School of Business and Administration

Department of Economics

ECO 351 - Econometrics

Determinants of Income Inequality in South Korea from 1990-2014

By:

Al-Baraa El-Hag 47451

Under the supervision of:

Dr. Oz Kaya

Table of Contents

Table of Figures & Tables

1. Literature Review

Income inequality refers to the disparity between the income distribution of a population. It is typically measured using the Gini Coefficient (after the Italian statistician and sociologist Corrado Gini). The coefficient takes values from 0 to 1, with one representing maximum inequality in a society (in which one person has all the wealth) and with zero standing for perfect equality (in which everyone has the same wealth). Today, most countries range from 0.25 to 0.7. There is no agreed upon ideal value of a Gini coefficient, mostly due to differing political opinions in the countries and the fact that the ratio is only an approximation (although it is the most widely used in research, even fields such as biodiversity and chemistry utilize it).

Kuznets (1955) tried to study the relationship between income inequality and economic growth. His theory was that income inequality will rise with initial periods of economic growth until reaching a peak and falling after that, creating an "inverted U-shape" that has come to be known as the Kuznet's curve. According to Fields and Yoo (2000), the economy of South Korea is an example of this. Due to reforms done by the government starting in the 1960s, income inequality initially rose rapidly but it then reached a peak and started to decrease; by their calculations it fell by 27% from between the years of 1976 to 1993.

However, the index has started to rise again, going from 0.254 in 1991 to a peak of 0.32 in 2009, a 25.9 % increase. Although still a debated issue, many studies have gone so far as trying to correlate population health and income inequality. Torre and Myrsyla (2011) use panel data to show that there is a highly significant positive relationship between male mortality and the Gini index at the ages of 15-49, with the p-value<0.01, although conversely they found that the female coefficient at the same age group was 40% smaller and only significant at a p-value<0.10. In South Korea, it was shown empirically by Cho, Park and

Kang (as cited in Lee, et al., 2013), that the rising Gini index in Korea was positively related to the rate of crime and suicide.

An important question in the research surrounding income inequality is what are its determinants. There are many factors that may have brought on the rapid rise of income inequality in South Korea, with this paper studying the period of 1990-2014. I will try to reconstruct Lee, Kim and Cin's (2013) paper on the determinants where they, by using an OLS model with income inequality as the determinant variable were able to reach an adjusted R^2 of 0.963. A final thing to be noted is that one of the factors that may have accelerated it is the policy change's that were meant to curtail the 1997 financial crisis, but trying to capture the effects during a crash is very difficult.

2. Economic Theory & Data

This section will introduce the eleven variables used for the regression and give a background of the economic theory surrounding them. All variables were in the window of time from 1990-2014.

2.1 Gini Index

As discussed in the literature review, the Gini coefficient will serve as my dependent variable. The source of the index data comes from KOSIS (Korean Statistical Information Service), specifically described as the market income value for urban households (2 and more non-farm). While the authors are using the series from 1980 to 2012, I could only find the data starting 1990 to 2014. On trying to look for alternatives I found the IMF carrying similar data and the World Bank only carrying the Gini coefficient for 1998.

$$Y = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 Z_t + e_t$$

Y = Gini Index

$$X_t = \ln \ln \left(\frac{GDP}{Population} \right)$$

Z = Vector of maro-economic variables

Above you can see the formulae that will be used to model income inequality. The reason we are including the second degree of Xt is to see if Kuznets' curve is valid. If his theory is about income inequality and growth than the U-shaped graph will cause the $\beta_1>0$ and $\beta_2<0$. In contradiction to Kuzents; it is argued by Barro (2000) that during the initial periods of growth the Gini index decreases until the economy reaches a level in which income inequality will increase. If Barro's theory is true than we should instead see that $\beta_1<0$ and $\beta_2>0$.

2.2 GDP per capita

My first set of independent variables will be log(GDP per Capita) and log(GDP per Capita) ^2. The source was the World Bank's "GDP (current LCU)".

2.3 Consumer Price Index

The CPI growth rate was used (inflation) and also sourced from the World Bank. The theory behind it is that there should be a negative relationship between it and the dependent variable. With the poor and middle class losing relative more then the richer class who are more insulated to price gains.

2.4 Unemployment Rate

The unemployment rate, defined as the numbers unemployed people divided by the labour force and it was sourced from the World Bank. Initially, data from KOSIS was used but it only started at 2000. The intuition would suggest that there also should be a negative relationship between unemployment and the Gini index, since unemployment usually affects the lower classes first.

2.5 Elderly Population Percentage

This is defined as the percentage of the population over 65 divided by the working age population and it is sourced from the World Bank. Many countries are facing an increasing number of elderly, due to low birth rates and longer life expectancy. The intuition would suggest that there should be a positive relationship with aging and the Gini coefficient.

2.6 Foreign Direct Investments

As defined by the World Bank "Foreign Direct Investments are the net inflows of investment to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. FDI is added here as a measure of globalization and there are arguments for the relationship to be positive or for it to be

negative. The argument for the former is that FDI would raise income by creating employment opportunities and advancing technology, for the latter it is argued that FDI's would be hampering with the receiving countries independence.

2.7 Female Percentage in Labor Force

Defined as the percentage of females participating in the labor force and it is sourced from the World Bank. If the coefficient turns out to be statistically significant I would estimate for it to be negative, since in many areas what is exacerbating income inequality is the unequal participation of females.

2.8 Trade Openness and Share of Imports

These are also a measure of globalization and are calculated as follows:

Trade Openness = (Exports + Imports)/GNI
Share of Imports = Imports/GNI

And it is sourced from the World Bank. The same argument could be made for their effects as the FDI's.

2.9 Share of Investments and Government Spending

This is calculated as follows:

Share of Investments = Investment/GDP
Government Spending = Government spending/GNI
I couldn't find the investments of Korea so I took national savings as a substitute. The intuition would suggest that the coefficient should be negative since through investments and government spending you are usually aiming towards the lower class.

2.10 Variables Not Used

There were three variables that could not be found, they are:

1. Share of middle school students

2. Growth rate of Agricultural Product

3. Share of self-employed

Table 1: Variable Definition

Variable	Description	Definition
X	Nominal GDP per capita (in Won)	ln(GDP/pop)
cpi	Yearly consumer price growth rate	% ΔCPI
old	Share of elderly population	Pop over 65 / working-age pop
Unemp	Unemployment rate	Unemployed / working-age pop
empf	Female employment rate	Female employment / working-age pop
log(FDI)	Foreign direct investment	ln(FDI)
xmy	Trade openness	(Export + Import) / GNI
iy	Share of investment	Investment / GDP
gy	Share of government spending	Government spending / GNI
my	Share of import	Import / GNI

3. Empirical Results

Table 2: Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
X	-0.377	-6.258	3.486*	-7.136***	-0.147	-7.040**	-0.783
	(0.885)	(5.641)	(2.007)	(2.306)	(1.875)	(2.432)	(1.498)
X^2	0.0133	0.195	-0.106	0.222***	0.00523	0.219**	0.0251
	(0.0266)	(0.175)	(0.0622)	(0.0714)	(0.0577)	(0.0754)	(0.0462)
CPI		0.00175	0.00118	-0.000225	0.000283	-0.000195	0.000569
		(0.00113)	(0.00127)	(0.00109)	(0.00135)	(0.00108)	(0.00133)
old		-0.0174	0.00869	-0.0291***	-0.00519	-0.0286***	-0.00695
		(0.0187)	(0.00764)	(0.00722)	(0.00593)	(0.00772)	(0.00446)
Unemp		0.00887***		0.00704***		0.00720***	
		(0.00299)		(0.00233)		(0.00221)	
empf			-0.00279***		-0.00160		-0.00193**
			(0.000807)		(0.00100)		(0.000869)
log(FDI)				8.42e-05	0.00538	0.000114	0.00564
				(0.00304)	(0.00359)	(0.00302)	(0.00335)
xmy				0.00866	-0.0128		
				(0.0141)	(0.0263)		
iy				-0.0992	-0.139	-0.0983	-0.114
				(0.101)	(0.142)	(0.102)	(0.138)
gy				1.504***	1.192***	1.495***	1.137***
				(0.267)	(0.281)	(0.270)	(0.272)
my						0.0155	-0.0539
						(0.0275)	(0.0488)
Constant	2.883	50.55	-28.35*	57.61***	1.156	56.83**	6.291
	(7.374)	(45.69)	(16.24)	(18.68)	(15.27)	(19.69)	(12.19)
R	0.811	0.912	0.904	0.972	0.959	0.972	0.961
Adjusted R2	0.794	0.889	0.879	0.956	0.935	0.955	0.937

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, *p<0.1

Table 2 on the previous page are my results for the regression; on comparing it with the papers I noticed that it was almost identical if not for some minor differences (see Appendix A). First of all, the highest adjusted R^2 for my regression was Model #6 explaining 95.5% of my dependent variable, they got their highest adjusted R^2 also at Model #6 albeit at a slightly higher value of 96.4% which could be explained from the addition of the variables representing growth rate of agriculture product and share of middle school students which were not added in my regression.

Second difference is that our X and X^2 are significant for three of the models, compared to the papers where it was only significant (at the 10% level also) for the first model. This could be because of the different sampling windows of the data taken. For them they have the period from 1980 – 2012, which included a down then up then down trend for X ending up cubic. But for the period of 1990 – 2014, it was mostly quadratic (Figure 1 shows a plot of Time vs Gini). Perhaps if they took an X^3 factor they would have had lower standard errors.

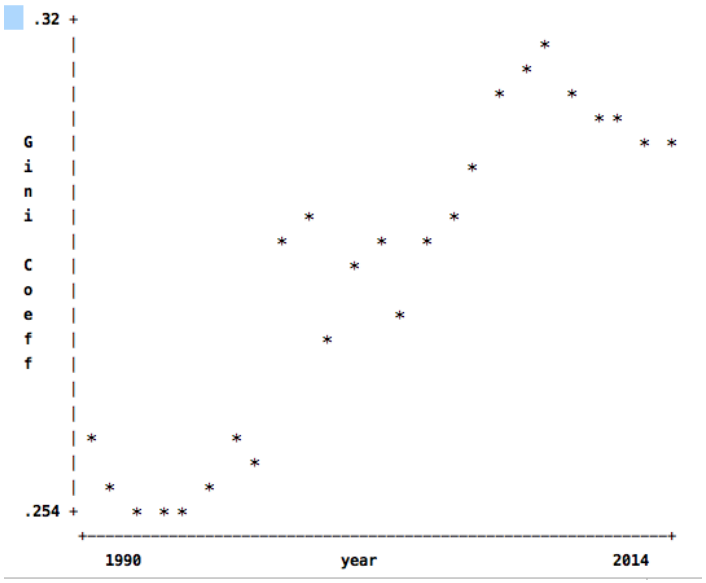

Figure 1: Time vs Gini

On the implications of the signs of co-efficients of GDP ($\beta_1<0$ and $\beta_2>0$) we see that the data proves Barro's theory over Kuznets'. Although, it must be noted that the sample time for my regression is for only 24 years and across only one country.

The coefficient of old was significant at the 10% level for two out of five of the models with a negative value. This means that with a larger number of elderly people, the income inequality goes down. This is opposite to that of intuition, but perhaps it could be that as the Korean population ages they like to spend their money instead of save it. Thereby transferring it to the younger generation who are usually in a lower class than the older generation.

Unemployment was significant in all models, with a positive coefficient, as predicted. The results seem to suggest that during higher levels of unemployment income inequality rises. Also significant in two models out of three was the female participation rate, with a negative coefficient, also as predicted.

Government Spending (gy) was also significant in all models at the 10% level. But in all four cases it came as a positive coefficient, meaning that with more government spending there is a higher level of income inequality. The research paper found the gy was insignificant at all levels. The discrepancy could be due to the fact that during recessions/depression governments have to take drastic actions to counter-act them and bringing the country back to stability, thus they increase their spending. During this time, the lower classes suffer the most and thus inequality rises. Therefore it might be a correlation but not a causation per say.

Finally, we find that the following variables were not significant at any level in any model: Inflation, FDI, XMY (Trade Openness), IY (Share of Investment) and MY (Share of Import).

This somewhat contradicts the earlier paper since at some point over the models they are significant, this could be due to a number of reasons; different windows of time, source of data differs, the missing variables not included.

4. Conclusion

In conclusion, we see that we were able to replicate the (Lee, et. all, 2013) results with minor differences. We found that GDP, age of population, unemployment rate, female employment rate and share of government spending are all significant variables for the dependent variable, income inequality. While inflation, foreign direct investment, trade openness, share of investment and share of import were not significant variables. Recommendations for future research would be to extend the window of time to the 1960's using experimental methodology to estimate the Gini coefficient. With more data points some variables will probably become significant and we could further test Kuznets' theory.

References

G. Fields and G. Yoo, “Falling Labor Income Inequality in Korea's Economic Growth: Patterns and Underlying Causes”, Review of Income and Wealth, vol. 46, (2000), pp. 139-159.

Lee, H. Y., Kim, J., & Cin, B. C. (2013). Empirical Analysis on the Determinants of Income Inequality in Korea. International Journal of Advanced Science and Technology, 53, 95-109.

R. J. Barro, “Inequality and Growth in a Panel of Countries”, Journal of Economic Growth, vol. 5, (2000), pp. 5-32.

S. Kuznets, “Economic Growth and Income Inequality”, American Economic Review, vol. 45, (1955), pp. 1-28.

Torre, R., & Myrskylä, M. (2011). Income inequality and population health: a panel data analysis on 21 developed countries (No. WP-2011-006). Max Planck Institute for Demographic Research, Rostock, Germany.

The World Bank, World Development Indicators (2014). *Korea, Rep.* [Data file]. Retrieved from http://data.worldbank.org/country/korea-republic

Korean Statistical Information Service (2015). *The Index of Income Distribution.* [Data file]. Retrieved from http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1L6E001&conn_path=I2&language=en

Appendix A

Table 3. Determinants of Income Inequality by OLS

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
X	-0.181***	0.003	0.071	-0.032	-0.014	-0.003	0.014	0.046
	(0.028)	(0.036)	(0.064)	(0.036)	(0.034)	(0.035)	(0.033)	(0.063)
X^2	0.015***	-0.003	-0.007	0.001	-0.001	-0.002	-0.003	-0.005
	(0.002)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
d97	-0.026**	-0.007	-0.007	-0.014***	-0.012**	-0.014***	-0.013***	-0.014***
	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ΔCPI		-0.093***	-0.046	-0.071**	-0.074**	-0.062**	-0.066**	-0.053*
		(0.025)	(0.031)	(0.029)	(0.031)	(0.028)	(0.028)	(0.031)
OLD		0.014***	0.015***	0.008**	0.010***	0.010***	0.012***	0.010***
		(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
MidST		-0.270***	-0.239***	-0.228**	-0.230**	-0.225**	-0.226**	-0.196*
		(0.084)	(0.088)	(0.101)	(0.107)	(0.096)	(0.099)	(0.104)
AggPr		-0.015	-0.019	-0.021*	-0.018	-0.020*	-0.019	-0.027*
		(0.015)	(0.014)	(0.012)	(0.013)	(0.012)	(0.012)	(0.014)
Unemp		0.268***		-0.190		-0.156		-0.181
		(0.097)		(0.151)		(0.148)		(0.166)
Emp_f			-0.190**		0.043		0.061	
			(0.085)		(0.115)		(0.112)	
SelfEmp			30.909**					12.311
			(13.292)					(15.171)
ln(FDI)				0.004	0.002	0.004	0.002	0.003
				(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
XMY				0.030*	0.030*			
				(0.016)	(0.017)			
IY				-0.079*	-0.062	-0.099**	-0.091*	-0.093
				(0.047)	(0.049)	(0.046)	(0.049)	(0.060)
GY				0.111	-0.008	-0.036	-0.153	0.036
				(0.242)	(0.231)	(0.259)	(0.244)	(0.291)
MY						0.061**	0.067**	0.070**
						(0.029)	(0.030)	(0.034)
$\bar{R}^2$	0.729	0.942	0.957	0.963	0.960	0.964	0.963	0.963
DW	0.37	1.27	1.55	2.13	1.93	2.14	2.05	2.25
DF-GLS	-2.91*	-3.12**	-4.00***	-4.27***	-4.62***	-4.25***	-4.52***	-4.27***

Standard errors in parentheses

***, **, *: 1%, 5%, and 10% statistical significance respectively