative correlation between the Gini coefficient and inflation, highlighted by the downward trend observed among the data points. In the context of 2020, an increase in inflation would lead to a fall in income inequality's level among countries. According to Monnin (2014), a spike in inflation can result in inequality reduction until it approaches a nadir at an inflation rate of approximately 13% before it begins to rise again.

Missing Values and Outliers

Within the dataset, there exist 45 nations with incomplete data pertaining to the Gini coefficient, while 7 countries exhibit missing data for unemployment and inflation rates. Furthermore, a total of 48, 68, and 61 countries have missing data related to income tax, trade tax, and education, respectively. One approach to addressing the issue of missing values is to substitute the missing values with their respective mean values.

The dataset has **numerous outliers** in each independent variable. Specifically, these variables, involving gini, inflation, unemployment, gdpgrowth, netexppergdp, fdipergdp, incometax, tradetax, and education, exhibit outliers, with respective counts of 10, 12, 8, 9, **2**, **25**, 13, 16, and 24. Nevertheless, the three indicators, namely inflation, **fdipergdp**, and trade tax each possess distinct outliers considered influential points, as they exhibit unusuality and possess big values. Hence, we should eliminate any observations containing influential data points to enhance the assessment’s accuracy of the correlation between these variables and the Gini.

*The excluded countries are listed in the table below:*

|  |  |  |
| --- | --- | --- |
| **Variables** | **Observations/Countries** | **Influential outliers** |
| **inflation** | Lebanon | 84.86% |
| Sudan | 163.26% |
| Zimbabwe | 557.20% |
| **fdipergdp** | Malta | -74.02% |
| Mauritius | 111.69% |
| **tradetax** | Namibia | 40.23% |

***Table 2.2:*** Countries with influential points

Mathematical Transformations

Our team has implemented the log transformation on variables with positive values. Nevertheless, by recreating the boxplots alongside scatter plots of the variables, the logarithmic transformation does not enhance the linearity between the dependent and independent variables, nor does it increase the normality of the data. An alternative approach is to employ the inverse hyperbolic sine transformation (IHST) (Brauw 2020). However, after applying the IHST to both the positive and negative variables, the results fail to match our desired criteria, as with the log transformation. This failure is due to the distortions introduced by the IHST for values below 10. Hence, we should not employ mathematical transformations on our dataset.

PART 3: MODEL SPECIFICATION AND ANALYSIS

Model specification

1. Dependent variable

A common indicator of economic inequality in a nation or population is the Gini coefficient; the higher the coefficient, the greater the inequality (Bussin 2022). According to the International Monetary Fund (2022), the coefficient ranges from 0 to 1, with 0 denoting complete equality and 1 denoting perfect inequality. Governments may access and analyze wealth and income disparity between nations or historical periods using the Gini coefficient, which offers insightful view on inequality (Cingano 2014). Therefore, the Gini coefficient is the ideal choice of dependent variables to illustrate how income inequality is measured.

1. Independent variables

Inflation, unemployment, GDP growth, FDI, net export, trade tax, education and income tax are determined to be factors that affect inequality as the following discusses.

Income inequality gains from inflation and unemployment, according to Jantti and Jenkins (2010). The income gap between rich and poor grows since inflation lowers poor people's purchasing power. Like how unemployment affects higher-income groups' wealth, it limits lower-income groups' revenue sources.

According to Majumdar and Partridge (2009), economic growth may worsen income inequality because it is frequently associated with higher investments, employment levels, and employment-generating processes, which increase access to jobs and income for more people in both rural and urban areas and decrease income inequality.

Nam HJ and Ryu D (2023) also note that foreign direct investment (FDI) boosts economic development by producing jobs and growth possibilities, which impacts society. Cho and Ramirez (2016) show that FDI decreases income inequality more in nations with well-developed absorptive capacity. However, FDI may damage income distribution in host nations with limited absorptive ability.

According to Nguyen and Su (2022), export quality has a negative effect on income inequality in low- and middle-income countries while having a favorable effect in high-income nations. Thus, net export per GDP is considered as independent variables.

According to Tanzi and Zee (2004), the international trade tax has two adverse consequences on income inequality. Particularly, the higher trade tariff may shield local businesses from international rivalry, resulting in employment growth in certain sectors. However, it may weaken industrial competition, which would slow down economic expansion.

The root causes of opportunity disparity may be addressed via education. The key to assisting nations in achieving long-term economic and social success will be to increase investment in better and more relevant education (OECD 2021).

According to Luebker (2011), most tax systems need higher income tax rates for those with high earnings than for people with low incomes. Therefore, progressive taxation decreases income inequality by reducing the difference between wealthy and poor families.

1. Population model specification

Based on the specified dependent and independent variables, a regression model for the population (henceforth known as Model 1) can be written as follows:

***Gini = β0 + β1\*inflation + β2\*unemployment + β3\*GDP growth + β4\*FDI per GDP+ β5\*net export per GDP+ β6\*trade tax + β7\*education + β8\*income tax +***

Estimation and interpretation

Null Hypothesis: *β2 = 0 (*independent variables have no effect on Gini index)

Alternative two-tailed Hypothesis *β2* ≠ 0 (independent variables have effect on Gini index)

The rejection region in two side is the significance level of 0.01, and the t-distribution is used as a critical value.

As such, the results for Model 1 estimation were given as below:

*Gini = 37.9+ 0.08\*inflation + 0.4\*unemployment – 0.02\*GDP growth + 7.8\*10-12\*FDI*

*(2.56)*  *(0.09)*  *(0.084)*  *(0.05)*  *(1.66\*10-11)*

*+ 0.037\*net export + 0.224\*trade tax - 0.83\*education + 0.09\*income tax +*

(0.037)  (0.08)  (0.2)  (0.043)  
n = 156, R-squared = 0.278, Adjusted R2 = 0.2387

T-statistics values for Model 1 are calculated below based on R results:

|  |  |
| --- | --- |
|  | **Dependent variable** |
|  | **Gini index** |
| **Explanatory variables** | **t-statistics** |
| Inflation | 0.737 |
| Unemployment | 4.922 |
| GDP growth | 0.090 |
| FDI | 1.172 |
| Net export | 0.553 |
| Trade tax | 2.747 |
| Education | -4.220 |
| Income tax | 2.272 |
| Significance level: α = 0.01  Degrees of freedom: d.f = 147  Critical t value (based on R result): ± 2.06 | |

***Figure 7:*** t-statistics of independent variables for Model 1

***Conclusion***

* t-statistics of unemployment (4.922), trade tax (2.747), and income tax (2.272) are greater than critical t value (2.06), we reject *H0.* Thus, unemployment, trade tax and income tax are statistically significant at 1% significance level and have effect on Gini coefficient.
* t-statistics of education (-4.220) are lower than critical t value (-2.06), we reject *H0*. Consequently, education is statistically significant at 1% significance level and have effect on Gini coefficient.
* We cannot reject H0 since t-statistic of inflation (0.737), GDP growth (0.090), FDI (-1.17), and net export (1.172) are more than -2.06 and lower than 2.06. Thus, at 1% significance level, inflation, GDP growth, FDI and net export does not affect Gini coefficient.

Interpretation

* Education has a coefficient of -0.83, which highlights the negative correlation between the Gini coefficient and education. The Gini coefficient will typically rise by 0.83 points for every point in education drop, ceteris paribus,
* GINI is positively correlated with unemployment, trade tax, and income tax coefficients of 0.4, 0.224, and 0.09, respectively. Thus, GINI will rise by 0.4, 0.224, and 0.09 points for each point increase of trade tax, income tax and unemployment, respectively.

Multicollinearity and Heteroskedasticity Evaluation for Model 1

A literature review and statistical tests to establish significant variance can be used to build a Gini model, although multicollinearity and heteroskedasticity must be investigated. This study determines if the report has the best linear unbiased estimator (BLUE) and evaluates variable significance results.

Create a matrix of correlation coefficients between variables to determine model multicollinearity. A study by Andrew University (2005) defines a high absolute correlation coefficient as greater than 0.7 and a low as less than 0.3. Consequently, based on appendix 1 there is no significant link between model variables.

Second, the Breusch-Pagan test must be used to detect model heteroskedasticity. The test hypothesis is as follows:

Appendix 2 show that R-Studio's test model yields a P-value of 0.002, indicating statistical significance. This shows that the report's findings may reject the null hypothesis (H0) due to statistical evidence that the model residual is not constant, showing heteroskedasticity. The model fails to achieve BLUE because the report on model 1's variable significance is no longer valid.

The current research will alter the previous model into model 2, the log-log model, to solve the problem. Inflation, GDP growth, FDIPerGDP, NetExportPerGDP, and trade tax have negative values, hence inapplicable logarithmic transformations. Write Model 2 as follows:

Log(Gini)= 3.55 + 0.002Inflation + 0.056Log(Unemployment) - 0.0002 GDPGrowth

(0.123) (0.002) (0.017) (0.001)

0.000FDIPerGDP - 0.000NetExportPerGDP + 0.052Log(IncomeTax) + 0.006TradeTax –

(0.002) (0.001) (0.051) (0.002)

0.114Log(Education)

(0.038)

n = 156, R-squared = 0.1855 , Adjusted R2 = 0.1411

Following appendix 3, The Breusch-Pagan test P-value in model 2 is 0.1018 owing to variable redefinition. The null hypothesis (H0) cannot be rejected with a high P-value, proving that the model holds the fifth assumption of homoskedasticity in the residuals.

|  |  |  |
| --- | --- | --- |
|  | **Dependent variables:** | |
|  | **gini** | **log(gini)** |
|  | **Model (1)** | **Model (2)** |
| inflation | 0.067 | 0.002 |
| unemployment | 0.412\*\*\* |  |
| log(unemployment) |  | 0.056\*\*\* |
| gdpgrowth | 0.005 | 0.0002 |
| fdipergdp | 0.085 | 0.003 |
| netexppergdp | 0.022 | -0.0003 |
| incometax | 0.098\*\* |  |
| log(incometax) |  | 0.052\*\* |
| tradetax | 0.228\*\*\* | 0.006\*\* |
| education | -0.820\*\*\* |  |
| log(education) |  | -0.114\*\*\* |
| Constant | 37.929\*\*\* | 3.546\*\*\* |
| R2 | 0.284 | 0.185 |
| Adjusted R2 | 0.245 | 0.141 |
| Residual Std. Error (df = 147) | 5.396 | 0.154 |
| **Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01** | | |

**Figure *8:*** Regression results for homoscedasticity of Model 1 and Model 2

Unemployment, income tax, trade tax, and education remain statistically significant after being reclassified from model 1 to model 2. Additionally, modified R-square dropped from 0.284 to 0.185. The decline in R-squared or modified R-squared values may be due to inconsistent logarithmic treatment of variables. In theory, heteroskedasticity should not affect R-squared. However, the model's residual standard error has decreased. Consequently, Model 2's adjusted R-squared indicates lesser explanatory capability but decreased residual standard error, indicating better predicted value consistency.

Income inequality in African countries

A dummy variable is added to Model 1 to examine wealth inequalities, notably in African countries, creating Model 3. The study will better understand income inequality in Africa and other regions with this upgrade.

|  |  |
| --- | --- |
|  | **Dependent variable:** |
|  | **gini** |
| inflation | 0.059 |
| unemployment | 0.400\*\*\* |
| gdpgrowth | -0.0003 |
| fdipergdp | 0.085 |
| netexppergdp | 0.021 |
| incometax | 0.092\*\* |
| tradetax | 0.214\*\* |
| education | -0.775\*\*\* |
| african | 0.963 |
| Constant | 37.579\*\*\* |
| R2 | 0.287 |
| Adjusted R2 | 0.243 |
| Residual Std. Error | 5.400 (df = 146) |
| **Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01** | |

***Figure 9:*** Regression results for Model 3

Model 3 shows that African countries have a coefficient 0.96 points higher than the rest of the world. Nevertheless, with a significant standard error of 1.095, the variable's coefficient lacks precision, suggesting that income disparity between African nations and other worldwide countries may be different than reported. Note that this variable is insignificant regarding the Gini index.

Endogeneity

Endogeneity occurs when the explanatory variable correlated with the residual. Lynch and Brown (2011) define simultaneity endogeneity, a subset of Endogeneity, as when changes in the dependent variable affect the explanatory factors. Thus, the dependent and independent variables explain each other's outcomes.

Model 1 may have simultaneity endogeneity due to education. The literature research and statistical analysis show that education is strongly correlated with the Gini index. Education reduces income gaps by improving human capital (Köse and Güven 2007). Nevertheless, simultaneity endogeneity is feasible in this variable. OurWorldInData (2021) found little income inequality in high-GDP countries. Thus, countries with minimal income inequality are often wealthier. Rich countries have more money to invest in human capital and tend to have high education levels, hence the Gini index may explain education level. Consequently, as education and Gini are each other explanatory variable, Simultaneity endogeneity occurred may occurred.

PART 4: CONCLUSION AND POLICY IMPLICATIONS

Conclusion and recommendations

Model 1 aims to investigate the effects of hypothesised independent variables on the Gini coefficient and has revealed 4 variables with significant effects on the Gini coefficient However, tests for Model 1 has found that the data collected has heteroskedasticity, rendering results for Model 1 less reliable due to the data points not following a normal distribution. To correct this, Model 2 proposes applying logarithmic transformation to almost all variables (except inflation, GDP growth, FDI, net export, and trade tax due to negative values present) and re-evaluate t-test afterwards. Model 2 therefore follows a log-log model, effectively showing to be more reliable than Model 1 in terms of results, which eliminates need for further correction of robust standard errors.

Results from Model 1 show that unemployment, trade tax, income tax and education have a significant effect on the Gini coefficient individually that is not dismissible as by chance, which has been confirmed by results from Model 2. Trade tax, notably serving as an indicator of a country’s openness to trade, has shown to have a positive effect on the Gini coefficient. This indicates that the more a country creates barrier for trade (ie increase trade tax), the more disparate income may become between the country’s population, and vice versa. Other variables found to be having effects on the Gini coefficient also have shown results as hypothesised, with unemployment having a positive effect and education having a negative effect. Income tax, hypothesised to have a negative effect as an indicator for the government’s effort of income redistribution, was found to have a positive effect on Gini coefficient, thus suggesting that income inequality may lower as income tax policies are relieved. Policy recommendations for countries looking to reduce income inequality drawn from these findings thus include reducing income tax and trade barriers, along with focusing on welfare aspects such as reducing unemployment and investing more in education. Singapore’s Prime Minister has declared its Gini coefficient reaching a record low due to tax reductions impacting low-income households (Min 2021). Conversely, reports show that welfare issues such as education and employment unavailability in the US during Covid-19 contributed largely to rising income inequality (Haaker and Ibarra 2021; Thorbecke and Mitropoulos 2020).

Analysis limitations

One limitation of this analysis is the lack of data for 2020 due to Covid-19, which renders interpretation of factors affecting income inequality less accurately. Income tax observations from each country collected for this year is a national average, which does not reflect the country’s progressive or regressive tax policies if applicable, thus not fully representing the effect of different tax policies on income inequality. More research could be done regarding categorisation of each country’s tax policies as dummy variables. Less could be done for the lack of data for variables Gini coefficient and Mean years of schooling in 2020 as this is uncontrollable circumstances.

PART 5: INDUSTRIAL TALK

According to Mr. Trung, descriptive analysis is frequently the first stage in comprehending complex econometrics and data analytics datasets. This process involves summing up the primary characteristics of a dataset, such as the mean, median, and mode, and dividing the data into smaller subsets based on specific criteria. In addition, a comprehensive comprehension of the subject matter is required for data analytics to extract meaningful insights from the data. This knowledge, or the academic rationale, is the foundation for the analysis process's assumptions. After conducting foundational analyses such as descriptive analysis, researchers use their expertise in the field to interpret the results and identify patterns or trends. These insights, obtained through analytical techniques and subject-matter expertise, inform decision-making across various industries (Sarker, 2021). A solid understanding of these assumptions' academic rationale is essential for practical data analysis. In addition, RTA analysis and data presentation are also carried out iteratively to implement the central limit theorem. This enables real-time reporting via on-premises code, parallelizing and automating conventional research procedures.

The advent of Artificial Intelligence (AI) has revolutionized the field of data analytics, which is the second significant aspect. AI can automate traditionally labour-intensive and time-consuming complex analytical duties. Machine learning may automate the analysis of vast and intricate datasets, identify trends, and make predictions or classifications. Over the past decade, venture funding in AI has increased 13 times, resulting in a proliferation of "usable" data and readily available tools, significantly increased computing capacity, and open-source techniques (Deveau et al. 2023). This has made data analysis more efficient and effective and enabled real-time data visualization, diminishing dependence on human expertise and making completing data analysis accessible and effective (Davenport and Fitts 2021). According to a McKinsey survey, over 80 percent of IT experts believed AI data insights helped reduce costs, with over a quarter reporting a reduction of more than 20% (Deveau et al. 2023). These statistics demonstrate the transformative effect of AI on statistical analysis and its capacity to increase efficiency and reduce costs.

Students planning a career in this discipline must be aware of these trends. AI's incorporation into data analytics transforms how data is analysed and how businesses operate and make decisions. Those who utilize AI will be in a prime position to dominate this swiftly developing field. Therefore, proficiency in AI and ML and a firm foundation in econometrics could provide a substantial advantage in the job market (Yuen, 2022).

REFERENCES

Acharyya R and Marjit S (2000) ‘Globalisation and Inequality: An Analytical Perspective’, *Economic and Political Weekly*, 35(39):3503–3510. <https://www.jstor.org/stable/4409777>.

Amadeo K (12 November 2021) ‘What Is the Natural Rate of Unemployment?’, *The Balance*, accessed 15 September 2023. <https://www.thebalancemoney.com/natural-rate-of-unemployment-definition-and-trends-3305950.>

Bardhan P (2007) ‘Poverty and Inequality in China and India: Elusive Link with Globalisation’, *Economic and Political Weekly*, 42(38):3849–3852, <https://www.jstor.org/stable/40276420>.

Baur D (2021) *Hyper-inflationary economies at 31 December 2020*, PwC website, accessed 15 September 2023. <https://www.pwc.ch/en/insights/accounting/hyper-inflationary-economies-at-31-december-2020.html.>

Brauw A and Herskowitz S (2020) *New methods aren’t always silver bullets: Challenges with use of the inverse hyperbolic sine transformation for elasticity estimation*, The International Food Policy Research Institute website, accessed 15 September 2023. <https://www.ifpri.org/blog/new-methods-arent-always-silver-bullets-challenges-use-inverse-hyperbolic-sine-transformation>.

Bussin M and Swart-Opperman C (2022) *Measuring income inequality – A holistic approach*, Job Evaluation in the New World of Work, United Kingdom.

Çelik S and Basdas U (2010) ‘How Does Globalization Affect Income Inequality? A Panel Data Analysis’, *International Advances in Economic Research*, 16:258-370, doi:10.1007/s11294-010-9281-0.

Cho HC and Ramirez MD (2016) ‘Foreign Direct Investment and Income Inequality in Southeast Asia: a Panel Unit Root and Panel Cointegration Analysis, 1990–2013’, *Atlantic economic journal,* 44(4):411–424. <https://doi.org/10.1007/s11293-016-9521-7>.

Cingano F (2014) ‘Trends in income inequality and its impact on economic growth’, *OECD Publishing*, 1(1):6-59. <https://doi.org/10.1787/5jxrjncwxv6j-en>.

Cysne RP (2009) ‘On the positive correlation between income inequality and unemployment’, *The Review of Economics and Statistics*, 91(1):218–226. <https://www.jstor.org/stable/25651330>.

Davenport T and Fitts J (2021) *AI can help companies tap new sources of data for analytics*, Harvard Business Review website, accessed 16 September 2023. <https://hbr.org/2021/03/ai-can-help-companies-tap-new-sources-of-data-for-analytics>.

Deveau R, Griffin S (2023) *AI-powered marketing and sales reach new heights with generative AI*, McKinsey website, accessed 16 September 2023. <https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/ai-powered-marketing-and-sales-reach-new-heights-with-generative-ai>.

Engemann K (2019) *The Fed’s Inflation Target: Why 2 Percent?*, The Federal Reserve Bank of St. Louis website, accessed 15 September 2023. <https://www.stlouisfed.org/open-vault/2019/january/fed-inflation-target-2-percent>

Erosa A and Ventura G (2002) ‘On inflation as a regressive consumption tax’, *Journal of Monetary Economics*, 49(4):761–795, doi:https://doi.org/10.1016/s0304-3932(02)00115-0.

Farris FA (2010) ‘The Gini Index and Measures of Inequality’, *The American mathematical monthly,* 117(10):851–864. <https://doi.org/10.4169/000298910x523344>.

Fellman J (1976) ‘The Effect of Transformations of Lorenz Curves’, *Econometrica*, 44(4):823, doi:https://doi.org/10.2307/1913450.

Gao S (2000) *Economic Globalization: Trends, Risks and Risk Prevention*, United Nations website, accessed 18 September 2023. <https://www.un.org/en/development/desa/policy/cdp/cdp_background_papers/bp2000_1.pdf>.

Gini C (1912) *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche. [Fasc. I.]*, Tipogr. di P. Cuppini, Google Books website, accessed 18 September 2023. <https://books.google.com.vn/books/about/Variabilit%C3%A0_e_mutabilit%C3%A0.html?id=fqjaBPMxB9kC&redir_esc=y>.

Goldberg PK and Pavcnik N (2007) ‘Distributional Effects of Globalization in Developing Countries’, *Journal of Economic Literature*, 45(1):39–82, <https://www.jstor.org/stable/27646747>.

Gozgor G and Jha P (2015) ‘Globalization, Inequality, and Redistribution: Theory and Evidence’, *SSRN Electronic Journal*, 5522(1), doi:https://doi.org/10.2139/ssrn.2675996.

Greenwood J (2022) Detecting Influential Points in Regression with DFBETA(S), accessed 15 September 2023, University of Virginia Library database.

Haaker JR and Ibarra GL (2021) *How COVID-19 is exacerbating inequality of opportunities in Latin America*, World Bank blogs website, accessed 19 September 2023. <https://blogs.worldbank.org/latinamerica/how-covid-19-exacerbating-inequality-opportunities-latin-america>.

Jäntti M and Jenkins SP (2010) ‘The impact of macroeconomic conditions on income inequality.’, *Journal of economic inequality,* 8(2):221–240, <https://doi.org/10.1007/s10888-009-9113-8>.

Köse S and Güven A (2007) ‘Government education expenditures and income inequality: evidence from provinces of Turkey’, *SEER: Journal for Labour and Social Affairs in Eastern Europe*, 10(1):79–101. <https://www.jstor.org/stable/43293202?seq=4>

Lloyds Bank (2023) *Tax rates in Belarus*, The International Trade Portal website, accessed 15 September 2023. <https://www.lloydsbanktrade.com/en/market-potential/belarus/taxes>

Lynch SM and Brown JS (2011) ‘Chapter 8 - Stratification and Inequality over the Life Course’, *ScienceDirect*, 8(1):105-117. <https://www.sciencedirect.com/science/article/pii/B9780123808806000083>.

Majumdar S and Partridge MD (2009) ‘Impact of Economic Growth on Income Inequality: A Regional Perspective’, *IDEAS Working Paper Series,* DOI: 10.22004/ag.econ.49270*.*

Milanovic B (2022) *The three eras of global inequality, 1820-2020 with the focus on the past thirty years*, Graduate Center City University of New York, OSF website, accessed 18 September 2023. <https://osf.io/preprints/socarxiv/yg2h9/download>.

Monnin P (2014) *Inflation and Income Inequality in Developed Economies*, The Council on Economic Policies (CEP) website, accessed 15 September 2023. <https://www.cepweb.org/wp-content/uploads/2014/05/CEP_WP_Inflation_and_Income_Inequality.pdf>

Nguyen CP and Su TD (2022) ‘Export Dynamics and Income Inequality: New Evidence on Export Quality’, *Social indicators research*, 163(3):1063–1113, <https://doi.org/10.1007/s11205-022-02935-4>.

OECD (2020) *Tax policies vary widely from country to country*, Organisation for Economic Co-operation and Development website, accessed 15 September 2023. <https://www.oecd.org/general/taxpoliciesvarywidelyfromcountrytocountryoecdstudyshows.htm>

OurWorldInData (2021) *Income inequality vs. GDP per capita*, *Our World in Data,* OurWorldInData website, accessed 18 September 2023. <https://ourworldindata.org/grapher/gini-coefficient-vs-gdp-per-capita-pip?xScale=linear.>

Owens J and Zhan J (2018) ‘Introduction to the Special Issue Trade, Investment and Taxation: Policy Linkages’, *Transnational Corporations Journal*, 25(2). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3617614>.

Polacko M (2021) ‘Causes and Consequences of Income Inequality – An Overview’, *Statistics, Politics and Policy*, 12(2):341–357, doi:https://doi.org/10.1515/spp-2021-0017.

Rubin A and Segal D (2015) ‘The effects of economic growth on income inequality in the US’, *Journal of Macroeconomics*, 45(1):258–273, doi:https://doi.org/10.1016/j.jmacro.2015.05.007.

Sarker I (2021) ‘Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective’, *Springer Link*, 2(5):377-378. DOI: 10.1007/s42979-021-00765-8.

Siegel R, Dam A and Werner E (28 January 2021) ‘2020 was the worst year for economic growth since World War II’, *The Washington Post*, accessed 15 September 2023. [https://www.washingtonpost.com/business/2021/01/28/gdp-2020-economy-recession/.](https://www.washingtonpost.com/business/2021/01/28/gdp-2020-economy-recession/)

Stolper WF and Samuelson PA (1941) ‘Protection and Real Wages’, *The Review of Economic Studies*, 9(1):58–73, doi:https://doi.org/10.2307/2967638.

Tamboli N (2023) *Effective Strategies to Handle Missing Values in Data Analysis*, Analytics Vidhya website, accessed 17 September 2023. [https://www.analyticsvidhya.com/blog/2021/10/handling-missing-value/.](https://www.analyticsvidhya.com/blog/2021/10/handling-missing-value/)

Tanzi V and Zee H (2004) ‘Tax Policy for Developing Countries. Place of publication not identified: International Monetary Fund.

Ulf Jakobsson (1976) ‘On the measurement of the degree of progression’, *Journal of Public Economics*, 5(1-2):161–168, doi:https://doi.org/10.1016/0047-2727(76)90066-9.

Thorbecke C and Mitropoulos A (2020) ‘‘Extreme Inequality Was the Preexisting condition’: How COVID-19 Widened America’s Wealth Gap’, *ABC News* accessed 19 September 2023. <https://abcnews.go.com/Business/extreme-inequality-preexisting-condition-covid-19-widened-americas/story?id=71401975>.

UNCTAD (United Nations Conference on Trade and Development) (24 January 2021) ‘Global foreign direct investment fell by 42% in 2020, outlook remains weak’, *United Nations Conference on Trade and Development*, accessed 15 September 2023. <https://unctad.org/news/global-foreign-direct-investment-fell-42-2020-outlook-remains-weak>

United Nations (2023) *SDG Report 2021 - Goal 8 - Decent Work and Economic Growth*, The Sustainable Development Goal indicators website, accessed 15 September 2023. <https://unstats.un.org/sdgs/report/2021/goal-08/>

Villanthenkodath M, Pal S and Mahalik M (2023) ‘Income Inequality in Globalization Context: Evidence from Global Data’, *Journal of the Knowledge Economy*, doi:10.1007/s13132-023-01342-5.

Wei S-J and Wu Y (2001) *NBER Working Paper Series Globalization and Inequality: Evidence from Within China*, National Bureau of Economic Research website, accessed 18 September 2023. <https://www.nber.org/system/files/working_papers/w8611/w8611.pdf>.

World Bank (2023) *DataBank - World Development Indicators [Excel]*, The World Data website, accessed 15 September 2023. <https://databank.worldbank.org/source/world-development-indicators>

Yang J and Qiu M (2016) ‘The impact of education on income inequality and intergenerational mobility’, *China Economic Review*, 37(1):110–125, doi:https://doi.org/10.1016/j.chieco.2015.12.009.

Yuen M (26 September 2022) ‘How big data warehouses and AI in data analysis are transforming business intelligence in 2022’, *Insider Intelligence*, accessed 16 September 2023. <https://www.insiderintelligence.com/insights/ai-data-analysis/>.

APPENDIX

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Inflation | Unemployment | GDPGrowth | FDI/GDP | Net Export / GDP | Income Tax | Trade Tax | Education |
| Inflation | 1 | 0.01 | 0.16 | 0.08 | -0.01 | -0.07 | 0.16 | -0.18 |
| Unemployment | 0.01 | 1 | -0.1 | -0.04 | -0.2 | -0.06 | -0.02 | -0.01 |
| GDPGrowth | 0.16 | -0.1 | 1 | -0.34 | 0.02 | 0.05 | -0.07 | -0.01 |
| FDI | 0.08 | -0.04 | -0.34 | 1 | 0.32 | -0.1 | -0.04 | 0.04 |
| Net Export | -0.01 | -0.2 | 0.02 | 0.32 | 1 | 0.05 | -0.17 | 0.25 |
| Income Tax | -0.07 | -0.06 | 0.05 | -0.1 | 0.05 | 1 | -0.13 | 0.03 |
| Trade Tax | 0.16 | -0.02 | -0.07 | -0.04 | -0.17 | -0.13 | 1 | -0.12 |
| Education | -0.18 | -0.01 | -0.01 | 0.04 | 0.25 | 0.03 | -0.12 | 1 |

***Appendix 1***: Correlation coefficient matrix between 8 variables of model 1

|  |  |  |  |
| --- | --- | --- | --- |
| Studentized Breusch-Pagan Test | | | |
| Model Name | Model 1 | BP | 23.448 |
| Degree of Freedom | 8 | p-value (LM) | 0.002834 |

***Appendix 2***: Breusch-Pagan test result for model 1

|  |  |  |  |
| --- | --- | --- | --- |
| Studentized Breusch-Pagan Test | | | |
| Model Name | Model 2 | BP | 13.304 |
| Degree of Freedom | 8 | p-value (LM) | 0.1018 |

***Appendix 3***: Breusch-Pagan test result for model 2