Quantitative Analysis of Global Economic Inequality after the 2008 Financial Crisis: A Mathematical Modeling Approach Using Multiple Linear Regression

Jiayuan Yang

Weihai No.1 High School, Weihai, China
Reginayangjy@outlook.com

Abstract. This study explores how globalization affects economic inequality across social classes through transnational economic interactions by a quantitative analysis and case studies. A multivariate linear regression model is constructed based on World Bank data after the 2008 financial crisis. The results show that export price levels(log(pl_x)) and capital prices(log(pl_k)) significantly increase income inequality, while consumer prices(log(pl_c)) and share of labor income(labsh) have a suppressive effect. The study found that domestic policies(such as the high welfare system in Northern Europe) can partially offset the negative impact of globalization, but the model has limited explanatory power for Latin American and African countries. The study advocates addressing the challenge of inequality under globalization through tax reform, strengthening social security and industrial diversification policies.

Keywords: Globalization, Economic inequality, GINI index, Multiple Linear Regression, Policy Intervention

1. Introduction

Nowadays, globalization has become a hot topic and is an important force to drive the development of the world economy, which drastically promotes economic relations and resource flows between countries. In the process of globalization, there are multiple economic interactions between countries through international trade, capital flow and labour migration, which brought number of positive effects such as scientific and technological progress, labour enrichment and international market expansion [1]. Meanwhile, the impact of globalization on income distribution is highly controversial. In some countries, market expansion has brought more benefits to capital owners, while the wages of workers and farmers have relatively slowly, resulting in widening income disparities and fewer middle-income groups [2]. Besides, the benefits of globalization, such as the improvement of science and technology, has possibilities exacerbate the economic development imbalance between different countries, and caused some developing countries face problems, like labor unemployment and limited market opportunities for small and micro enterprises, thereby further exacerbating domestic inequalities.

With more details, from the perspective of developed countries and developing countries relatively, developed countries are inclined to set up manufacturing factories in lower-cost developing countries [3], globalization has led to increased competition in developing countries. For example, Chile's largest steelmaker, Compania de Acero del Pacifico, which said in 2023 that it would shut down operations at its Huachipato steel mill because it could no longer compete with Chinese steel imports, which are 40% cheaper than Chilean steel. This threatens more than 22,000 jobs. Although the Chilean government has proposed a 15% tariff on Chinese steel imports, companies believe this will not be enough to offset the low price shock [4]. The two phenomena revealed unemployment or lower incomes for national workers.

Consequently, this study aims to explore how globalization affects economic inequality of all social strata through transnational economic interactions, and quantitatively analyze the impact on GDP, employment, and national accounts of countries around the world in the year after the 2008 economic crisis through multiple linear regression in mathematical modeling, and discuss possible countermeasures to achieve a more equitable distribution of wealth.

Firstly, the study will analysis how income distribution is distributed between developed and developing countries from a qualitative and quantitative perspective, and compare the differences in income distribution between the two kinds of countries. Secondly, the study will reveal the method of income distribution between developed and developing countries exacerbated the disparity between social classes. Finally, combined with the results of the study, some theoretical foundations for policymaking are provided.

2. Literature review

Globalization covers a wide range of time. Thomas L Friedman divides the history of globalization into three period. The Globalization 3.0 involves individuals and has been going on since 2000 [5]. Globalization has always been widely discussed in society. Since the beginning of the 21st century, the in-depth development of economic globalization and multinational corporations has not only brought significant impetus to world trade, but also contributed to economic inequality within the societies of various countries. In response, experts and scholars from all walks of life have actively pointed out profound problems and explored the causes and solutions to this phenomenon.

2.1. The impact of globalization on economic inequality

Although globalization is considered to be the driving force of economic development, its substantial impact on the economics of different kinds of countries is not single. Nowadays, many scholars believed that globalization has exacerbated the gap between the rich and the poor, especially between developing and developed countries.

On the one hand, Gereffi believes that globalization has promoted the global utilization of resources through channels such as international trade and capital flows, but this has enabled developed countries to benefit from capital and high-tech, while developing countries can only obtain lower wages through labour processing, thereby exacerbating the income inequality [6]. Wade also questioned whether globalization has really reduced poverty and inequality? In his opinion, even though globalization has promoted economic growth, due to the unreasonable distribution of benefits, many developing countries and vulnerable groups have not benefited and even faced greater economic risks [7].

On the other hand, inequality in developed countries is also becoming increasingly apparent. Low-skilled workers and the middle class in developed countries have been hit by globalization,

leading to income stagnation or decline. According to analysis from Dorn and Hanson of the impact of Chinese import completion on the US labour market, it was found that globalization has caused a reduction in US manufacturing employment and exacerbated income inequality between regions and skills [8].

Furthermore, "Globalization Trilemma" from Rodrik figured it is difficult to have globalization, national sovereignty and democratic politics at the same time. Globalization promotes the transformation of industrial structure, which has brought profound negative impacts to some social classes, especially the wealth and social security level of the middle and low-income groups have not been improved accordingly, leading to the aggravation of social inequality [9].

2.2. Quantitative analysis: the impact of globalization on GDP, employment and national accounts

In order to make the analysis more reasonable, many scholars have adopted quantitative methods, especially statistical tools such as regression analysis. Harrison found through regression analysis of globalization process in the 1990s and 2000s that globalization has promoted GDP growth in some developed countries, but this growth has not significantly improved the overall welfare of society. On the contrary, employment opportunities for some low-skilled workers have reduced and income inequality has raised [10].

In the multivariate linear regression model, variables affecting globalization such as international trade are used as independent variables, and the impact on various aspects of economy is quantified by analyzing the relationship between these variables and GDP, employment rate, GINI index, national accounts, etc. For example, Dollar and Kraay used panel data regression analysis to analyze globalization indicators(such as trade openness) and examine the relationship between GDP growth, poverty rate and GINI index. They found that there is a positive correlation between economic growth and poverty reduction though cross-national data, and proving that globalization indirectly reduces poverty, but its impact on inequality is more complicated [11].

2.3. Countermeasures to economic inequality after globalization

After discovering that globalization has led to economic inequality, scholars have proposed corresponding policies. Piketty points out that globalization and high-tech have intensified the distribution injustice between capital and labour. In order to achieve a more fair distribution of wealth, he proposed a series of tax reform measures, including global wealth taxation and cross-border tax cooperation, to reduce the concentration of capital gains and improve the efficiency of wealth redistribution [12].

2.4. Summary and relevance

In summary, while globalization promotes economic growth through transnational economic interactions, it also leads to economic inequality. After the 2008 economic crisis, globalization had different influence on GDP and employment in the short term, but in the long run, the challenges faced by low-income groups and developing countries remain severe. Quantitative analysis can more clearly see the muti-dimensional effects, so that policy makers have more possibilities to take measures that are more suitable for national conditions and promote more equal wealth distribution.

3. Exploratory data analysis

3.1. Data description and processing

The dataset involved in this paper comes from the website of the World Bank Group. The World Bank Group (composed of five parts: the International Bank for Reconstruction and Development, the International Development Association, Investment Guarantee Agency, and the International Center for Settlement of Investment Disputes) is an international organization established at the end of World War II. Its initial mission was to help rebuild countries destroyed in World War II. Nowadays, this organization provides low-interest loans, interest-free credits and grants to developing countries to help them overcome poverty. Each institution plays a unique role in the mission of alleviating poverty and improving living standards. The data contains 12377 rows and 52 features, covering various information such as real GDP, employment and population levels, current price GDP, capital and TFP, national accounts, exchange rate and GDP price level etc. The target variable is GINI index, which used to comprehensively examine the differences in the internal income distribution of residents in developed and developing countries and which mainly affect it.

Since this paper principally quantitively analyzes the impact of the 2008 economic crisis on factors such as GDP, employment, and national accounts of countries, various data from 180 countries in 2009 were selected during data preprocessing. There are two methods to deal with missing values: the first is to delete this factor if the data for this factor is mission in most countries, and it does not participate in the overall evaluation; the second is to take the average of data for surrounding years or directly delete this country if a factor is missing in some countries, and it does not participate in the overall evaluation. After completing the above operations, the data is reduced to 77 rows and 43 features. Then, added the GINI index data of the remaining countries in 2009. If there is missing data, add the data of 2008 to supplement it. Finally, 0 and 1 are used to distinguish developed and developing countries(0 represents developed countries, 1 represents developing countries). It can be seen that the data includes 32 developed countries and 45 developing countries.

A more detailed interpretation of data in the Table 1 from the perspective of range and standard deviation reveals that the two data related to economic aggregate (GDP(rgdpe, rgdpo, rgdpna), capital stock (cn, cda), etc.) are both large, indicating that there are obvious differences in economic scale of countries. At the same time, the population(pop, emp) and employment data are also large, indicating the diversity of population sizes in various countries.

Although the two variables related to price level(pl_con, pl_c, pl_x) are relatively small, they still show significant differences in price levels and purchasing power among countries. The standard deviation of the GINI index is moderate, indicating that there are certain differences in the degree of income inequality worldwide, but it will not be as extreme as GDP.

Table 1. Data dictionary

Variable	Descrpition	Unit	Range	Average	Standard deviation
country	Name of country	None	None	None	None
rgdpe	Expenditure-side real GDP at chained PPPs	in mil. 2011US$	3687 - 15029481	861801	2196762
rgdpo	Output-side real GDP at chained PPPs	in mil. 2011US$	3700 - 14865562	847600	2168571
pop	Population	in millions	0.3161 - 1352.0681	50.2162	158.5244

Table 1. (continued)

emp	Number of persons engaged	in millions	0.1632 - 772.4278	23.8574	89.36748
hc	Human capital index, based on years of schooling and returns to education; see Human capital in PWT9.	None	1.156 - 3.687	2.789	0.6283391
ccon	Real consumption of households and government, at current PPPs	in mil. 2011US$	3797 - 12464253	634668	1616044
cda	Real domestic absorption, (real consumption plus investment), at current PPPs	in mil. 2011US$	4165 - 15369474	851510	2198358
cgdpe	Expenditure-side real GDP at current PPPs	in mil. 2011US$	3657 - 14959026	858564	2194053
cgdpo	Output-side real GDP at current PPPs	in mil. 2011US$	3688 - 14922572	853913	2182575
cn	Capital stock at current PPPs	in mil. 2011US$	12524 - 51777660	3398614	8106787
ck	Capital services levels at current PPPs	USA=1	0.0002094 - 1	0.059526	0.1480003
ctfp	TFP level at current PPPs	USA=1	0.2513 - 1.1959	0.672	0.2151609
cwtfp	Welfare-relevant TFP levels at current PPPs	USA=1	0.2204 - 1.0404	0.6623	0.1973763
rgdpna	Real GDP at constant 2011 national prices	in mil. 2011US$	3897 - 14922642	888643	2218163
rconna	Real consumption at constant 2011 national prices	in mil. 2011US$	3824 - 12682013	658891	1650422
rdana	Real domestic absorption at constant 2011 national prices	in mil. 2011US$	4302 - 15505308	880334	2229080
rnna	Capital stock at constant 2011 national prices	in mil. 2011US$	15086 - 50871128	3507006	8117300
rkna	Capital services at constant 2011 national prices	2011=1	0.6401 - 1.1714	0.935	0.07009802
rtfpna	TFP at constant national prices	2011=1	0.8652 - 1.2108	0.9806	0.04751854
rwtfpna	Welfare-relevant TFP at constant national prices	2011=1	0.7394 - 1.2065	0.9726	0.06810633
labsh	Share of labour compensation in GDP at current national prices	None	0.1643 - 0.6789	0.5161	0.1120583
irr	Real internal rate of return	%	0.01 - 0.43007	0.10159	0.07496294
delta	Average depreciation rate of the capital stock	%	0.0243 - 0.08432	0.0432	0.0106293
xr	Exchange rate, national currency/USD (market+estimated)	%	0.642 - 9864.303	280.436	1268.126
pl_con	Price level of CCON (PPP/XR)	price level of USA GDPo in 2011=1	0.2774 - 1.5307	0.7006	0.3488037
pl_da	Price level of CDA (PPP/XR)	price level of USA GDPo in 2011=1	0.3023 - 1.4413	0.6923	0.313794

Table 1. (continued)

pl_gdpo	Price level of CGDPo (PPP/XR)	price level of USA GDPo in 2011=1	0.2586 - 1.8868	0.7035	0.3355154
csh_c	Share of household consumption at current PPPs	%	0.34 - 0.9227	0.6218	0.1243015
csh_i	Share of gross capital formation at current PPPs	%	0.06297 - 0.47464	0.217	0.07366998
csh_g	Share of government consumption at current PPPs	%	0.0498 - 0.2985	0.186	0.05685012
csh_x	Share of merchandise exports at current PPPs	%	0.01576 - 1.16884	0.29268	0.2255672
csh_m	Share of merchandise imports at current PPPs	%	(-1.19551) - 0.07204	-0.34221	0.2185283
csh_r	Share of residual trade and GDP statistical discrepancy at current PPPs	%	(-0.3492819) - 0.4429156	0.0247448	0.08995006
pl_c	Price level of household consumption	price level of USA GDPo in 2011=1	0.2773 - 1.4485	0.6985	0.3223899
pl_i	Price level of capital formation	price level of USA GDPo in 2011=1	0.3464 - 1.2523	0.7057	0.226736
pl_g	Price level of government consumption	price level of USA GDPo in 2011=1	0.1745 - 1.8473	0.724	0.4375541
pl_x	Price level of exports	price level of USA GDPo in 2011=1	0.5097 - 0.74	0.6861	0.05175123
pl_m	Price level of imports	price level of USA GDPo in 2011=1	0.558 - 0.8026	0.6537	0.0436251
pl_n	Price level of the capital stock	price level of USA in 2011=1	0.2452 - 1.2999	0.638	0.2884795
pl_k	Price level of the capital services	price level of USA=1	0.2927 - 2.1354	0.8817	0.3786422
GINI index	The extent to which the distribution of income or consumption among individuals or households within an economy deviates from a perfectly equal distribution.	None	24.8 - 63.4	38.21	9.566965
country_type	Distinguish between developed and developing countries	None	0 or 1	None	None

3.2. Data visualization

Since the data set is large and involves a number of variables, histogram, box plot and bar chart are first used to identify patterns, trends and anomalies in income distribution.

3.2.1. Histogram analysis

The concentrated distribution of data from countries is analyzed by the skewness of the histogram distribution, the kernel density curve (KDE) and whether it is a bimodal distribution in the

histogram. For example, according to Figure 1, it shows a positive skewed distribution, which means that the income inequality in most countries is low, while the GINI index of a few countries is particularly high, and the extreme gap between the rich and the poor.
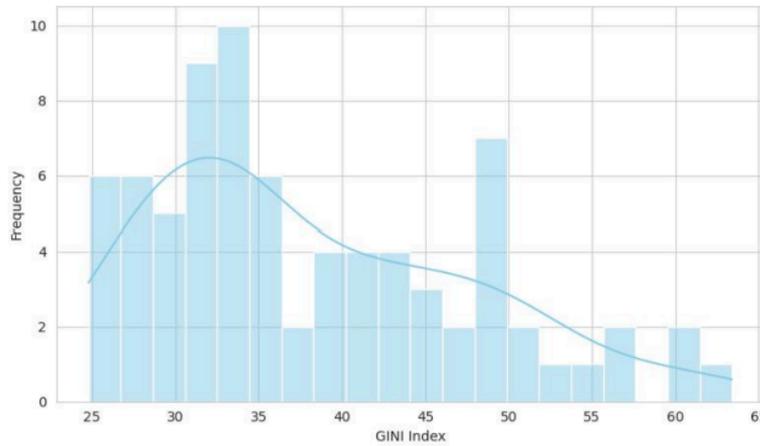

Figure 1. Histogram of GINI index

3.2.2. Box plot analysis

The median, interquartile range (IQR) and outliers in the box plot, are used to identify outliers and data ranges. The data used in the Figure 2 has no outliers, which can demonstrate which countries have income inequality far above or far below the global average. For example, countries with extremely high GINI indexes may lack effective wealth redistribution policies, while countries with lower indexes usually have sound social welfare systems.
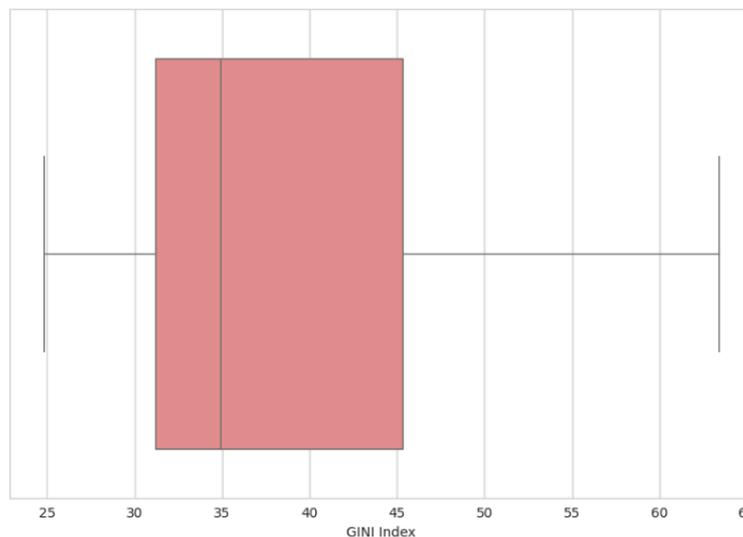
Figure 2. Box plot of GINI index

3.2.3. Bar chart analysis

According to Figure 3, the bar chart sorted by GINI index from small to large clearly shows the ranking of income inequality in various countries. The lower the GINI index is, the more evenly distributed the annual income is in countries, and they usually have fairer tax policies and social

security systems. On the other hand, countries with higher GINI index have more unequal annual income distribution and often face the problem of increasing income disparity. Through this chart, income distribution trends in specific regions can also be identified. For example, Nordic countries like Norway(26.3) and Czech Republic(26.2) are generally at a low level, while some countries in Latin America or Africa like South Africa(63.4) have a higher GINI index.
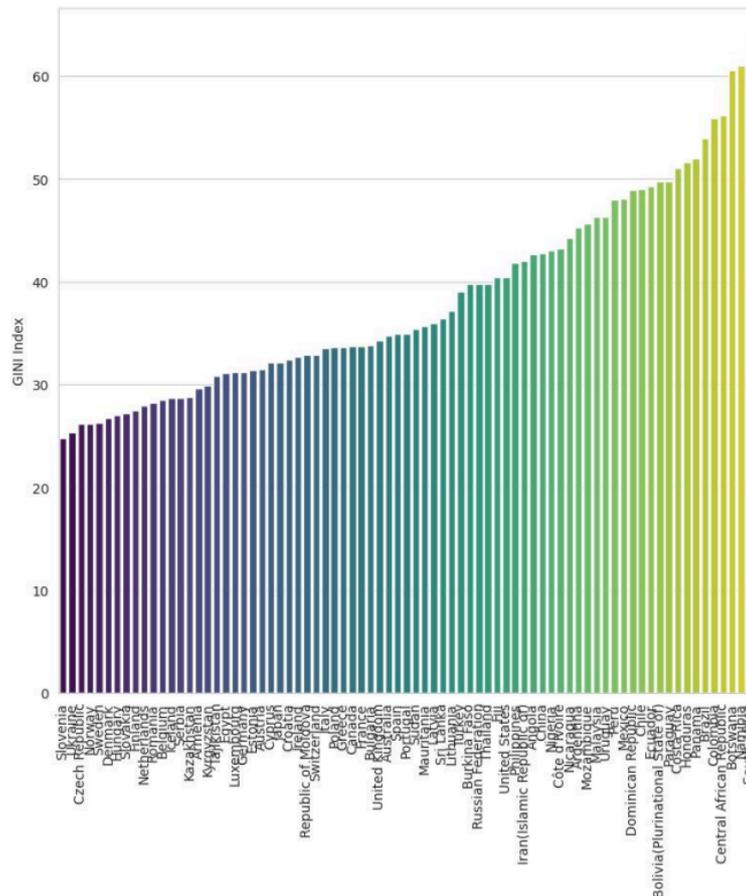

Figure 3. Bar chart of GINI index

4. Model building

4.1. Model fitting

4.1.1. Multicollinearity problem checking

As the study uses a multiple linear regression model, it is necessary to check whether there is a multicollinearity problem. The variance inflation factors(VIF) is a very common evaluation indicator. In statistics, VIF measures the severity of multicollinearity in a regression model, quantifying how much the variance of a regression coefficient is inflated due to this correlation. VIF is calculated using the following formula:

$$VIF = \frac{1}{1 - R^2} \tag{1}$$

where R^2 is the R-squared value obtained from regression the predictor variable on the other predictors.

When VIF = 1, it means there is no multicollinearity. When the VIF value exceeds 1, it means there is multicollinearity, and the higher the VIF is, the more serious the multicollinearity will be. If VIF is greater than 5, it is usually regarded as a sigh of multicollinearity. In this study, variables with VIF greater than 7(hc, rwtfpna, labsh, pl_m)are deleted and not included in the subsequent stepwise regression analysis.

4.1.2. Stepwise regression variable screening

The core of stepwise regression is to determine whether a variable should be added or removed based on statistical significance tests. The covariates were checked from three aspects: "forward selection" , "backward elimination" and "both elimination" through stepwise regression. The principle of "forward selection" is to start with no variables, add one variable at a time, and then choose the variable that optimally improves the model until no variable can significantly improve the model. Meanwhile, the principle of "backward elimination" is to start with a complete model containing all variables, calculate the p-value of each variable, and remove the variable if a variable's p-value exceeds a set threshold, removing one variable at each step until all remaining variables are statistically significant. Besides, "both elimination" combines the forward selection and backward elimination methods, adding variables according to the forward selection method, and performing a backward elimination check after each new variables is added to ensure that all variables in the model are still significant until no variables can be added or removed.

Therefore, this study compares the Akaike Information Criterion(AIC) values of the three methods. The AIC of "forward selection" is 316.57, that of "backward elimination" is 293.6 and that of "both elimination" is 293.48. Based on the above results, using the "both elimination" to complete and the regression model in algebraic form is as follows:

$$Y_i = \beta_0 + \beta^T X + \epsilon_i \tag{2}$$

where Y_i is the GINI index for i^{th} country in the world.

$$\beta^T = \left(\beta_{rgdpo}, \beta_{emp}, \beta_{cda}, \beta_{cgdpe}, \beta_{cgdpo}, \beta_{cn}, \beta_{ck}, \beta_{rgdpna}, \beta_{rdana}, \beta_{rnna}, \beta_{pl\_da}, \beta_{csh\_c}, \right.$$

$$\left. \beta_{csh\_i}, \beta_{csh\_x}, \beta_{csh\_m}, \beta_{pl\_c}, \beta_{pl\_x}, \beta_{pl\_n}, \beta_{pl\_k}\right)$$

4.2. Model checking

4.2.1. Regression diagnostics

As the study is using a linear regression model to estimate the relationship between GINI index and the other covariates, looking for any evidence that Gauss-Markov assumption do not hold is imperative.

In the analysis of histogram in 3.2.1, it is found that the distribution was positively skewed, so it does not satisfy the normal distribution assumed by regression model. Meanwhile, for the assumption of linearity, check the scatterplots of the standard residuals versus the numeric covariates, as shown in Figure 4. However, this results does not satisfy the linearity problem of the regression model. Consequently, the study decided to formulate a new linear model via logarithmic transformation on these variables to make them meet the basic assumption. The Residuals versus Fitted plot after transformation is shown in Figure 5. The amplitude of the red line fluctuating in the horizontal direction has become smaller, which indicates that the problem of non-linearity described previously has been corrected to a certain extent.
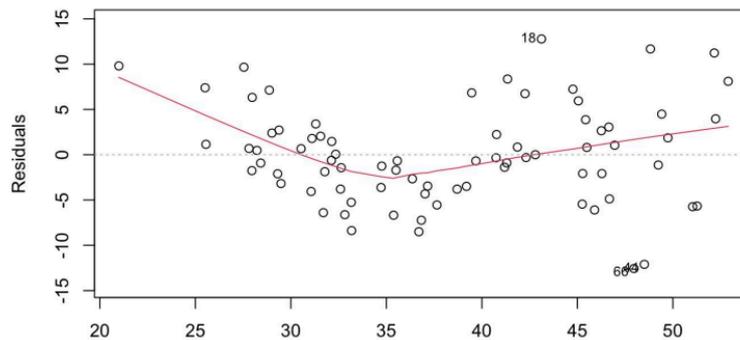

Figure 4. Residuals vs. fitted plot before logarithm transformation correction
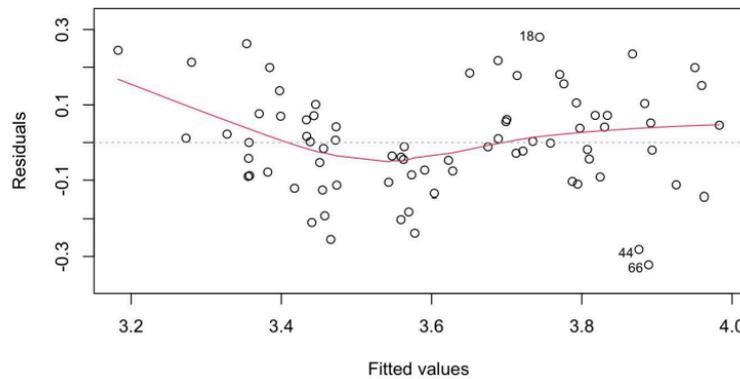

Figure 5. Residuals vs. fitted plot after logarithm transformation correction

After transformation for explanatory variables, for the assumption of normally distributed errors, Figure 6 shows that with the exception of observations the #66 and #72, most part of observations fall close to the straight line and more that 95% of the data fall within the confidence interval of studentized residual $[-2,2]$, marked by the blue area.Thus, there is no evidence against the normality assumption. In addition, according to the error distribution shown in Figure 7, not only the data basically obeys the normal distribution, but also the Normal Curve conforms to the theoretical shape of the standard normal distribution. Hence, the assumption of independence of errors is reasonable since each data point is a unique country.
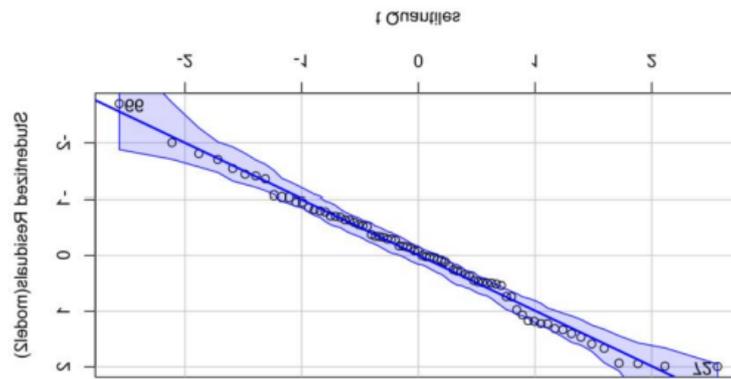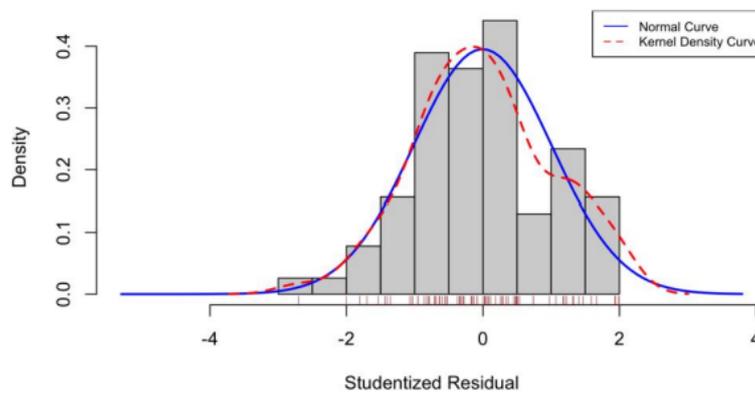
Figure 6. Q-Q plot

Figure 7. Distribution of errors

For the assumption of independence of homoscedasticity, in the Figure 8, the overall performance is relatively stable. and plenty of the points are randomly distributed above and below the red line. Furthermore, the p-value(0.82717) of this group of data is larger than 0.05, so there is no heteroscedasticity and it is homoscedasticity, so the assumption is passed.
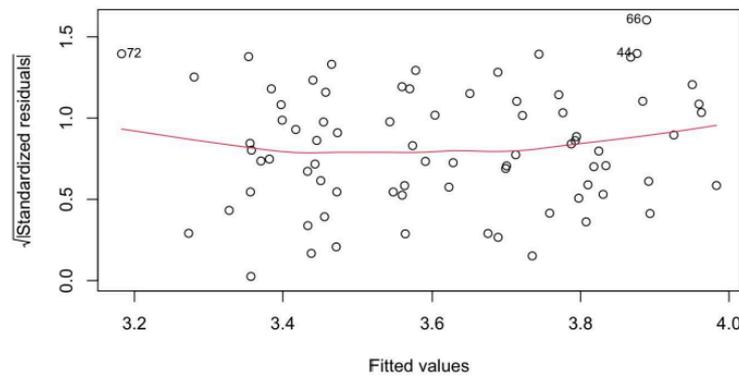

Figure 8. Scale-location plot

4.2.2. Unusual observation

The Cook's distance plot (Figure 9) and Index Plot of Hat Values(Figure.10) show that #76 has the largest Cook's distance, which shows that it has a great impact on fitted value and then means that if

this data is removed, the regression coefficient of model will change greatly. Also, its highest leverage value shows that the economic variables are very different from other countries. Specifically for Uruguay, this country has a special economic structure, a small open economy, mainly based on agriculture, animal husbandry and service industry. However, this export-oriented economy means the economy is vulnerable to global market fluctuations. The high GINI index in 2008(46.3) probably resulted from, against the backdrop of global economic fluctuations, the economy relied on agriculture and the decline in commodity prices has damaged exports, affecting the livelihoods of low-income groups. Moreover, a system of the high welfare and high tax may inhibit corporate investment and economic vitality. Once the rich evade taxes, it would most likely lead to wealth accumulation in the hands of a few people, thereby increasing income inequality.

#66 has the largest negative variance and is overestimated by the regression model, indicating that its economic data has strong volatility or extreme characteristics, as show in the Q-Q plot(Figure 6), #66 is the Russian Federation, which has an unstable economic growth mode. It is an oligarchic economy formed after the collapse of the Soviet Union, with a large amount of wealth concentrated in the hands of a few people. Furthermore, this economy is highly dependent on energy. As one of the world's largest energy exporters, oil prices guide economic trends, in the mid-2000s, rising energy prices drove economic growth, but the benefits numerously went to large state-owned enterprises and a few wealthy people. The wages of ordinary workers grew slowly, while the income of executives in the energy and financial industries grew rapidly, widening the income gap. The financial crisis in 2008 had an impact on the Russian economy. The devaluation of the ruble and the stock market crash led to a decline in the living standard of ordinary people. The rich could easily cope with the crisis through capital operations, further widening the income gap between the rich and the poor.

#16 and #41 both have high leverage values(Figure 10), medium Cook's distances(Figure.9) and normal performance in Q-Q plot(Figure 6), which reflects that they have an impact on the model but do not completely dominate the results. Moreover, since the economic conditions of the two countries are different, the reasons leading to the above situations are not the same.

#16 is Chile. whose GINI index has always been at a high level, reaching 49 in 2008. This probably because the country's economic growth is highly dependent on copper exports(which account for a high proportion of GDP(csh_x = 0.294)). The international copper price rose sharply from 2003 to 2007. Although the economy grew rapidly during the period, wealth was mainly concentrated in a few wealthy classes, and the income of the middle class and low-income groups grew slowly. However, after the outbreak of the financial crisis in 2008, the copper price plummeted, which impacted its fiscal situation and led to debt growth. On the one hand, for wealthy class, their leveraged investments are more concentrated in the capital market and overseas assets. After the outbreak of the financial crisis in 2008, investment losses exacerbated debt pressure and increased the overall leverage ratio. On the other hand, for low-income groups, their reliance on credit consumption is even more serious. They need to borrow money to buy basic needs such as housing, education and medical care. In order to meet these needs, financial institutions provide them with higher-risk loans, resulting in an increase in household leverage.

#41 is Italy. Firstly, its labsh(0.53) is relatively low, indicating that the proportion of labor income to GDP is low, which is specifically reflected in slow wage growth, high youth unemployment. Many families rely on credit to survive. After the 2008 financial crisis, many borrowers were unable to repay their debts, which further increased the household debt leverage ratio. Secondly, the banking system is fragile and the bad debt problem is serious, which affects credit expansion and investment. The government is forced to invervene, increasing the public debt burden. Lastly, Italy

is one of the countries with heaviest debt burden in the EU. The imperfect tax policy has caused the government's fiscal deficit to continue to expand. After the 2008 crisis, government debt has further soared.
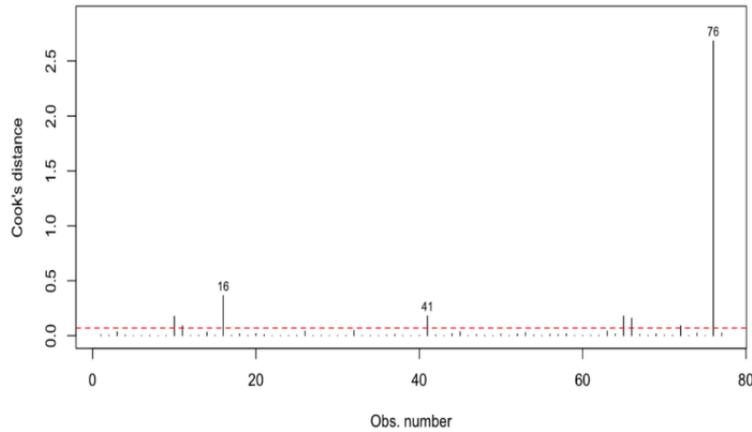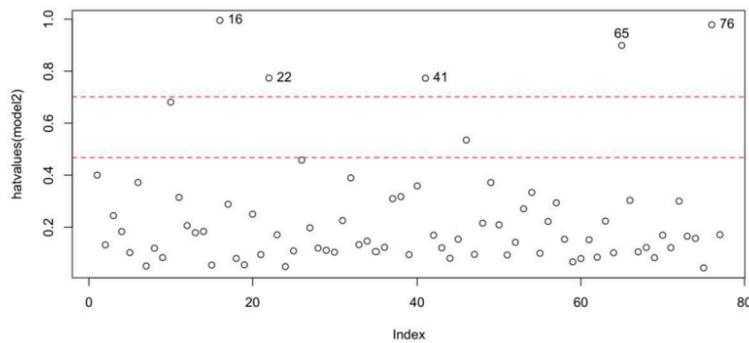

Figure 9. Cook's distance plot

Figure 10. Index plot of hat values

4.3. Model results

The summary of final model is shown in Table 2. Based on the final multiple regression model, income inequality is positively correlated to log(pl_x), log(pl_k), interaction term and negatively correlated to log(pl_n), log(pl_c). Particularly, log(cn), log(rmna), log(pl_n) and log(rdana) were highly significant. In, addition, for categorical variables, consumption and labor productivity play an important role in determining the impact of income inequality within a country. However, considering only these two factors cannot fully explain why the variables show abnormal conditions. In other words, factors such as the stability of a country's pillar industries, whether the banking system is stables, whether households will borrow money to maintain normal demand, whether the country's fiscal policy is well-established and so on will influence income inequality and amplify the phenomenon of the inequality during an economic crisis.

Table 2. Estimates for the linear model with covariates

Coefficients	Estimate	Std. Error	t-value	p-value	
Intercept	2.744e+00	4.195e-01	6.541	1.61e-08	***
log(rgdpo)	6.516e-06	4.031e-06	1.617	0.111268	
log(emp)	-1.60E-03	8.034e-04	-1.989	0.051398	.
log(cda)	2.919e-06	1.117e-06	2.614	0.011346	*
log(cgdpo)	-6.34E-06	3.812e-06	-1.662	0.101731	
log(cn)	-4.48E-07	1.221e-07	-3.67	0.000523	***
log(rgdpna)	9.110e-07	4.691e-07	1.942	0.056916	.
log(rdana)	-3.89E-06	1.145e-06	-3.395	0.001233	**
log(rnna)	4.538e-07	1.238e-07	3.666	0.000529	***
log(pl_da)	1.981e+00	7.269e-01	2.725	0.008451	**
log(csh_c)	8.389e-01	2.929e-01	2.864	0.005789	**
log(csh_i)	1.166e+00	4.481e-01	2.601	0.011724	*
log(csh_x)	3.767e-01	2.072e-01	1.818	0.074152	.
log(csh_m)	5.061e-01	2.025e-01	2.499	0.015261	*
log(pl_c)	-1.68E+00	6.655e-01	-2.529	0.01414	*
log(pl_x)	5.576e-01	4.937e-01	1.129	0.263286	
log(pl_n)	-8.31E-01	1.613e-01	-5.15	3.14e-06	***
log(pl_k)	1.069e-01	5.626e-02	1.9	0.062294	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Overall, the final model "explain" 69.9% variability in income inequality; equivalently, the study also use cross-validation in order to test the performance of the model in the real world. The result shows the change in R-square only approximately 0.4, which also indicates the model demonstrates better generalizability. In addition, as previously discussed, the model is overestimating Russian Federation(#66). It is noteworthy that the reason the observation is removed is because the study need to realize the maximum fitting effect just for this dataset rather than #66 is incorrect. From another perspective, this is a limitation for our model as it does not fit well for some subjects which are from Latin America. Furthermore, another limitation of our final model is due to the shortcomings of dataset itself. For one thing, artificially insert some inaccurate data for the missing values of the original data could greatly bias for our fitted model; for another, even though our model has statistically and economically significant implication, the variables chosen are probably not the best measures of economic inequality.

5. Discussion and conclusion

5.1. Discussion

5.1.1. Differences from literature review

Literature reviews generally believe that globalization has exacerbated income inequality between and within developed and developing countries through capital flows and technological advantages

[6,9]. However, the quantitative analysis in this study shows that the impact of globalization on economic inequality is multidimensional: although international trade(such as export price levels) is positively correlated with the GINI index, the significant negative correlation of certain domestic economic variables(such as the share of labor income and consumer price levels) indicates that policy intervention can partially offset the negative effects of globalization. For example, there are also vision models in the literature showing that Nordic countries have significantly reduced inequality through high welfare and tax policies, and providing proposals related to wealth taxes, but the literature rarely emphasizes the immediate regulatory role of domestic policies [12].

5.1.2. Limitations

The study has the following limitations. At the beginning, for the dataset, missing values for some countries in the World Bank data are processed through interpolation or deletion, which may contribute to model bias. For example, the economic particularities of Russia Federation(#66) and Uruguay(#76) are not fully captured. Moreover, in the term of model construction, multivariate linear regression assumes a linear relationship between variables, but the interaction between globalization and inequality may show non-linear or dynamic characteristics. Besides, in terms of geographical coverage, the model has low fit for Latin American and African countries, which may be owing to insufficient data representativeness or structural differences.

5.2. Conclusion and future

5.2.1. Conclusion

The study uses a multivariate linear regression model to analyze the impact of globalization on income inequality after 2008 economic crisis and finds the following key results. First of all, the export price levels(log(pl_x)) and capital price(log(pl_x)) are significantly positively correlated with GINI index, indicating that globalization has exacerbated income inequality through trade and capital flows. Secondly, consumer price (log(pl_c)) and labor income share(labsh) are negatively correlated, proving that domestic policies(such as price regulation and labor income distribution) can effectively curb income inequality. Last but not least, the GINI index of high-welfare countries(like Norway and the Czech Republic) is lower than that of countries that rely on a singke industry(such as Chile, South Africa and Russian Federation), highlighting the importance of institutional design and that R&D investment is an important balancing force.

5.2.2. Policy recommendations

Based on the research results, the following three policy recommendations are put forward. First, reform the tax system, implement progressive capital tax and cross-border tax cooperation to reduce the concentration of capital gains [12], and strengthen tax audits on high-income groups. Second, strengthen social security, expand unemployment insurance coverage and increase job transfer training for low-skilled workers who are vulnerable to the impact of globalization to enhance their economic resilience. Third, diversify industries and provide international assistance to developing countries that rely on a single energy industry(like Chile) to promote technological upgrading and market access for small and medium-sized enterprises to reduce the shock of price fluctuations on income distribution.

References

[1] Furusawa, T., Konishi, H. and Tran, D.L.A. (2019). International Trade and Income Inequality*. The Scandinavian Journal of Economics, 122(3), pp.993–1026. doi: https: //doi.org/10.1111/sjoe.12360.

[2] Shin, I. (2012). Income inequality and economic growth. Economic Modelling, 29(5), pp.2049–2057. doi: https: //doi.org/10.1016/j.econmod.2012.02.011.

[3] Roser, M. and Cuaresma, J.C. (2014). Why is Income Inequality Increasing in the Developed World? Review of Income and Wealth, [online] 62(1), pp.1–27. doi: https: //doi.org/10.1111/roiw.12153.

[4] Global Energy Monitor (2024). CAP Acero Huachipato steel plant - Global Energy Monitor. [online] Global Energy Monitor. Available at: https: //www.gem.wiki/CAP_Acero_Huachipato_steel_plant.

[5] Friedman, T. (2005) 'It's a Flat World, After All', The New York Times, 3 April. Available at: https: //www.nytimes.com/2005/04/03/magazine/its-a-flat-world-after-all.html.

[6] Gereffi, G. (2019) 'Global value chains and development: redefining the contours of 21st century capitalism', Journal of Economic Geography, 19(5), pp. 1169–1170. Available at: https: //doi.org/10.1093/jeg/lbz015.

[7] Wade, R.H. (2004) 'Is Globalization Reducing Poverty and Inequality?', World Development, 32(4), pp. 567–589.

[8] Dorn, D. and Hanson, Gordon.H. (2013) 'The China Syndrome: Local Labor Market Effects of Import Competition in the United States', American Economic Review, 103(6), pp. 2121–2168. Available at: https: //ideas.repec.org/a/aea/aecrev/v103y2013i6p2121-68.html.

[9] Rodrik, D. (2011) The globalization paradox, eprints.lse.ac.uk. Edited by D. Rodrik. Oxford, UK: Oxford University Press. Available at: https: //eprints.lse.ac.uk/57253/.

[10] Harrison, A. (2007) Title: Globalization and Poverty: An Introduction. Available at: https: //www.nber.org/system/files/chapters/c10713/c10713.pdf.

[11] Dollar, D. and Kraay, A. (2002) 'Growth is Good for the Poor', Journal of Economic Growth, 7(3), pp. 195–225. Available at: https: //doi.org/10.1023/a: 1020139631000.

[12] Piketty, T. (2014) Capital in the Twenty-First Century — Harvard University Press, Harvard University Press. Available at: https: //www.hup.harvard.edu/books/9780674430006.