

**The Effects of Tertiary Education Attainment and Globalization through FDI on the
Within-country Income Inequality**

Empirical evidence from 34 OECD countries in 2005-2015

Linh Anh Le

Beloit College

Introduction

According to Keeley (2015), starting from the last decades of the 20th century, income inequality increased within some rich economies countries and countries with high economic rise, causing great concern among economists and policy makers. From a cross-country panel data of 84 countries over 3 decades from 1965, Barro(2000) showed that the influence of income inequality on rates of economic growth and investment varied across richer and poorer countries. “Growth tends to fall with greater inequality when per capita GDP is below around \$2000 (1985 U.S.dollars) and to rise with inequality when per capita GDP is above \$2000” (p. 29). Additionally, income inequality could lead to many other social issues besides economic growth. Malderen et al.(2013) proved that the determinants of wealth-related inequality, including household wealth and occupation, negatively affect under-5 mortality in 13 African countries. Stiglitz(2012) made a critical claim that a society with an extreme inequality can lead to the malfunction of the whole system, “if the belief takes hold that the political system is stacked, that it’s unfair, individuals will feel released from the obligations of civic virtue” (p.109). This means people lose trust in their government and social bonds are disrupted.

Acknowledging many severe impacts of within-country income inequality, researchers have focused on trying to understand its driving forces, including globalization (Yeung, 2002; Ha, 2012; Meschi&Vivarelli, 2009). While the definition of globalization is broad and ambiguous, this research paper will focus on the economic globalization as the cause of income inequality. Yeung (2002) highlighted two interesting political perspectives towards economic globalization. One perspective values globalization as an inevitable route towards prosperity and economic growth, which means “if a national economy is not performing well, it must be because the economy is not having "enough" economic globalization” (Yeung, p.288). On the other hand, the wave of anti-globalization or anti-globalism has been blaming basically every problem that a country confronts with on globalization. President of the United States, Donald Trump, during his presidential term from 2017, has supported the latter perspective through both his public statements and also policies. He claimed that the U.S. “will no longer allow our workers to be victimized, our companies to be cheated and our wealth to be plundered and transferred” under the belief that international trade has detrimental effects on the growth of the developed economies like the U.S’s. Developing and poorer countries also witness some negative impacts of globalization, especially on the distribution of wealth. Trade flow and FDI caused income inequality in a less developed

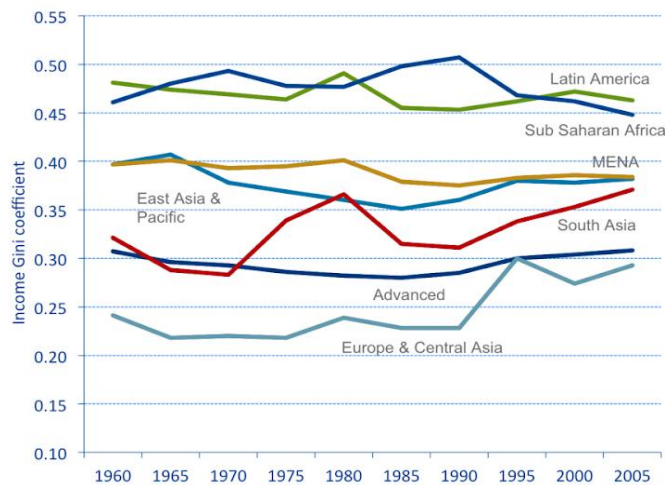
economy, especially in the countries with the rightist political regimes (Ha, 2012). Meschi and Vivarelli (2009) stated that trading with higher income countries worsened the wealth distribution in developing countries. For developing countries, it is harder to reverse the ongoing cycle of globalization or to choose to limit their exposure, unlike the U.S. Consequently, in order to tackle inequality issues in these countries, it is necessary to understand how globalization, through foreign investment and trading, affect the countries, for better policy implementation.

Education is the solution that the United Nations has consistently promoted to reduce income inequality and other poverty-related issues. In the 2017 report, *Learning to realize education's promises*, the U.N. advocated for education and education coverage to better support poor people, and encourage social mobility. Behrmana et al. (2017) argued that even though increasing parental education does not necessarily have significant impacts on the next generation's poverty, it is positively correlated with welfare among current poor households, or current generation's inequality.

In summary, globalization and education are both important factors that have been proved to significantly influence the within-country income inequality. Kishan (2017), through econometrics model, showed that the interaction of globalization and education on income inequality varied substantially based on where the country is in education attainment level. This means when the country starts with a low level of education, increasing education and globalization will help alleviate income inequality. However, there is a turn-around point when more educated people, especially in higher education, globalization worsens income inequality. Keeley (2015) wrote that within the OECD groups, some countries showed an increase in their Gini Coefficient and some had relatively low income inequality, such as Denmark, Norway, the Slovak Republic and Slovenia. Using fixed effect model, this research aims at answering the question whether globalization and higher education attainment level affect income inequality measured by Gini index in 34 OECD bloc countries during the 2005-2015 period. Globalization through FDI is expected to have negative impacts on wealth distribution and worsened the income inequality. The push towards higher education, which has made more people have college degrees, especially in technical fields, should lower the income inequality.

Literature Review

Growing income inequality in many countries and regions has been one of the most popular topics among sociologists and economists. A number of research has proved that inequality could lead to many social issues that capture the attention of politicians all around the world, such as health access inequality (Van Malderen, 2013), political instability and policy volatility (Dutt and Mitra, 2007; Stiglitz, 2012), and an obstacle for economic growth (Berg and Ostry, 2011). According to Castelló-Climent and Doménech (2014), growing income inequality is, indeed, an international trend.



Graph 1: Evolution of the income Gini coefficient across regions, 1960-2005

(Castello-Climent and Domenech, 2014, extracted from World Bank)

To capture the development of income inequality, researchers have been using different indices of measurement. Some researchers use methods such as the mean logarithmic deviation of income (MLD) index of inequality (Sala-i-Martin, 2006), or UTIP-UNIDO and EHII from University of Texas Inequality Project (UTIP) database (Meschi and Vivareli, 2009). Among those inequality measures, Gini index has been by far the most popular (Goldberg and Pavcnik, 2007; Ha, 2012; Alderson and Nielson, 2012). Gini index measures how much the incomes of individuals and households deviate from a perfectly equal distribution. This index ranges from 0, meaning perfect equality, to 1, perfect inequality. Mills (2009) pointed out some serious conceptual, methodological, and definitional issues when using Gini index to measure cross-country income inequality comparison. However, since I focus on the determining factors that contribute to the within-country inequality so these problems would not affect the analysis's results.

It is important to also understand the difference between income and wealth inequality, both of which have been increasing in developed and developing countries within the last 30 years. Wealth and income do not always have a direct relationship. Balestra and Tonkin (2018) argue that Denmark features “a low level of income inequality, ranked 22nd out of the 27 OECD countries included but the third highest level of wealth inequality” (p.25). In addition, “Wealth inequality is always higher than labor income inequality, due to income originated from wealth (or capital income), such as rents, dividends or royalties and the increase of asset values” (Bernan, p.19). Wealth is about everything a person or household owns. On the other hand, income, according to the common way of measurement like Gini index, only stems from labor. This makes wealth harder to calculate, even though it could be more comprehensive with regards to a person or a country’s economic inequality. In this research, I choose income inequality because it is still worth investigating and there is more data available than wealth inequality.

Globalization, according to Yeung(2002), is the cause of every problem one country encounters for anti-globalism group including income inequality. To measure the impact of globalization on income inequality, some research uses foreign direct investment (FDI). Alderson & Nielson (2002) uses generalized linear model to regress income inequality measured by Gini Index on offshore direct investment, female labor force participation, proportions of workforce in agriculture and manufacturing, as well as other explanatory factors. This research shows that direct investment has negative impacts on income inequality in 16 OECD countries during 1967-1992 period. This may happen through different channels. Similarly, Ha(2012) argues that globalization through FDI, instead of yielding positive impacts like more capital stock and labor productivity, causes inequality in low developing countries. The influence of political regime, whether it is more rightist or leftist, could also significantly intensify or alleviate that relationship between FDI and income inequality.

Some research uses trade as a measurement of globalization’s impacts on income inequality. Meschi and Vivarelli (2009) suggest that trade did not significantly impact the within country income inequality for 65 developing countries from 1980 to 1999. However, they point out that trading with higher income countries worsened the wealth distribution in those countries. The use of inflation rate in this research increases the validity since different inflation rate would affect real wage and consumption power of the individuals. However, the impacts of trade on

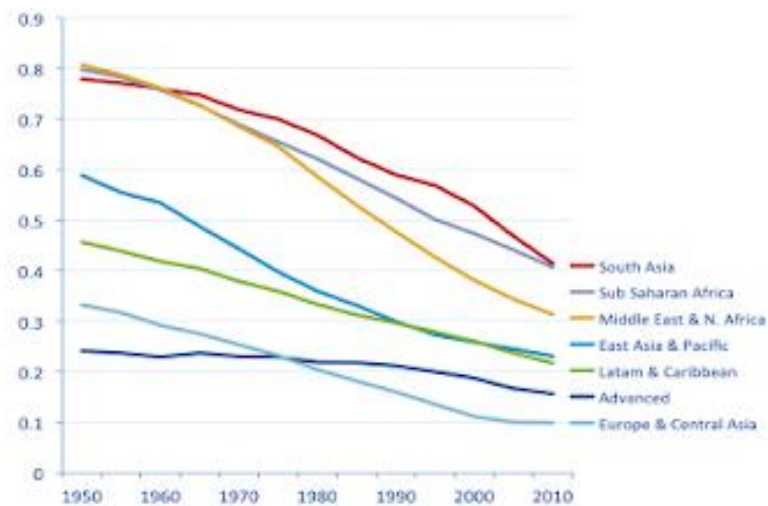
income inequality could be is not straightforward, and geographical and political characteristics may play a great role in this (Silva, 2007; Babones and Zhang, 2008; Goldberg and Pavcnik, 2007).

Education is the solution that the United Nations has consistently promoted to reduce income inequality and poverty-related issues. Under the continuous trend of globalization, there is also a push towards better education and higher education to meet the international and local needs in developing countries (Green et.al, 1997). According to Nayyar(2007), globalization influences the format, courses and needs for higher education, “market forces, driven by the threat of competition or the lure of profit, have led to the emergence of higher education as business” (p.32). However, education, in some cases, increases inequality within a nation. Through logistic regression, Pfeffer(2018) argues the positive influence of higher education on household wealth was statistically significance, and also household’s income impact the child’s education attainment. The college premium that leads to wealth gap expands among households with different education level and it may continue to the next generation.

Although there has been multiple research on the influence of education and globalization on income inequality separately, there has not been much work done on the combined impact of both factors. Kishan(2017) uses panel data analysis to examine the impact of globalization on income inequality and how education intervenes as a mediating force. Kishan (2017) uses World Bank’s public data on their website, together with some international sources of countries that had available data during the 1980-2005 period. To measure globalization, this research uses two indices, including Gini and EHII. The independent variables include KOF index of globalization (measured by World Bank), and mean years of schooling. This way of choosing the measurement for globalization is different from most other research, which Kishan (2017) argues that could make up for the incompleteness of using either trade or/and FDI. Overall, this research concludes that globalization helps lower inequality within a country when the country has a high proportion of its population being illiterate and low educated. When the education attainment level increases, globalization’s restorative impact on income inequality is weakened. These findings are provocative, however, the way the authors randomly chooses which countries to be in the sample makes the regression susceptible to selection bias.

Castelló-Climenta and Doménecha (2017) find that there was a consistent decrease in education Gini coefficient, a measure of education inequality (graph 2). This is contrary to the

increasing income Gini coefficient across regions during 1960 and 2005 (graph 1). The authors analyze 146 countries with the data from World Bank. Variance in education inequality is larger than that in income inequality. In both large sample and across world region data sets, while income inequality is stable, and tend to increase, human capital inequality plummeted. With the assumption that the changes in technology and globalization impact these opposite trends, Castelló-Climent and Doménech (2017) use OLS estimation and regress natural log of income Gini index on proportion of low educated and higher educated works with ratio of net export on GDP. One finding of this research is the fall in human capital inequality leads to the decrease in income inequality, but is offset by the increase in demand for high-tech workers and globalization. This is slightly different from Kishan(2017), but mostly due to the difference in their initial research questions and their direction of including education to the model.



Graph 2: Evolution of the education Gini coefficient across regions, 1950-2005
(Castello-Climent and Domenech, 2014, extracted from World Bank)

Overall, there is a consensus among previous research that globalization may cause detrimental impacts, including worsened income inequality. Education is evidently a factor that impacts income inequality, however, the direction of effect is not clear yet. However, there are many other factors that also affect the income inequality of one country. Using fixed effect regression with both countries and years, I hope to examine which factors affect 34 OECD countries, with a focus on FDI and tertiary education's attainment level. I also include other variables that measure the economic development and population composition of one country.

Empirical Methodology

Conceptual and Theoretical model

Income inequality within a nation is an evidence of a stratified society and a hindrance for economic development. It is important for policy makers to understand the causal factors of income inequality. Previous research highlights that globalization and tertiary education together have strong influences on the level of inequality for both household and individual's income (Kishan, 2017, Castello-Climent&Domenecha, 2017). The promotion for higher education is prevalent in both developed and developing countries. The UN also mentioned giving people more chances for higher education is also a way to alleviate income inequality, by giving people more skills to make their ways out of poverty. However, this is still debatable. Indeed, there is an increasing amount of work that claimed the opposite trend that more people going to college leads to higher wage premium, or income inequality (Hendel et al., 2003).

Within the scope of this research paper, I chose to focus on more developed economies, particularly OECD countries during the period of 2005-2015. "In all OECD and partner countries, except Israel and South Africa, the share of younger adults (25-34 year-olds) with tertiary qualifications is larger than that of older adults (55-64 year-olds) with that level of qualification" (IMHE, 2017). FDI could also represent globalization's influence in OECD countries fairly well, because they are both the major distributors and receivers of FDI. Using fixed effect regression model, I would investigate the impact of growing population having tertiary education, and FDI on the income inequality. The regression model also includes some controlled variables related to the economic development and population composition of those countries. The first hypothesis is that globalization through FDI had negative impacts on wealth distribution and worsened the income inequality in OECD countries. The second hypothesis is that the push towards higher education, which has made more people have college degrees, especially in technical fields, combined with globalization, positively affects the income gap.

Data

This research paper focuses on the factors that affect income inequality in the OECD countries during the 2005-2015 period.

I pay close attention to select the data that best represents the level of inequality within a country to be my independent variable. There are many different ways that income inequality has been collected and measured, and no method is completely superior than others. To test the robustness of the tertiary education's and FDI's estimates on income inequality, I use Standardized World Income Inequality Database (SWIID)'s Gini index on disposable income (post tax, post transfer) collected by Solt (2019). This is the most comprehensive data set on Gini index, and there is no missing value. I have 374 observations for 34 countries in 11 years. Solt (2019) claims that this index is calculated "to meet the needs of those engaged in broadly cross-national research by maximizing the comparability of income inequality data while maintaining the widest possible coverage across countries and over time". Despite the completeness of this data set, a lot of researchers and economists have shown concerns about its validity. Jenkins (2015) did some thorough work on proving the problem of SWIID Gini coefficients due to the multiple imputation that Solt used to generate the missing values. Even though I acknowledged the limitation of this variable, within the time frame of this research, I could not find a better proxy that measures income inequality continuously for all 34 countries I have in my sample. Income inequality is a major concern and there should be a more consistent data on this

For the two independent variables of interest, FDI and tertiary education attainment, I collected the data from two sources, namely World Bank and OECD's public data called Adult education level, respectively. FDI data from World Bank was fairly comprehensive with all 374 observations. However, there are 17 missing values in the highest level of education (tertiary education) completed by the 25-64 year-old population. The countries with missing values are New Zealand, Chile and Israel. I lost 9 out of 17 data points for New Zealand, 7 for Chile and 1 for Israel.

Since I mostly use World Bank as my main data source, their data is usually of good quality and continuity. For variables like inflation, GDP growth rate (consumer price), rural population, urban population, and total population, there are all 374 observations. However, for immigrant population, which is not available through World Bank, I extracted the data from OECD's public data on their website. There are 15 missing observations. The countries with missing values are Slovenia, Greece and Turkey, especially in the case of Turkey, there is only one available data point of year 2010.

Table 1: Variable Descriptions Table

Variable name	Description	Measurement units
gini_disp	SWIID Gini coefficient of the disposable income, which is post-tax, post-transfer income	Percentage (%)
top10income	Income share held by highest 10%	Percentage (%)
FDI	Net foreign direct investment, equal to net FDI outflow (asset) minus net FDI inflow (liabilities)	Billions of US dollar
terEduc	The proportion of the 25-64 year-old population completing tertiary education as their highest level of education	Percentage (%)
GDP_gr	The rate at which GDP at market prices based on constant local currency grows annually	Percentage (%)
Inflation	The annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly	Percentage(%)
crisis	Dummy variable for economic crisis (2005-2007: 0, 2008-2015: 1)	0 or 1
Rur_Pop	The number of people living in rural area	millions of people
Urb_Pop	The number of people living in urban area	millions of people
tpop	Total population	millions of people
imgr	The inflow of foreign population (measured through work permit, population registration, etc.)	millions of people
rurate	the proportion of people living in rural area out of total population	Percentage (%)

In order to have a clearer perspective on my research, I conducted summary statistics using STATA in table 2 (below)

Table 2: Summary statistics for all quantitative variables

	Mean	Standard deviation	Min	Max
Dependent variable				
gini_disp	30.76	5.43	23.4	46.7
Independent variables				
FDI	3.87	7.80	-2.97	734
terEduc	30.605	10.077	10.239	55.174
GDP_gr	1.963	3.434	-14.72	25.162
Inflation	2.326	2.117	-4.478	12.694
Rur_Pop	7.635	11.361	.0204	59.495
Urb_Pop	28.699	46.838	.0204	261.95
tpop	36.335	57.647	.297	320.74
imgr	.174	.261	.00098	2.016
rurate	22.78	11.381	2.12	48.467

The gini_disp variable has a relatively large range from 46.7% down to 23.4%. This means some countries in OECD experienced much more significant income inequality than the others. This is reasonable because even though they are in the same OECD block, some entered earlier than the other, some had more economic advantage over the others. Especially if we compare the U.S with Mexico, both are members of OECD but they are really different socially and economically. Rich people are highly capable of continuing to stay richer than the rest of the country, so the share they have is sustainably high. However, there are still 14 outliers.

Some negatives values in table 2 are also worth investigating. For FDI and inflation, there is no particular trend. However, for GDP growth rate, it was all positive for every country before the start of the world economic crisis in 2008. During and after the crisis, some OECD still struggled to recover and their GDP growth rate is negative.

The maximum value of the immigrant population was surprising to me that it could reach over 2 millions of people. When I looked more closely into the data, that country is Germany in year 2015. It is justifiable itself because Germany among other developed and more economically advanced countries are much more open to take in immigrants. However, this number is

substantially higher than other OECD countries and Germany itself in the previous years, which made me concern what happened during 2015 and who those immigrants were.

There are two variables that I planned to put into my regression, is the rural and urban population. Their correlation coefficient is 0.9974. I did not expect the correlation to be this high and also positive. There are couple of reasons why this happens, one of which is the growth in overall population leads to the growth in both rural and urban area. However, including both variables in the regression will reduce the unbiasedness of the estimates. Instead of doing this, I will include the rurrate and the tpop variables in my regression, and they do have low correlation.

One limitation in this data set relates to the immigrant population. Since the way the data is collected through each OECD country based on whether the immigrants are documented and rightfully stay in that country. Besides them, there are many other immigrants who came without documents, or came because of their home countries problem. Working without documents and law protection may make them vulnerable to exploitation and low payment. The selection bias will cause the coefficients to be overestimated. However, omitting this variable may also lead to omitted relevant variables that may have very strong relationship with FDI, tertiary education and also income inequality, especially with OECD. In addition, I also expect the benefit-cost trade-off between using gini_disp and top10income, due to the questionable validity of the former and the high missing observations caused by the latter, which may affect the fit of the model.

Model Specification

After reviewing literatures, I came up with two hypotheses about what affects income inequality. Globalization through FDI had negative impacts on wealth distribution and worsened the income inequality in OECD countries. The push towards higher education, which has made more people have college degrees, especially in technical fields, combined with globalization, postively affects the income gap. To test these hypotheses, I will run pooled OLS and fixed effect regression model

$$\text{gini_disp} = \beta_0 + \beta_1 * \text{FDI} + \beta_2 * \text{terEduc} + \beta_3 * \text{Xs} + u$$

My expectation for this model is that tertiary education's attainment level will be significant factor that affect the income inequality. For tertiary education, opposite to the trend in the previous literatures, I expect the result to show that more people going to tertiary education

means more people making it out of the poverty. Since my sample mostly contains developed countries, and strong economy will offer more opportunities to talented people. FDI, however, could be ambiguous because the inconsistency in the countries' policies, where some countries invested offshores more than receiving. But since the economy is already strong, more foreign investment may not necessarily lead to a really big impact to everyone in the society. Other economic development factors could be also ambiguous, however, I expect GDP growth to be significantly positive, especially after the economic crisis. For population compositions, I think rural population rate would be negative, which means the more people living in rural area, the more income inequality, particularly between rural and urban people. Total population may have negative impacts on income inequality, more people lead to more competition, and there will always be a loser and a winner. However, for immigrant population, one hypothesis I have is that they will have positive impacts, since this documented foreign population will contribute their abilities to the whole economy, and most of them will join the middle class, that equalizes the inequality between the rich and the poor in a country. There is a potential for heteroscedasticity issue, which is the result of variance of the residual depending on one independent variable is different from the variance of that variable. In order to check this and make sure that the empirical results is robust, I conduct robustness check along with my regression.

Results

To estimate the impact of FDI, as a representative for globalization, and tertiary education attainment, measured by the proportion of adult population having upper secondary degrees, I started with a pooled ordinary least squared model (Model 1). However, since this research deals with panel data, I am aware of some omitted time invariant variables that could create serious issues to the validity of OLS estimates. As a consequence, I use fixed effect model control for those characteristics and to assess the net effect of independent variables on SWIID Gini index

TABLE 1 Estimated effects of SWIID Gini Coefficient on disposable income on FDI and tertiary education attainment for OLS, fixed country effect and fixed country-year effect model specifications

Dependent variable: SWIID Gini index on disposable income					
	(1)	(2)	(3)	(4)	(5)
	Pooled OLS	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
Variables of interest					
FDI	-0.00758*** (0.00220)	0.000439 (0.000667)	0.000333 (0.000637)	0.000676 (0.000676)	0.000541 (0.000623)
Tertiary education attainment (%)	-0.0810** (0.0383)	-0.00823 (0.0245)	0.00944 (0.0235)	-0.0105 (0.0472)	-0.00859 (0.0509)
Economic growth variables					
Economic crisis (dummy)	1.019* (0.583)	-0.0201 (0.178)	0.0413 (0.159)	-0.0780 (0.600)	0.175 (0.525)
GDP growth rate	0.134* (0.0724)	0.00764 (0.0124)	0.0111 (0.0127)	0.0168 (0.0193)	0.0185 (0.0200)
Inflation rate	0.248** (0.122)	0.0158 (0.0419)	0.00952 (0.0425)	0.0320 (0.0471)	0.0326 (0.0516)
Population composition variables					
Rural Population Rate	-0.0363 (0.0272)	-0.189* (0.0989)		-0.188* (0.104)	
Total Population	0.0625*** (0.00658)	0.00967 (0.0413)		0.00840 (0.0435)	
Number of Immigrants	-4.430*** (1.276)	-0.667 (1.034)	-0.794 (1.001)	-0.653 (1.078)	-0.836 (1.002)
Rural Population			-0.314*** (0.0748)		-0.302*** (0.0705)
Constant	30.76*** (1.705)	34.43*** (3.034)	32.29*** (1.003)	34.51*** (3.163)	32.68*** (1.543)
Observations	342	342	342	342	342
R-squared	0.347	0.091	0.093	0.101	0.105
Number of countries		34	34	34	34
Country FE		YES	YES	YES	YES
Year FE				YES	YES

Note: Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

For all models, I estimate the impacts of FDI and higher education attainment while controlling for population characteristics, such as rural population, immigrants and economic development characteristics, such as GDP growth rate, Inflation rate, and Economic Crisis Dummy (1 for after and 0 for before the 2008 World economic crisis). In model 1, using OLS regression, I find that the majority of the independent variables are highly significant, at 1% to 10% level of significance. Among 34 OECD countries, one billion dollar increases in Net FDI could decrease the Gini coefficient by -0.00758 units, and 1% increase in the proportion of adults attaining tertiary education could decrease the Gini coefficient by -0.0810, holding other characteristics constant. Total population and the number of immigrants are highly significant. Interestingly, *ceteris paribus*, the more immigrants a country receives, the lower the Gini index. This estimate could be upwardly biased since there are still a lot of undocumented immigrants who did different kind of jobs and got paid low without law protection.

Moving from the pooled OLS model to other fixed effect models, I see a dramatic decrease in the significance level of most independent variables, in model 2, 3, 4, 5. This is the evidence of omitted variable bias in pooled OLS model for not taking the time invariant factors into consideration. FDI, contrary to the estimate of OLS, has positive effect on Gini Index, however, the impact is not significant. Tertiary education also seems to alleviate the income inequality, but, similar to FDI, not statistically significant even at 10% level of significance. This result is striking, since it is a common belief that education is a tool to get ones out of poverty and accumulate human capital. Previous research has found statistically significant evidence that higher education attainment among labor force contributes to income inequality, either worsened it (Park, 1996) or contributed to more equal income distribution (Gregorio&Lee,2003).

There is an interesting shift in the significant level of rural population rate and total population from model 1 to model 2 and 4. In model 2 that controls for country invariant factors and model 4 that controls for both country and time invariant factors, the proportion of people living in rural area is significant, even only at 10% level of significant. Total population, despite its high t-value of 9.51 in the OLS estimation, turns out to be not statistically significant at all. In both model 2 and model 4, an 1% increase in rural population rate leads to Gini Index decreasing by approximately 0.189 point, *ceteris paribus*.

Model 3 and model 5 are almost identical to model 2 and 4, except that I use the actual number of people living in rural area of all 34 OECD countries within the 2005-2015 timeframe, instead of the rate. The number of people living in rural area is highly significant at 1% level. According to model 3 and model 5, on average, if the population of people in rural area of these OECD countries increased by 1 million people, the SWIID Gini coefficient on disposable income will decrease by 0.314, and 0.302 point, respectively, holding other factors constant. This resembles the trend that I describe in the previous paragraph. Overall, rural population has negative impacts on the income inequality. In another words, the more people living in rural area, the lower the income inequality of an OECD country.

Another way to deal with panel data is through random effect model. However, when I ran fixed effects models, I noticed that the correlation coefficient between the residuals and the independent variables be much different from zero, so this violated the requirement for the random effects model. For example, that correlation coefficient for model 2 is -0.2896, and for model 3 is -0.8356. Fixed effects, therefore, is the most suitable model. In addition, I conducted Hausman test with the null hypothesis that the correlation between the time invariant factors and my independent variables is 0. I tried Hausman test with both types of condition, one says base both (co)variance matrices on disturbance variance estimate from efficient estimator, and the other says base both (co)variance matrices on disturbance variance estimate from consistent estimator. Both tests show the same results of $\text{Prob} > \chi^2$ equal to 0.00. So we reject the null hypothesis at 0.1% level of significance, meaning using fixed effect model is better.

In summary, using fixed effect estimations, I could not conclude whether there is an positive, negative or no impact of net FDI on income inequality. This result is different from the findings of Alderson and Nielson (2002) and Ha (2012) that FDI showed very significant influence on income inequality. In particular, Alderson and Nielson (2002) also looked at 16 OECD countries, but in a different time period, from 1960s to 1990s. They used generalized linear model to work with the panel data, and controlled for some labor market characteristics, such as female participation, and employment in manufacturing. These were likely to be the reasons why Alderson and Nielson (2002) could find a positive and highly statistically significant estimate of FDI on income inequality. However, previous research on globalization's impacts on inequality through trade and FDI showed mixed results, and depends a lot on the country characteristics, such as

geographic location or political regime (Silva, 2007; Babones and Zhang, 2008; Goldberg and Pavcnik, 2007). Similar to FDI, there is no statistically significant evidence of the influence on income inequality from the proportion of people with tertiary education degrees in an OECD country. Rural population is the only factor that has statistically significant influence on equalizing the income distribution.

Discussion and Conclusion

The results from this paper yield some important discussion on how to counter income inequality, especially among developed countries like OECD countries. According to an OECD's report called Education at a glance 2019, the proportion of tertiary-educated 25-34 year-olds increased by 9 percentage points, on average, across OECD countries between 2008 and 2018, while the share of adults with less than upper secondary education fell from 19% to 15%. On average, in 2018, 44% of the adults in OECD countries got training in tertiary education institutes. Especially for South Korea, this figure was almost 70%. With the income inequality continuing to rise in OECD countries, and the suggestion from this research data that tertiary education and inequality are not related, the push towards higher education may not be an ideal solution. It is still a good sign that higher education attainment among the workforce is not worsening the inequality like 59 countries in Park (1996). However, there needs to be better policy so that higher education attainment could be more effective in balancing the inequality, especially there are OECD countries, like the U.S.A, the U.K, and Australia, who have renowned higher education institutes.

This research has some potential limitation. The countries in this research are all from OECD countries with a more developed economy compared with most of the remaining countries all around the world. This may make the results rather exclusive for OECD countries during the 2005-2015 period, and hard to be applicable to other countries. This also explains why rural population could be so highly significant and offset the impacts of other explanatory variables on SWIID Gini Index. Another issue that could potentially harm the validity of my regression model is the way Solt (2018) calculated and imputed SWIID Gini Index so that there is no missing value. This could potentially be the reason why my regression model of income inequality did not show statistically significant results for tertiary education and globalization through international trade and investment like Castelló-Climent and Doménecha (2017) and Kishan (2017).

From this research, there is no significant evidence that globalization through FDI could lead to changes in within countries income inequality. In this research, I use net FDI that is equal to net FDI outflow (asset) minus net FDI inflow (liabilities). Another factor that also relates to globalization I included in this research, in terms of the transfer of human labor, is immigrants. Even though the number of immigrants was not statistically significant in fixed effect models, it is hard to conclude that globalization has no impacts on income inequality. Future research for OECD countries should look at other proxy for globalization, such as net export, and trade with lower or more developed economy to see whether the income inequality within that country could be changed.

An interesting finding from this research is the impact of rural population on the whole country's income inequality, meaning more people living in rural area becomes an equalizer for income distribution. This contradicts to the common belief that rural area stays stagnant behind urban area and worsens the whole country's income inequality. According to the U.S Census Bureau, 2011-2015, the poverty rate among people living in rural area is so much lower than urban area. The result from this research also suggests that the income inequality, as depicted from SWIID Gini index, is caused by the widening gap between rural and urban area. Policy makers should consider this finding as a solution to alleviate income inequality and promote better living quality for everyone, so that people will see rural area as a good place to live. The efforts to accelerate the urbanization process may not yield the good outcome, especially for income distribution. Future research should also include rural population when measuring factors that affect the within countries inequality, especially with the countries that are being urbanized at a high rate.

Reference

- Alderson, A.S., Nielson, J. (2002). Globalization and the Great U-Turn: Income Inequality Trends in 16 OECD Countries. *American Journal of Sociology*, 107(5), 1244-1299
- Barbones, S., Zhang, X. (2008) HETEROGENEITY IN THE GLOBAL SOUTH IN THE RELATIONSHIP BETWEEN INCOME INEQUALITY AND FOREIGN TRADE, *International Review of Modern Sociology*, 34(1): 95-108
- Behrman, J.R., Schott, W., Mani, S., Crookston, B.T., Dearden K, Duc, L.T., Fernald, L.C.H., Stein, A.D. (2017). Intergenerational Transmission of Poverty and Inequality: Parental Resources and Schooling Attainment and Children's Human Capital in Ethiopia, India, Peru, and Vietnam. *Economic Development Culture Change*, 65(4), 657–697
- Barro, R.J. (2002). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5, 5–32
- Castelló-Climenta, A. and Doménech, R.(2014). Human Capital and Income Inequality: Some Facts and Some Puzzles
- Dutt, P., Mitra. D. (2007). Inequality and the Instability of Polity and Policy. Retrieved at <https://ssrn.com/abstract=748704>
- Goldberg, P.K., Pavcnik, N.(2007). Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature*, 45(1), 39-82
- Green, A. (1997). *Education, Globalization and the Nation State*
- Ha, E. (2012). Globalization, Government Ideology, and Income Inequality in Developing Countries. *The Journal of Politics*, 74 (2) ,541-557
- International Monetary Fund (IMF). (2015). *Causes and Consequences of Income Inequality: A Global Perspective*
- Ismail, M. (7/17/2018). Southeast Asia's widening inequalities. Retrieved from <https://theaseanpost.com/article/southeast-asias-widening-inequalities>
- Kishan, P. K. V. (2017). Globalization and inequality - A pathway through education. Retrieved from <http://hdl.handle.net/11718/20215>
- Malderen, C.V, Oyen, H.V. and Speybroeck, N.(2013). Contributing determinants of overall and wealth-related inequality in under-5 mortality in 13 African countries. *Journal of Epidemiology and Community Health*, 67(8): 667-676
- Meschi, E., Vivarelli, M.(2009). Trade and Income Inequality in Developing Countries. *World Development*, 37(2): 287-302

- Mills, M. (2009). Globalization and Inequality. *European Sociological Review*, 25(1), 1-8
- Ostry, J.D., Berg, A., and Tsangarides, C.G. (2014). *Redistribution, Inequality, and Growth*
- Park, K.H. (1996). Educational expansion and educational inequality on income distribution
Economics of Education Review, 15 (1): 51-58
- Sala-i-Martin, X. (2006). The World Distribution of Income: Falling Poverty and ...Convergence
- Stiglitz, J. E. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future*. United States: W. W. Norton.
- Silva, J.A. (2007). Trade and Income Inequality in a Less Developed Country: The Case of Mozambique. *Economic Geography*, 83(2): 111-136
- Yeung, H.W. (2009). The Limits to Globalization Theory: A Geographic Perspective on Global Economic Change. *Economic Geography*, 78(3): 285-305

Data Appendix

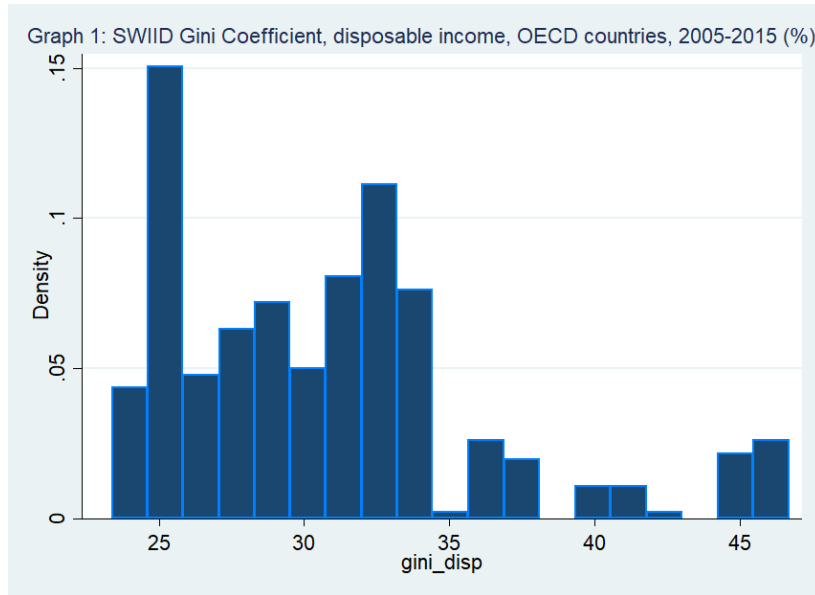
This appendix provides further description of our data sources and some manipulations to the data set.

Gini coefficient on disposable income (data label: gini_disp). Standardized World Income Inequality Database (SWIID) is the most comprehensive data set on income inequality, with a wide coverage of 196 countries from 1960 to the present for as many years as possible (Solt, 2019). The SWIID Gini coefficient ranges from 0 to 1, with 0 meaning that people have exact the same amount of money, and 1 meaning perfect inequality. Solt (2019) reports different types of Gini coefficient, including disposable income, market income, absolute distribution and relative redistribution. In this research, I choose the Gini coefficient of the disposable income, which is post-tax, post-transfer income, because it better represents the actual money and the consumption power for a household. The dataset I extracted the data from is SWIID Version 8.2, November 2019. Solt (2019) calculated Gini coefficient based on the collected data from different sources, including the United Nations University's World Income Inequality Database, the OECD Income Distribution Database, the Socio-Economic Database for Latin America and the Caribbean generated by CEDLAS and the World Bank, Eurostat, the World Bank's PovcalNet, the UN Economic Commission for Latin America and the Caribbean, the World Top Incomes Database, national statistical offices around the world, and many other sources. After that, the author applied multiple imputation estimation techniques to deal with missing data, which raised concerned among researchers about the validity of the data set.

Table1: Summary Statistics for SWIID Gini Coefficient, disposable income,

OECD countries, 2005-2015 (unit: %)

Variable	Obs.	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
gini_disp	374	30.76	5.43	23.4	26.1	30.2	33	46.7

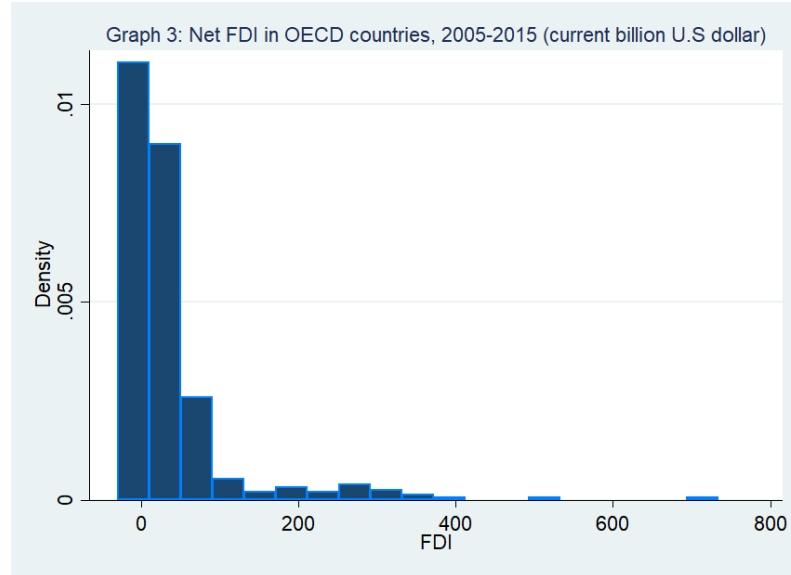


Overall, there is no missing value for SWIID Gini index, so there are 374 observations of 34 countries in 11 years. The distribution of SWIID Gini Coefficient of disposable income in OECD countries is positively skewed. The graph suggests the existence of outliers. During this 11-year period, the income inequality of OECD countries stayed between 23.4% and 46.7%.

Net foreign direct investment (data label: FDI). Foreign direct investment in this research is collected by World Bank from different sources, including International Monetary Fund, Balance of Payments Statistics Yearbook and data files. Net FDI is equal to net FDI outflow (asset) minus net FDI inflow (liabilities). The data is in current billions of US dollar.

Table2: Summary Statistics for FDI in OECD countries, 2005-2015, (unit: billion of US dollars)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
FDI	374	3.87	7.80	-2.97	3.77	13.48	38.338	734

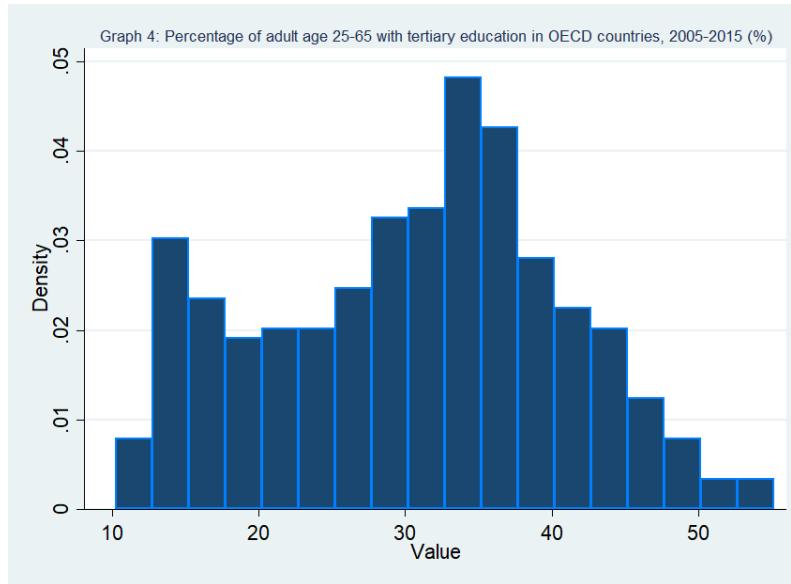


Net FDI in OECD countries during 2005-2015 was very positively skewed, with the gap between -2.97 billion dollar and 7.34 billion dollar. Interestingly, there are some countries with negative FDI. However, there are not a particular trend in terms of countries or years among these countries. The country with the highest net FDI is Netherland in 2007. FDI has 34 outliers and they all look extreme.

Tertiary Education (data label: *terEduc*). Data on tertiary education of OECD countries are collected from their website data.oecd.org, and the primary source is Education at a glance: Educational attainment and labour-force status. In this research, I use the highest level of education completed by the 25-64 year-old population as a measurement for tertiary education impact on income inequality. The data is measure in the scale of 100, and its unit is percentage. There are 17 missing values in this variable I collected from OECD's website.

Table 3: Summary Statistics for Proportion of adult 25-65 years old with tertiary education in OECD countries, 2005-2015, (%)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
FDI	357	30.605	10.077	10.239	22.637	32.041	37.584	55.174

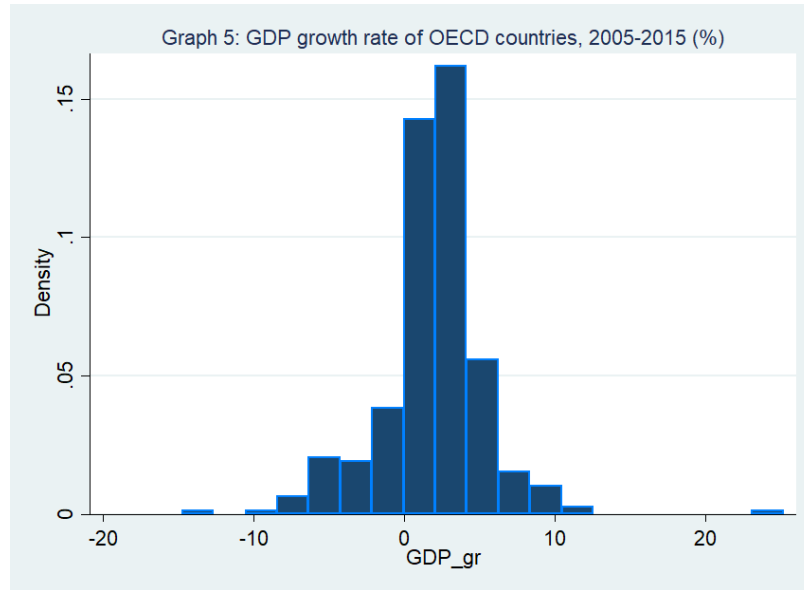


Tertiary education is an interesting variable for its normal distribution and having no outliers. During these 11 years, there are countries that only have very low amount of its population got a college degree, of only 10.077% while there are countries that had more than half (55.17%) of its people secure a college degree. With an urge towards normalizing higher education, this data is actually not very impressive, however, since this has people who are older generation at the age of 60s, their data could pull the average down because the policy did not take place during their time.

GDP growth rate (data label: GDP_gr). This data is downloaded directly from the data website of World Bank. The note there showed that they collected data from their own source, which is World Bank national accounts data, and OECD National Accounts data files. GDP growth rate I use in this research is the rate at which GDP at market prices based on constant local currency grows annually. It is measured in the 100 scales and the unit is percentage.

Table 4: Summary Statistics for GDP growth rate in OECD countries, 2005-2015, (%)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
GDP_gr	374	1.963	3.434	-14.72	.691	2.174	3.658	25.162

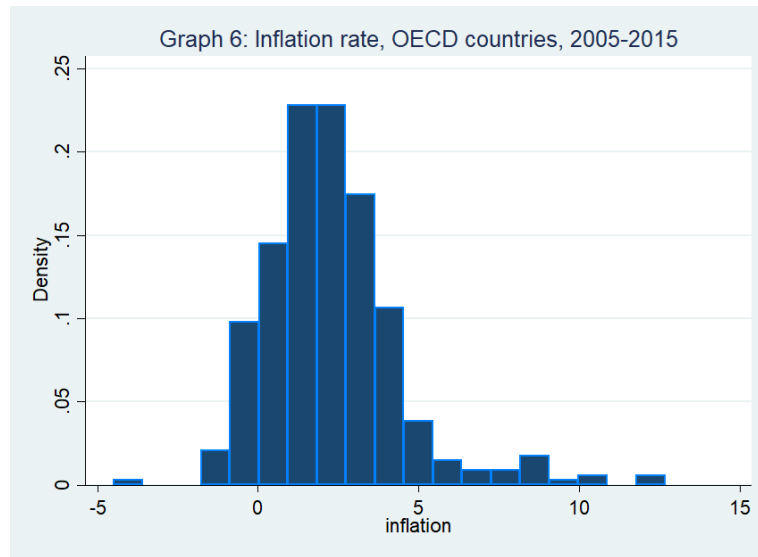


GDP growth is distributed in a normal scale during this period. There were countries that have their GDP growth be negative. My investigation into the data show that all negative GDP growth is during the after crisis period. Ireland is the country that had experience the highest growth in GDP in 2015, when their figure was 25.162%. GDP growth data has no missing values and 37 outliers to both negative and positive signs.

Inflation (data label: inflation). According to World Bank, where I downloaded the data set for inflation rate, there are different concepts of inflation rate. However, in this research, I chose inflation rate with a relation to consumer price. The definition by World Bank is “the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly.”

Table 5: Summary Statistics for inflation in OECD countries, 2005-2015

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
inflation	374	2.326	2.117	-4.478	1.0193	2.129	3.341	12.694

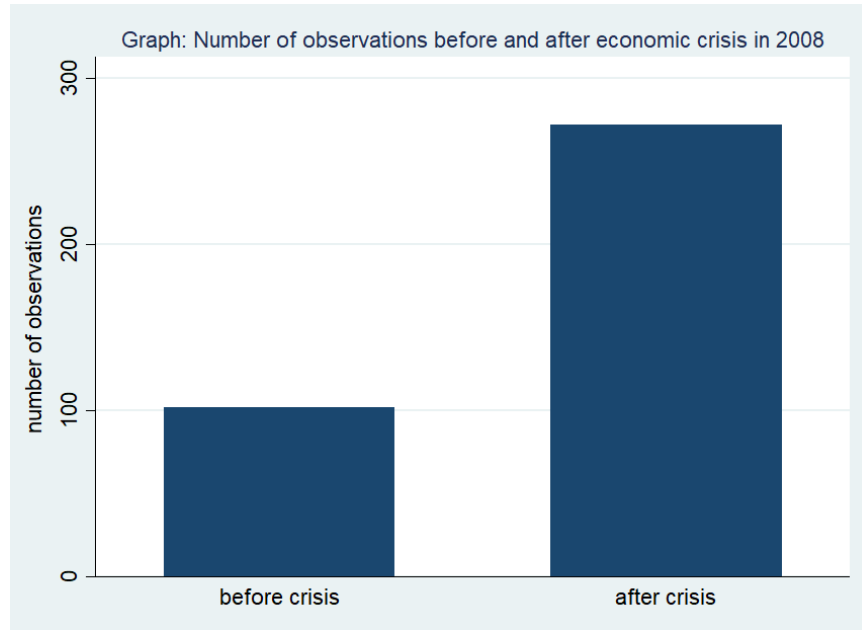


In general, the inflation rate of OECD during this time is at acceptable rate around 2%. Their median is 2.129%. However, there are still some countries during some years that had extremely high inflation rate. Using the 1st quartile of 1.0193% and 3rd quartile of 3.341%, I calculated the interquartile and upper level of the box plot, which was 6.82%. Using STATA, I found out that Turkey is the country that had the most inflation data that exceed this 6.82% benchmark. The inflation rate of Turkey is outliers of this data set for almost 11 years, except in 2009 and 2011.

Economic crisis (data label: crisis). I generated a new dummy variable, called crisis, since I think whether a nation's economy and the whole world's economy are in crisis contributes a lot on income inequality. Years from 2005-2007 are given the value of 0, which means they are before the crisis in 2008. The remaining years from 2008 to 2015 are 1.

Table 6: Summary statistics for crisis dummy variable, 2005-2015

Crisis	Frequency	Percentage	Cumulative percentage
0 (Before crisis)	102	27.27	27.27
1 (After crisis)	272	72.73	100
Total	374	100	



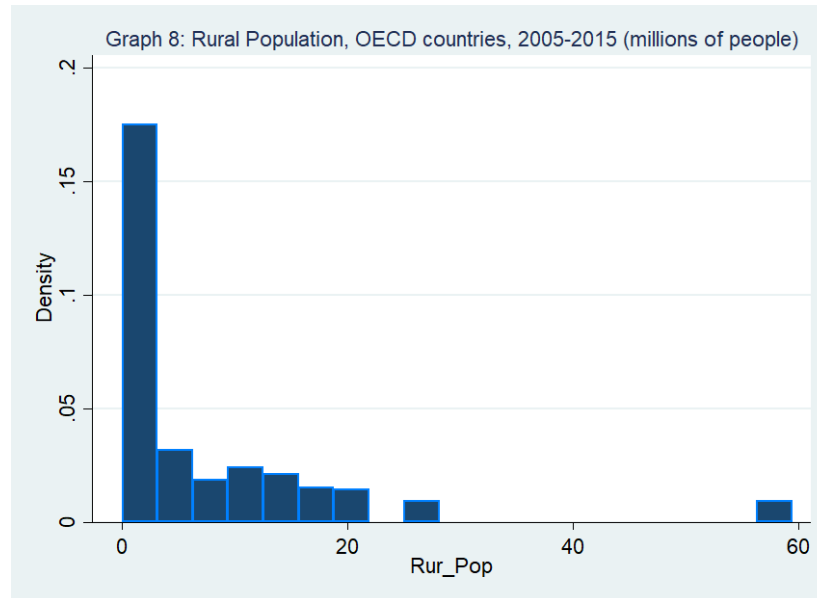
From the frequency table, we could see that there is no missing value, since all countries in OECD had open economy, which means they are susceptible to the 2008 recession, even to different levels. The time frame I chose made it quite disproportionate, meaning 3 years before and 7 years after the crisis, when it came to measure the impact of crisis on income inequality. However, this is still a relevant variable to be included in the regression.

Rural Population (data label: Rur_Pop). Rural population is simply the number of people living in rural area as defined by different countries' statistic offices. The data is from World Bank website, so it is very continuous for all OECD countries, without any missing values. Even though there is no universal definition of rural and urban, World Bank proposed one way to distinguish in the 2005 report, which is fundamentally base on the low or high density of the population, remoteness level and how active the market is. When I collected the data from World Bank's original data source, it is head counts. To better interpret and represent the data, I change the scale of the data by dividing it by 1,000,000. The units I use for all population related data in this research are in millions of people, likewise.

Table 7: Summary Statistics for rural population in OECD countries, 2005-2015

(million of people)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
Rur_Pop	374	7.635	11.361	.0204	.967	2.751	11.394	59.495



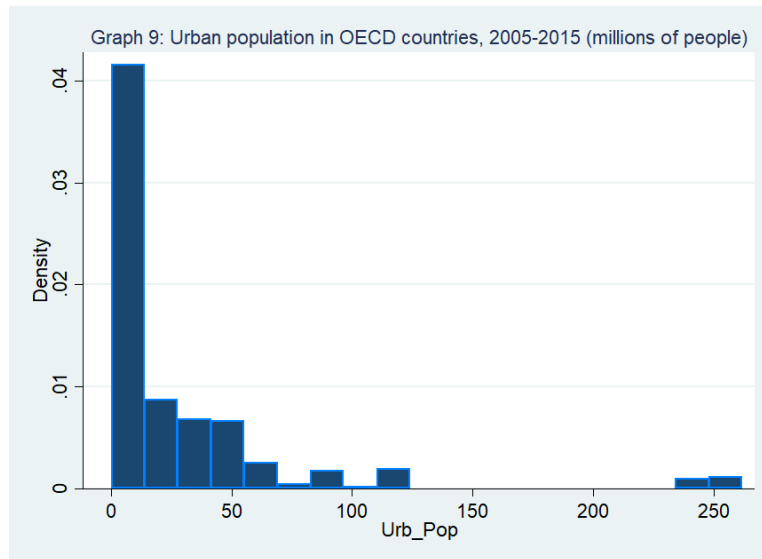
Graph 8 showed that rural population was positively skewed with 11 outliers. The majority of the countries have very low rural population of less than 5 million people..

Urban Population (data label: Urb_Pop). Similar to rural population, I also extract data on urban population from World Bank. According to its website, “urban population refers to people living in urban areas as defined by national statistical offices. It is calculated using World Bank population estimates and urban ratios from the United Nations World Urbanization Prospects”. I also conducted scale change, similar to rural population, so my unit of measurement for urban population is millions of people. There is also no missing value.

Table 8: Summary Statistics for urban population in OECD countries, 2005-2015

(millions of people)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
Urb_Pop	374	28.699	46.838	.0204	4.493	8.279	39.740	261.95



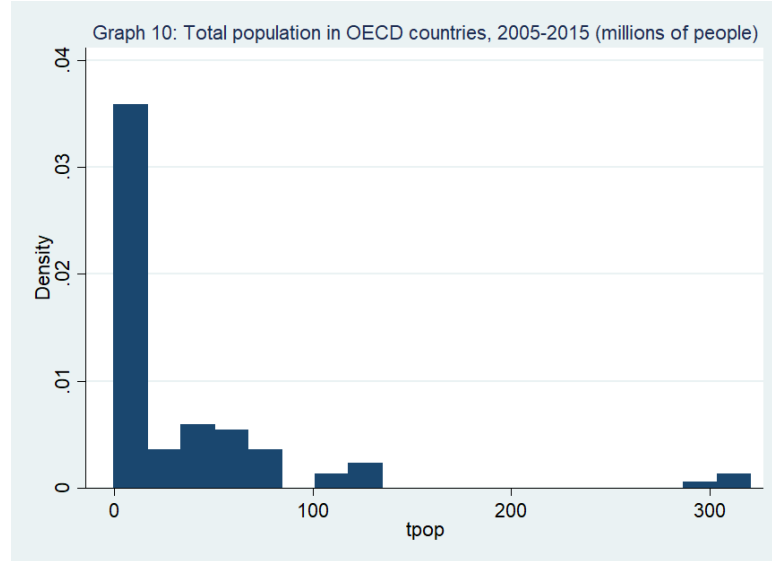
Similar to Graph 9 and the trend of rural population, urban population's skewness is positive. However, as can be seen from the graphs, urban population is so much higher than the rural population, which could be justifiable. All 34 countries in OECD could be classified as a more advanced economy with the urbanization starting much earlier.

Total Population (data label: *tpop*). The data for total population was from World Bank public data set. The data description in their website shows different sources that World Bank collected data and compiled this data set, including United Nations Population Division. World Population Prospects: 2019 Revision., Census reports and other statistical publications from national statistical offices, Eurostat: Demographic Statistics, United Nations Statistical Division. Population and Vital Statistics Reprot, U.S. Census Bureau: International Database, and Secretariat of the Pacific Community: Statistics and Demography Programme. The data set is very complete for the time frame that I chose with 374 observations for 34 countries in 11 years. The reasons why I did not use STATA to calculate total population but downloaded from World Bank is because of the concerns about how rural and urban are defined could lead to some calculation errors. "Aggregation of urban and rural population may not add up to total population because of different country coverages." (World Bank report, 2019).

Table 9: Summary Statistics for total population in OECD countries, 2005-2015

(millions of people)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
totpop	374	36.335	57.647	.297	5.414	10.563	49.554	320.74



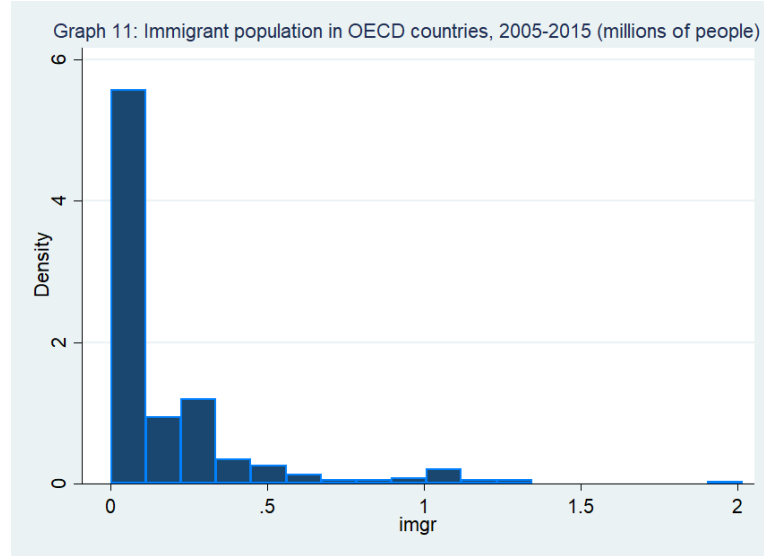
Similar trend is shown in both total population, rural population and urban population. The gap between maximum population of 320.74 millions of people and minimum population of only 0.3 millions is extremely large. There are 26 outliers for this variable.

Immigrant population (data label: imgr). Immigrant population is measured based inflow of foreign population. The data I used in this research is collected from International Migration Database publicized in stats.oecd.org. The way they measured this inflow is based on population registers (more permanent) and also work permits (more temporary). However, this data only focuses on legal immigration, since there are also illegal and undocumented immigrants, especially in advanced economic countries like OECD. Due to the difficulties in counting and keeping track of immigrants, there are 15 missing observations. The countries with missing values are Slovenia, Greece and Turkey, especially in the case of Turkey, there is only one available data point of year 2010.

Table 10: Summary Statistics for immigrant population in OECD countries, 2005-2015

(millions of people)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
imgr	359	.174	.261	.00098	.024	.066	.245	2.016

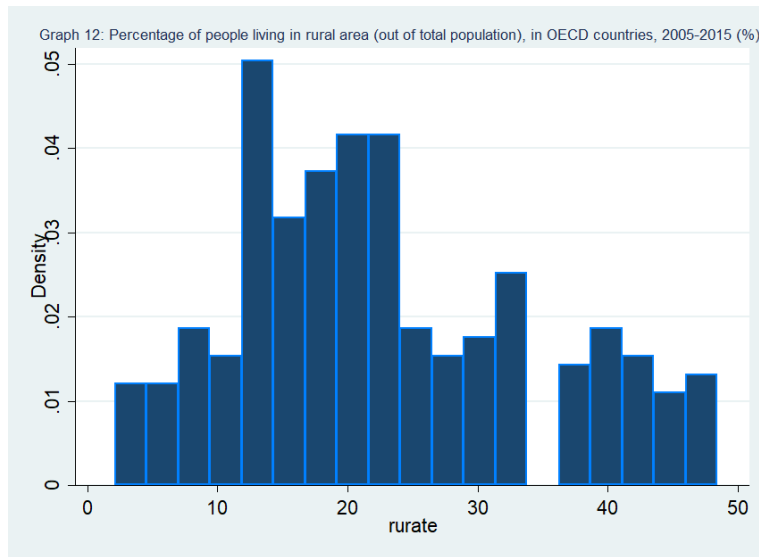


In general, comparing to the total population, the number of immigrants in OECD countries is not substantial. The country with the highest immigrant population in this period is Germany in 2015 with 2.016 millions of people. This is far higher than the rest of OECD countries, and also compared to Germany in the previous years.

Proportion of people living in rural area (data label: rurate). I generate this variable using the data collected from World Bank on total population and rural population. This is the proportion of people living in rural area out of total population. The data is calculated in the scale of 100, with the unit of percentage. There is no missing value in this variable.

Table 11: Summary Statistics for immigrant population in OECD countries, 2005-2015 (%)

Variable	Obs	Mean	Std. dev.	Min	1 st quartile	Median	3 rd quartile	Max
rurate	374	22.78	11.381	2.12	13.9	20.866	31.021	48.467



The histogram shows that the percentage of rural area inhabitants is normally distributed over this 11 year period. There is no outliers in this variable. 75% of the countries in this time frame has less than 30% of their population living in rural area.