The Effects of Macroeconomic Variables on Income Inequality Based on Cross-Section Analysis

Bingyue Bai(✉)

College of Science and Technology, Guizhou University, Guiyang, Guizhou, China
bingyuebai9786@163.com

Abstract. Almost all countries in the world are experiencing economic growth and technological progress, but the gap between rich and poor is also steadily widening. As a result, the income inequality has become an important issue that every country must face. Using cross-sectional data from 54 different countries in 2013, this study examines the influencing factors of income inequality. Data analysis shows that external balance, R&D expenditure and inflation have a positive impact on the Gini index. Moreover, higher tariff levels and exports of goods and services will reduce income inequality. This paper helps to explain the problems with macroeconomic tools and the causes of income inequality.

Keywords: Income inequality · R&D expenditure · GINI index · Tariff rate

1 Introduction

Economic development often leads to a widening gap between rich and poor. Given the complexity of income inequality and its causes, further research is necessary to better understand the impact of macroeconomic factors on income redistribution, such as R&D expenditure and tariff rate, and how to manage these factors to reduce income inequality. There exists a considerable body of literature on the determinants of income inequality. Hoffmann et al. (2020) emphasize the U.S. national income gap has become wider because of the progressively larger educated workforce, which has higher wages [1]. Therefore, education is one of the important factors that should be taken into consideration. Using panel data from India and Pakistan, Munir and Sultan (2017) found that fertility was one of the determinants of income inequality from 1973 to 2015 [2]. In addition, Ataguba (2021) explores the relationship between financing health services and income inequality. Based on Ataguba's research, funding for health services can reverse a widening income gap between the rich and poor [3]. In addition to health services, Włodarczyk (2017) believes that there is also a complex relationship between innovation and income distribution. Wlodarczyk (2017) points out that higher r&d spending increases income inequality, but higher patent numbers actually reduce it [4]. Overall, these studies highlight the need for focus the effect of the different macroeconomic factors on income inequality.

F. Balli et al. (Eds.): ESFCT 2022, AEBMR 226, pp. 345–353, 2022.
https://doi.org/10.2991/978-94-6463-052-7_40

Previous studies based on time series regression model analysis, almost only for specific countries or regions. The purpose of this paper is to analyze the cross-sectional regression model based on selected macroeconomic factors. The variables related to international trade can be measured by the balance of foreign trade in goods, services exports of goods and services and tariff rates. Besides, government expenditure and the inflation will be considered, such as health service and research and development. The data will be retrieved from 54 different countries and regions in 2013 through World bank database.

2 Data and Methodology

This study is going to use general-to-specific modeling as the main approach. While there are many ways to measure the income gap, the Gini coefficient is the most commonly used method to measure the income gap between people. Meanwhile, in order to avoid some possible problems, the functional form of GINI index will adjusted to natural logarithmic forml. On this basis, we apply ordinary least squares (OLS) to estimate variables to attenuate the adverse effects of measurement errors on individual observations (Sarel, 1997) [5]. Furthermore, in order to ensure the accuracy of regression model, Jarque-Bera test and Breusch-Godfrey LM test were carried out. Then, the Breusch-Pagan and Ramsey reset tests will be used to examine the heteroskedasticity and functional form. After the regression model has passed all diagnostic tests, the overall significance of model will be tested through F-test. Moreover, we will remove some non-significant explanatory variables. Finally, the model with the minimum AIC value will be the best-fitted model through the eviews output.

As reported in Table 1, data are collected from World Bank database including six macroeconomic factors as explanatory variables and one indicator as a explained variable. These data are mainly used to measure three measures of international trade (external balance on goods and services, exports of goods and services and tariff rate), two measures of government expenditure (research and development expenditure and health expenditure) and other variables (inflation consumer prices). The Gini coefficient is the dependent variable of the model.

Table 1. Determinants of income inequality

Variables	Index (Data: World Bank Database)
Y	**GINI coefficient**
X1	**Research and development expenditure (% of GDP)**
X2	**Health expenditure public (% of GDP)**
X3	**Inflation consumer prices (annual %)**
X4	**External balance on goods and services (% of GDP)**
X5	**Tariff rate, applied, weighted mean, all products (%)**
X6	**Exports of goods and services (% of GDP)**

3 Main Results

The first step is to use the logarithm of the GINI index, which ranges roughly from 0 to 100. Then the basic cross section model is established:

$$\mathrm{LogY} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u \quad (1)$$

The optimal linear unbiased regression model should satisfy six classical assumptions, so it is necessary to perform some examination. First, multicollinearity leads to estimation bias, so it must be tested with VIF. According to the eviews output, the VIF value of each explanatory variables is less than 10, which means the basic model does not have serious multicollearity problem. Second, the model passes the Jarque-Bera test, indicating that the residual term conforms to normal distribution.

Several diagnostic tests should be performed on the base model in step 2, such as Breusch-Pagan-Godfrey test, Breush-Godfrey LM test, Ramsey RESET test. From the results of eviews, the basic model (model 1) does not have heteroskedasticity and serial correlation problems, because the p-values of both F-statistics are greater than 0.05. Moreover, it is necessary to check whether the model has square terms or interaction terms. According to the results of the Ramsey RESET test, the model has passed the test. The results of eviews are shown in Fig. 2, the P-value of F-statistic for the basic model is 0.1987, which is less than 0.05, showing the basic model is reasonable.

The third step is to select some independent variables according to the significance of variables, so as to find a better regression model. As shown in Fig. 1, it is clear that X2 and X3 need to be taken into account, because they both have P-values above 0.05, which indicates that they are insignificant at the 5% level of significance. Then remove them from the model and regression again to get two different regression models. Firstly, the model (2) will be established by removing X3. As reported in Fig. 3, although the model passes all diagnostic tests, there are still some unimportant variables in the model. Take X2 and X5 as an example, their P-values are higher than 0.05, which means they are insignificant at the 5% level of significance.

$$\mathrm{LogY} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u \quad (2)$$

Variable	t-Statistic	Prob.
C	34.61777	0
X1	-3.04162	0.0038
X2	1.218499	0.2291
X3	-1.95318	0.0568
X4	2.637461	0.0113
X5	2.557755	0.0138
X6	-2.47564	0.017

Fig. 1. t-Statistics for model (1)

	Ramsey-RESET		Breush-Pagan-Godfrey		Breusch-Godfrey-LM	
	F-Statistic	P-value	F-Statistic	P-value	F-Statistic	P-value
Model(1)	1.675405	0.1987	1.017281	0.4258	0.550828	0.5803
Model (2)	1.567794	0.2194	1.472196	0.2164	1.113583	0.3371
Model (3)	0.363476	0.6972	1.675156	0.1587	0.604923	0.5504

Fig. 2. Diagnose tests for three models

Variable	t-Statistic	Prob.
C	34.21982	0
X1	-2.86825	0.0061
X2	1.336771	0.1876
X4	2.510629	0.0155
X5	1.920952	0.0607
X6	-2.27976	0.0271

Fig. 3. t-Statistics for model (2)

Both model (3) and model (2) pass all diagnostic tests, but all explanatory variables are significant and AIC is minimal. After comparing the various scenarios, model 3 becomes the final model.

$$\mathrm{Log}Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u \tag{3}$$

According to the Fig. 4, it is obvious that the coefficient of research and development expenditure, inflation and exports goods and services are negative, indicating that these variables can significantly reduce the Gini index. In other words, these three variables help narrow the income gap between rich and poor. On the contrary, external balance on goods and services and tariff rate have positive relationship with GINI index, which means that these two variables will increase income inequality.

Additionally, as each country's level of development varies, the impact of these variables may vary. It is necessary to use dummy variable to identify the omitted categorical effect, which might get different results. Therefore, the countries are divided into developing and developed countries according to their level of development in this model.

$$\mathrm{Log}Y = \beta_0 + \delta_0 \mathrm{developed} + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u \tag{4}$$

Developed = 1, when the country is a developed country.

$$\mathrm{Log}Y = \beta_0 + \delta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u$$

Dependent variable: Log GINI index					
C	X1	X3	X4	X5	X6
3.7221	-	-	0.0045	0.0298	-
21	0.0882	0.0164	27	56	0.0022
	27	97			89
R^2: 0.389309					

Fig. 4. Coefficients of model (3)

$$E(LogY|developed = 1) = \beta_0 + \delta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$$

Developed = 0, when the country is a developing country.

$$LogY = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u$$

$$E(LogY|developed = 0) = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$$

δ0 is the difference in log Y between developed and developing countries given the same level of X1, X3, X4, X5, X6 and residual term. In order to work out whether these two categories have impact on outcome or not, t-test would be used.

H0: $\delta 0 = 0$;
H1: $\delta 0 \neq 0$
$t - \text{Statistic} = \frac{\hat{\delta}_0 - 0}{s.e.\left(\hat{\delta}_0\right)} = 0.430453$
For 5% significance level, $df = n - k - 1 = 47$, t–critical = 2.011
t – Statistic < t –critical; reject H1.

According to the results, there is no difference between developing countries and developed countries.

4 Analysis and Limitation

Due to the data limitations, the cross-sectional regression model in 2013 only include 54 countries and regions. At the same time, the developed countries are different from developing countries in general, but it is indifferent between developed countries and developing countries in this model. This result differs from other studies. Taking the exports of goods and services as an example, this study finds that exports of goods and services are negatively correlated with income inequality. Hazama (2017) believes that the impact of exports of goods and services was different from that of high-income and low-income developing countries. Hazama illustrates that export levels may widen the income gap in low-income developing countries, but not significantly in high-income developing countries [6]. Aradhyula et al. (2007) studied the impact of the trade openness on income inequality, they also compared the difference between developed countries and developing countries. They found the same results as Hazama's research, and they assume that developed countries are less sensitive to exports because they have better democratic governance [7]. Thus, there may be some problems with the model, and the next step should be to increase the number of countries to calibrate the model.

The empirical results of this study provide some enlightenment. In order to reduce income inequality, government should focus on international trade, so decreasing the level of tariff rate and external balance may be effective ways to reverse the widening income gap. Vallejos and Turnovsky (2017) pointed out that the reduction of tariff rates may aggravate the inequality of income distribution in the long run. Meanwhile, the impact of tariffs on income inequality depends on a country's development level. In addition, the inflation has negative relationship with GINI index [8], and the results of this study are quite different to Law and Soon (2020). They found that the higher inflation rate will exacerbate the income inequality according to their research on different developed and developing countries. Moreover, government's spending on research and development may reduce income inequality [9]. Chu and Cozzi (2017) obtained different results, finding that increasing patent protection increases income inequality and consumption inequality, while increasing R&D subsidies reduces income inequality and consumption inequality [10].

5 Conclusion

This study explores the relationship between income inequality and several macroeconomic variables based on the data from 54 different countries in 2013. On the one hand, government's spending on research and development, inflation and exports of goods and service reduce income inequality. According to the research results, although technological progress is important for every country, the government should balance the relationship between technology spending and the inequality of income distribution. At the same time, policy makers should also pay attention to the impact of inflation and external balance. On the other hand, tariff rate and external balance on goods and service increase income inequality. However, the conclusions about tariffs differ from other studies. Furthermore, the developed countries and developing countries are indifference in this model, which is a questionable result. There could be more attention and work to analyze income inequality in the future.

Appendix

Model (1) $\text{LogY} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u$

Model (2) $\text{LogY} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u$

Model (3) $\text{LogY} = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + u$

Model (1)

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 04/17/18 Time: 20:04
Sample: 1 54
Included observations: 54

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.635928	0.105031	34.61777	0.0000
X1	-0.126440	0.041570	-3.041623	0.0038
X2	0.021507	0.017650	1.218499	0.2291
X3	-0.015730	0.008054	-1.953180	0.0568
X4	0.004526	0.001716	2.637461	0.0113
X5	0.033376	0.013049	2.557755	0.0138
X6	-0.002213	0.000894	-2.475642	0.0170

R-squared	0.408010	Mean dependent var	3.528237
Adjusted R-squared	0.332437	S.D. dependent var	0.202203
S.E. of regression	0.165209	Akaike info criterion	-0.642787
Sum squared resid	1.282821	Schwarz criterion	-0.384955
Log likelihood	24.35524	Hannan-Quinn criter.	-0.543351
F-statistic	5.398868	Durbin-Watson stat	1.833645
Prob(F-statistic)	0.000261		

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	1.017281	Prob. F(6,47)	0.4258
Obs*R-squared	6.206709	Prob. Chi-Square(6)	0.4004
Scaled explained SS	5.271676	Prob. Chi-Square(6)	0.5095

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.550828	Prob. F(2,45)	0.5803
Obs*R-squared	1.290396	Prob. Chi-Square(2)	0.5246

Ramsey RESET Test
Equation: UNTITLED
Specification: LOG(Y) C X1 X2 X3 X4 X5 X6
Omitted Variables: Powers of fitted values from 2 to 3

	Value	df	Probability
F-statistic	1.675405	(2, 45)	0.1987
Likelihood ratio	3.878306	2	0.1438

Model (3)

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 04/17/18 Time: 20:17
Sample: 1 54
Included observations: 54

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.722121	0.078029	47.70166	0.0000
X1	-0.088227	0.027424	-3.217134	0.0023
X3	-0.016497	0.008069	-2.044435	0.0464
X4	0.004527	0.001725	2.625132	0.0116
X5	0.029856	0.012789	2.334472	0.0238
X6	-0.002289	0.000896	-2.553843	0.0139

R-squared	0.389309	Mean dependent var	3.528237
Adjusted R-squared	0.325695	S.D. dependent var	0.202203
S.E. of regression	0.166041	Akaike info criterion	-0.648722
Sum squared resid	1.323345	Schwarz criterion	-0.427724
Log likelihood	23.51550	Hannan-Quinn criter.	-0.563492
F-statistic	6.119890	Durbin-Watson stat	1.862833
Prob(F-statistic)	0.000184		

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	1.675156	Prob. F(5,48)	0.1587
Obs*R-squared	8.022810	Prob. Chi-Square(5)	0.1550
Scaled explained SS	8.569613	Prob. Chi-Square(5)	0.1275

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.604923	Prob. F(2,46)	0.5504
Obs*R-squared	1.383857	Prob. Chi-Square(2)	0.5006

Ramsey RESET Test
Equation: UNTITLED
Specification: LOG(Y) C X1 X3 X4 X5 X6
Omitted Variables: Powers of fitted values from 2 to 3

	Value	df	Probability
F-statistic	0.363476	(2, 46)	0.6972
Likelihood ratio	0.846706	2	0.6548

Dummy variable

Dependent Variable: LOG(Y)
Method: Least Squares
Date: 04/20/18 Time: 20:48
Sample: 1 54
Included observations: 54

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.726158	0.079257	47.01380	0.0000
X1	-0.099898	0.038732	-2.579178	0.0131
X3	-0.016353	0.008146	-2.007622	0.0505
X4	0.004562	0.001741	2.619971	0.0118
X5	0.029954	0.012901	2.321803	0.0246
X6	-0.002273	0.000905	-2.511489	0.0155
DEVELOPED	0.035212	0.081802	0.430453	0.6688

R-squared	0.391707	Mean dependent var	3.528237
Adjusted R-squared	0.314052	S.D. dependent var	0.202203
S.E. of regression	0.167469	Akaike info criterion	-0.615620
Sum squared resid	1.318148	Schwarz criterion	-0.357788
Log likelihood	23.62173	Hannan-Quinn criter.	-0.516184
F-statistic	5.044229	Durbin-Watson stat	1.883604
Prob(F-statistic)	0.000459		

References

1. F. Hoffmann, D.S. Lee, T. Lemieux, "Growing income inequality in the United States and other advanced economies." *Journal of Economic Perspectives* 34.4 (2020): 52-78.
2. K. Munir, M. Sultan, “Macroeconomic determinants of income inequality in India and Pakistan.” *Theoretical & Applied Economics* 24.4 (2017).
3. J. E. Ataguba, “The impact of financing health services on income inequality in an unequal Society: the case of South Africa.” *Applied Health Economics and Health Policy* 19.5 (2021): 721-733
4. J. Wlodarczyk, “Innovations and income inequalities – a comparative study.” *Journal of international studies* 10.4 (2017): 166-178.
5. M. M. Sarel, How macroeconomic factors affect income distribution: The cross-country evidence. *International Monetary Fund*, 1997.
6. Y. Hazama, “The impact of exports on income inequality in developing countries.” (2017): 261-0014.
7. S. Aradhyula, T. Rahman, K. Seenivasan, Impact of international trade on income and income inequality. No. 381-2016-22241. 2007.

8. J. R. Vallejos, S. J. Turnovsky, "Tariff reduction and income inequality: Some empirical evidence." *Open Economies Review* 28.4 (2017): 603-631.
9. C. H. Law, S. V. Soon, "The impact of inflation on income inequality: the role of institutional quality." Applied Economics Letters 27.21 (2020): 1735-1738.
10. A. C. Chu, G. Cozzi, "Effects of patents versus R&D subsidies on income inequality." Review of Economic Dynamics 29 (2018): 68-84.