

The Correlation Between Democratic Freedom and Income Disparity Across OECD Member States

Siyuan Chai¹

¹Department of Economics, School of Political Science and Economics, Waseda University, Tokyo, 163-8001, Japan.

Abstract

In this paper, I will investigate and estimate the relationship between the level of democracy (how democratically free a country's governmental and social systems are) and income/wealth distribution inequality across countries.

Keywords

Income Disparity; Democratic Freedom; OECD Member States.

1. Paper Theme Overview and General Research Background

For this term paper, I will investigate and estimate the relationship between the level of democracy (how democratically free a country's governmental and social systems are) and income/wealth distribution inequality across countries. The independent variable (x-axis) in my statistical representation is the level of democratic freedoms, and the dependent variable (y-axis) is the income inequality. In order to illustrate quantitatively the value of the two variables, Freedom House index is used for representing countries' democratic freedoms level, and the widely applied economic measure, Gini coefficient from income distribution, is used for demonstrating the degree of income inequality for each society. Data set will be collected for the current 36 OECD (Organization for Economic Co-operation and Development) member states as research subjects and a correlation coefficient, a number between -1 and 1, between the two variables will be derived from a simple linear regression analysis, which involves a linear function. (Linear regression rather than multiple regression as for two variables, a nonlinear function)

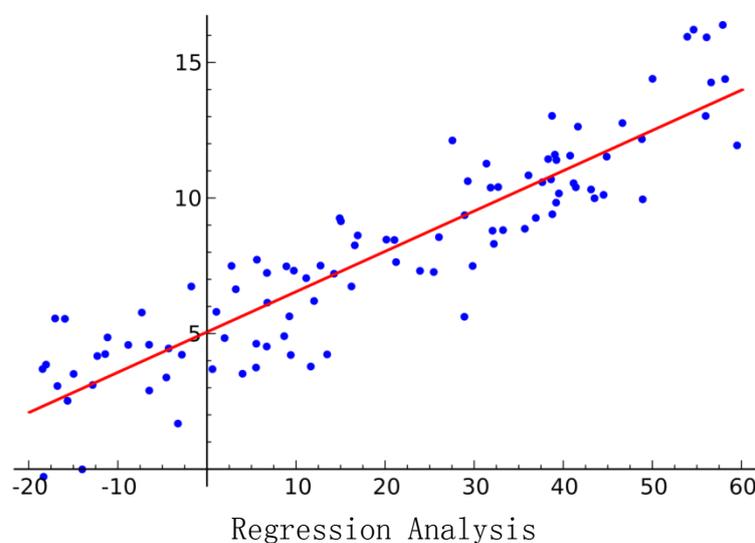


Figure 1. Illustration of scatter plot and linear regression on a data set

My presumptive hypothesis is that there will be a negative correlation pattern between democratic freedoms level and income distribution disparity. (A positive correlation pattern between democratic freedoms level and income distribution equality) In other words, as countries become more democratically free, the income of its residents is distributed more equally and evenly. However, the strength of this negative correlation represented by the value of the correlation coefficient will be obtained through a regression analysis, and the its statistical significance, goodness of fit and standard deviations, etc will be shown through a summary outputs table indicating the relatibility/genuineness/credibility of the linear regression analysis model without complex and detailed statistical explanations.[1]

In the term paper, my presumptive hypothesis will be tested and possible explanations will be summarized based on statistical analysis procedures. Regression graph/curve in the form of a scatter plot created by Excel software package will be included in the paper showing individual countries as dots, and a regression analysis summary output table constructed also from Excel software package will be shown for intuitive quantitative outputs. This paper will then utilize a prominent research from the World Bank Group in 2001 as an external source of explanatory tool that illuminate the regression analysis results in relation to ideology, culture and value systems. Finally, this paper will conclude by stating clearly the limitations and the reliability of the research procedures, and the potential inaccuracies in the statistical results.

2. Introduction of Regression Variables and Measures

It is important to explain the choices for the theme, data set and the quantitative measures of the two variables. As for the choice of this paper theme, the degree of income/wealth equality is an important demonstration of the efficiency and effectiveness of social welfare distribution mechanisms across countries. It is generally expected that with a more politically and economically free and democratic social environment, income distribution among citizens should become more even too, as high equality is supposedly an essential feature for stable and sustainable democracies. Moreover, income inequality is inevitably related to the management of corruption and the rule of law on a national level in world democracies. Much academic interest in the field of contemporary political science is given to the relationship between those two politically-orientated and economically-focused social factors.

As for the choice of the 36 OECD member states as my data set subjects, I aim to obtain information from a wide range of mostly developed and high-income economies with a high HDI (Human Development Index) across different world regions. OECD is generally a forum of states, founded in 1961 to stimulate economic progress and world trade, claiming themselves as committed to democracy and the market economy. [2]As a result, non-democratic countries and closed economies are excluded from the data set as they deviate from and are hardly in accordance with the core purpose of this paper's statistical comparison. In 2017, OECD member states collectively comprised 62.2% of global nominal GDP [3] and 42.8% of global GDP at purchasing power disparity. [4] Therefore, I regard the data set of OECD member states very effective for serving as the representation of economically important democracies on the world scale, especially that OECD is an official UN observer, [5] with relatively speedy data information collection due to time and other practical limitations, although the sample size for representation of the world is likely to be inefficient.

Furthermore, the choice of using Freedom House Index, which is a weighted average based on public opinion surveys, composes of numerical ratings and supporting descriptive texts for 195 countries and 14 territories based on assessment of each state's degree of political freedoms, civil liberties and human rights performance, [6] is due to its high applicability and citation rates by political scientists, policy-makers and journalists, although frequently being criticized that it's biased towards US interests as the NGO (Non-governmental Organization) Freedom

House is funded largely by the US government. [7] [8] Freedom House publishes its annual report, Freedom in the World since 1972, and assigns numbers to each country on a scale from 1 (most free) to 7 (least free), and countries are ranked and categorized as “free”, (1.0-2.5) “partly free”, (3.0-5.0) and “not free” (5.5-7.0) based on these scores. [9] Nevertheless, the organization has existed for more than 70 years and its data and rankings are generally informative and illustrative to political researchers.

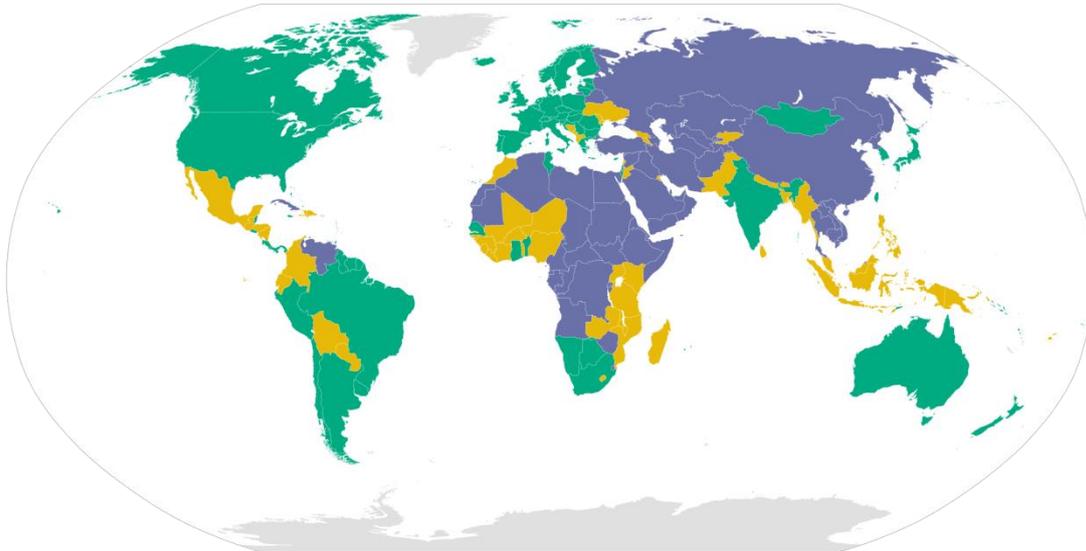


Figure 2. Country ratings from Freedom House's Freedom in the World 2018 survey, concerning the state of world freedom in 2017.
green-free (87), yellow-partly free(59), blue-not free(49)

Country ratings from Freedom House's Freedom in the World 2018 survey, concerning the state of world freedom in 2017 [10]

On the other hand, Gini coefficient/index/ratio, developed by the Italian statistician and sociologist Corrado Gini, is a measure of statistical dispersion among values of a frequency distribution, and a mathematical representation of a nation's Lorenz curve, (a graphical representation of the Gini coefficient) ranging from 0 to 1, which is currently the most commonly used measure of inequality. It measures the inequality among values of a frequency distribution of income. A Gini coefficient of 1 represents maximal inequality whereas a Gini coefficient of 0 represents perfect equality. [11] [12] In reality, for majority of world regions, their Gini coefficients fluctuate around 0.5. [13] For example, for OECD countries, in the late 20th century, considering the effect of taxes and transfer payments, the income Gini coefficient ranged from 0.24 to 0.49, with Slovenia being the lowest and Chile the highest. [14] Besides, the global income Gini coefficient in 2005 has been estimated to be between 0.61 and 0.68. [15] [16] In reality, Gini coefficients are usually represented by percentages on graphs and tables. There are several common criticism for the application of Gini coefficient, including the omission of demographic statures, e.g., an aging population, baby boom, etc, and the complexity of pre and post tax issues. [17] [18] [19] Besides, it doesn't capture dark market economic activities in the current economic research.

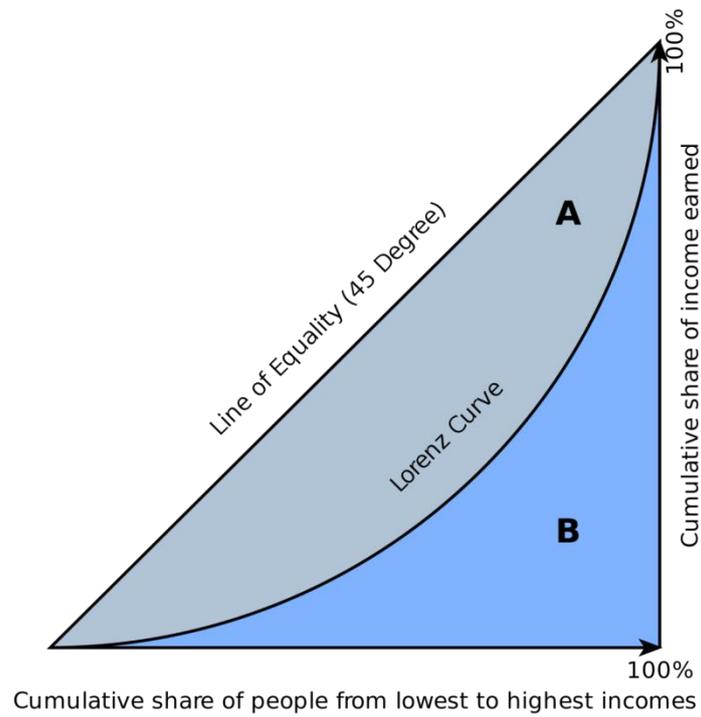


Figure 3. Lorenz Curve

Graphical representation of Gini coefficient [20]

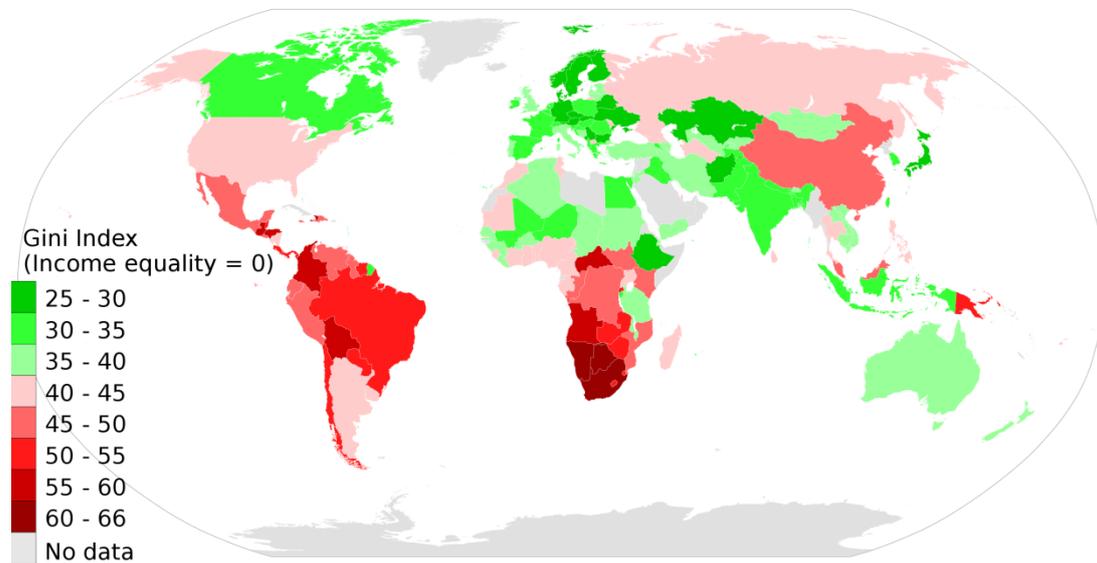


Figure 4. Gini Coefficient World

Countries' income inequality according to their most recent reported Gini index values [21]

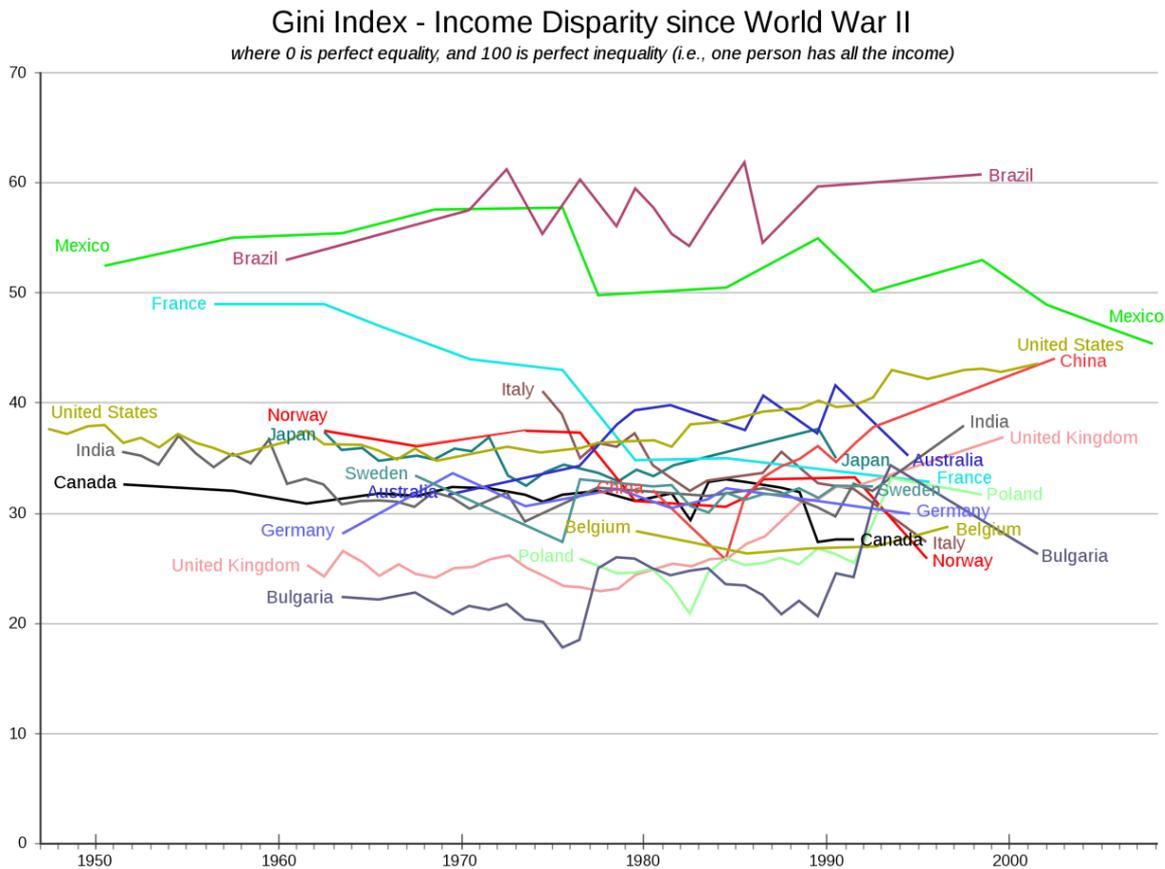


Figure 5. Gini Coefficient World After WW2

Gini coefficients for countries since WWII [22]

Expressing my presumptive hypothesis in terms of Freedom House index (x-axis) and Gini coefficient (y-axis), based on their definitions in according to their scores' assignments, a positive correlation pattern is expected to appear on the regression graph.

3. Explanation of Theoretical Basis and Data Sources

Mathematically, a linear regression function/equation is defined by this equation:

$$y = bx + a (+ \epsilon)$$

In this paper's linear regression analysis, x is the independent variable (a.k.a. criterion variable) is Freedom House score, and y is the dependent variable (a.k.a. Explanatory variable, or predictors) is Gini coefficient. b is the slope of the regression curve and a is the y-intercept. ϵ in parenthesis is the random error term, which is the difference between the actual value of a dependent variable and its predicted value. The linear regression function/equation always has an error term because, in real life, predictors are never perfectly precise and independent variables are never perfect predictors of the dependent variables. It is in parenthesis because in Excel, the software in which this linear regression analysis is conducted, does the random error term calculation behind the scenes.[23]

The source of Gini coefficients for the total population across OECD member states used in this paper's linear regression analysis is the OECD's official database (Other organizations also publish tables and graphs on Gini coefficient across countries such as the World Bank and CIA's the World Factbook) The report from different organizations vary in Gini coefficients' values due to different measurement procedures., e.g., different ways of dealing with taxes and public transfers in the calculation formula.

The figures in OECD report are from the latest years in which information is available, e.g., 2014, 2015, 2016 and mostly from 2017 for different countries. Based on the figures on the table, it can be observed that all OECD member states have a Gini coefficient lower than 0.5, which means they are relatively income equal and corruption free compared to the rest of the states. [24]

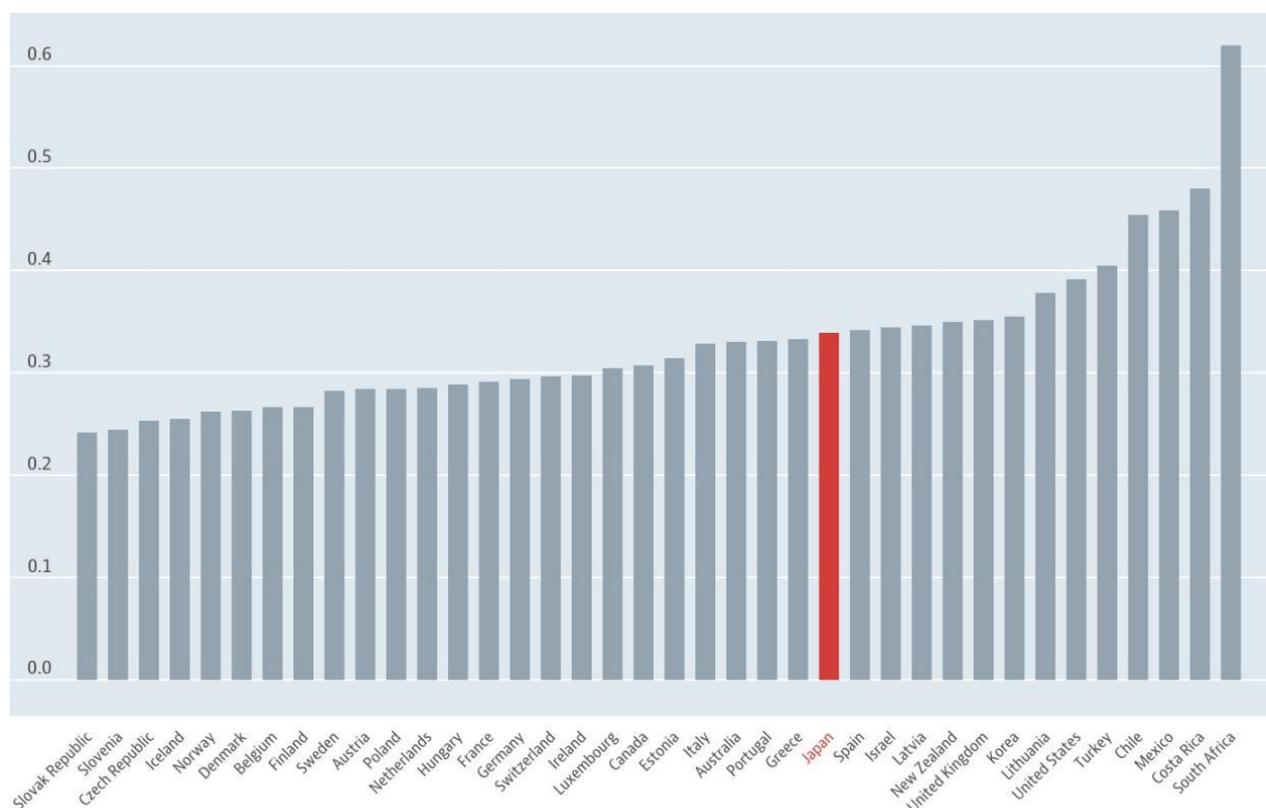


Figure 6. Gini Coefficient Across OECD member states

Gini coefficient, 0 = complete equality; 1 = complete inequality, 2017 or latest available
Income inequality across OECD member states (also includes Costa Rica and South Africa) measured by Gini coefficient from the latest years [25]

The data of Freedom House scores is obtained from the Freedom in the World 2018 Edition's Table of Country Scores (concerning the democratic freedoms level of world regions in the previous year, 2017) on the Freedom House official website. The reason for referring to data of 2017 is to correspond to the Gini coefficients of the most recent periods (mostly from 2017) for analytical consistency, as the information on Gini coefficients is less available and less frequently updated than the annually published Freedom House scores. The data from the column Aggregate Score is used as a combinations of Political Rights (PR) , Civil Liberties (CL) and Freedom Rating ranging from 0 (exception of -1 as in the case of Syria) to 100.

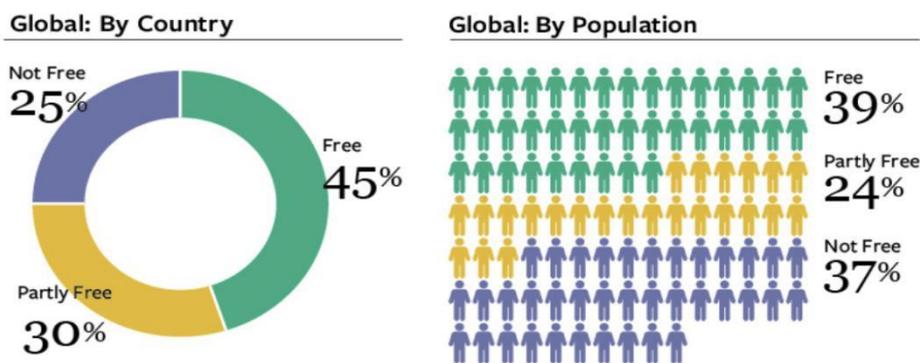
Show 50 entries

Country or Territory	Freedom Status	PR	CL	Freedom Rating	Aggregate Score
Timor-Leste	Free	2	3	2.5	69
Tonga	Free	2	2	2.0	75
Trinidad and Tobago	Free	2	2	2.0	81
Tunisia	Free	2	3	2.5	70
United States	Free	2	1	1.5	86
Vanuatu	Free	2	2	2.0	81
Andorra	Free	1	1	1.0	96
Australia	Free	1	1	1.0	98
Austria	Free	1	1	1.0	94
Bahamas	Free	1	1	1.0	91
Barbados	Free	1	1	1.0	96
Belgium	Free	1	1	1.0	95
Belize	Free	1	2	1.5	86
Canada	Free	1	1	1.0	99
Cape Verde	Free	1	1	1.0	90
Chile	Free	1	1	1.0	94
Costa Rica	Free	1	1	1.0	91
Croatia	Free	1	2	1.5	86
Cyprus	Free	1	1	1.0	94
Czech Republic	Free	1	1	1.0	93
Denmark	Free	1	1	1.0	97
Dominica	Free	1	1	1.0	93
Estonia	Free	1	1	1.0	94
Finland	Free	1	1	1.0	100
France	Free	1	2	1.5	90
Germany	Free	1	1	1.0	94

Freedom House Table

0-least free; 100-most free

Extracted from the Table of Country Scores of Freedom in the World 2018 Edition [26]



Freedom House status by country and population

Freedom in the World 2018, status by population and country [27]

Based on the observations from the Table of Country Scores, only two OECD member states are categorized as “partly free” and “not free” in 2017, which are Mexico and Turkey respectively. On the other hand, Finland, Norway and Sweden demonstrate the top world democratic freedom having scores of 100 during 2017.

The summary statistics are summarized in Excel as the following.

1	Countries	FH Aggregate Score	Gini Coefficient
2	United Kindom	94	0.351
3	United States	86	0.391
4	Australia	98	0.33
5	Austria	94	0.284
6	Belgium	95	0.266
7	Canada	99	0.307
8	Chile	94	0.454
9	Czech Republic	93	0.253
10	Denmark	97	0.263
11	Estonia	94	0.314
12	Finland	100	0.266
13	France	90	0.291
14	Germany	94	0.294
15	Greece	85	0.333
16	Hungary	72	0.288
17	Iceland	95	0.255
18	Ireland	96	0.297
19	Israel	79	0.344
20	Italy	89	0.328
21	Japan	96	0.339
22	Latvia	87	0.346
23	Lithuania	91	0.378
24	Luxembourg	98	0.304

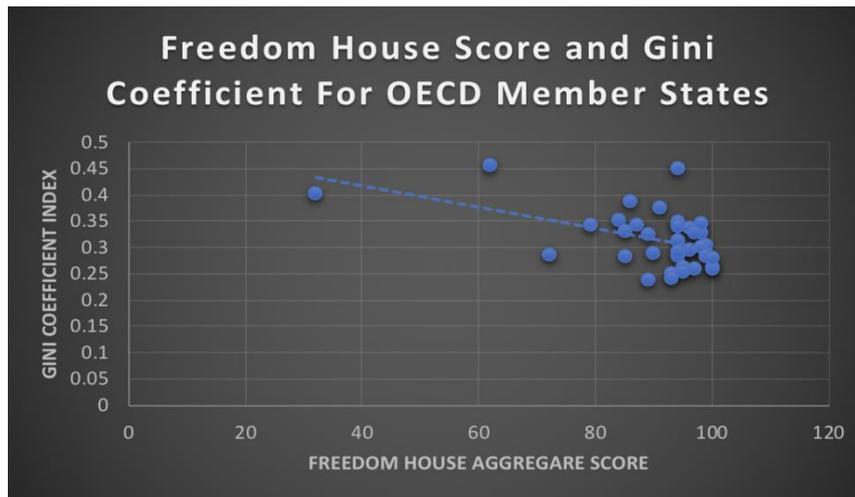
Excel Table

Excel table constructed from online database statistics.[28]

4. Graph From Regression Analysis and Interpretation of Statistical Results

For my presumptive hypothesis to be valid, a negative correlation pattern and a downward-sloping linear line is expected based on the two variables on the x and y-axes.

After running linear regression on Excel, software package, the following linear regression graph is obtained.



Regression Graph

Regression graph constructed from Excel software package [29]

As we can see from the above regression graph, my presumptive hypothesis stated in the introduction part of the paper is confirmed. A negative downward-sloping negative correlation pattern is observed. It is also worthy to highlight that there appear to have four outliers around the linear regression line that deviate from the predictions of the regression analysis. The countries the three outlier dots represent are Turkey, Mexico, Hungary and Chile from the left to right. Turkey and Hungary have a lower degree of income inequality than predicted by the regression curve, whereas Mexico and Chile exhibits a higher income inequality than predictions from the regression graph. Nevertheless, another basic observation which can be highlighted is that the spread of data points above and below the linear curve is pretty even.

The regression analysis summary output table, which provides more details is obtained by inserting the statistics into Excel software package is as follows, showing that the correlation has statistical significance at 95% level. The column of reference (P-value) which indicates statistical significance is filled with blue within the table. Other columns of information are to be interpreted on your own, as reference of all the columns and statistical terms and figures can be extremely complex and tedious, out of line with the purpose of this paper.

SUMMARY OUTPUT TABLE						
Multiple R	0.481050767					
R Square	0.23140984					
Adjusted R Square	0.208804248					
S.D.	0.047590855					
Observed Value	36					
Variance Analysis						
	df	SS	MS	F	Significance F	
Regression Analysis	1	0.023185312	0.023185312	10.23684012	0.002977872	
Residuals	34	0.077006244	0.00226489			
Sum	35	0.100191556				
	Coefficients	S.D.	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.498202239	0.057221569	8.706546257	3.57146E-10	0.381914019	0.61449
X Variable	-0.002011489	0.000628687	-3.19950623	0.002977872	-0.003289135	-0.00073

Regression Output Table

Regression analysis summary output table constructed from Excel software package [30]

5. Research Conclusion and Model Evaluation

The statistical results from the linear regression analysis confirms my presumptive hypothesis. However, it is important to note that the relationship is not necessarily linear over time, as linear regression is chosen for simplicity due to the statistical limitations. It would also be interesting if we extend our research in two ways. First, incorporate non-democratic and developing countries, and underdeveloped and closed economies into the statistical comparison, as the OECD sample is small and biased towards mostly democratic and economically developed states. As a result, a wider number, span and spread of data points can be analyzed. (A greater range of Gini coefficients and Freedom House scores) can be Second, beyond the latest information from OECD and Freedom House, data and figures from the past few decades can also be analyzed and compared to identify different trends in the correlation. By extending through these two ways, a more comprehensive comparative statistical research across countries (cross-sectional research) and through different time periods (longitudinal research) can be achieved, although it would require a much larger data set and a more complex regression analysis process. Alternatively, societies can be categorized based on social, cultural, ideological and religious structures showing different patterns of correlation, so that comparisons between regional and ideological subgroups can be made, instead of analyzing correlation within a whole general group of countries.

Moreover, for greater accuracy of the statistical analysis, other measures of political freedom from different organizations and other sources of Gini coefficients across world regions need to be applied and synthesized. The difficulty in obtaining the latest information of Gini coefficients from online database is another restriction on the research accuracy, as figures from inconsistent years are used.

Lastly, this paper only investigates and concludes the correlation between the two variables, it doesn't suggest or imply any causation in any direction. There is likely to be a range of confounding variables, and cultural, ideological, religious, socio-economic factors and value systems that can influence the democratic freedom level and the degree of income/wealth inequality simultaneously, and possibly, in multiple or opposite directions, statistically speaking. In order to investigate multiple effects and correlations of several variables, a multiple regression analysis is crucial. Interesting, studies are also done on the opposite direction between the two variables, about whether wealth and income distribution inequality poses a threat to existing democratic political regimes.

This paper devotes majority of its contents in the explanation of its theme, composing variables, and the statistical analysis procedures and the results' interpretations. Less effort is devoted in explaining the underlying and hidden reasons on why the correlation exists. For explaining more explicitly and thoroughly the correlation, other political theories and models from various political scientists and scholars are necessary. Therefore, it is beyond my own capacity and this paper's purpose and domain to offer a comprehensive set of valid, viable and possibly complicated explanations behind the simple linear regression results. (Due to the simplicity of this regression analysis method, insights and implications are also highly limited) For further inquiry, more observations, data collection, robust model building and statistical analysis are essential. Due to time and resource limitations and the specific requirements of this term paper, these issues are not fully addressed.

6. External Source of Explanation and Further Thinking

To obtain further insights of the research, I will utilize one prominent external research from World Bank Group with presumably high accuracy and reliability with large dataset and relatively long period of observation, on democracy and income inequality to illuminate the

regression analysis results. Even though a range of existing past research and published papers on this topic are available, I find this paper most concise and appropriate in explaining the correlation results from the regression analysis for illustrative and explanatory purposes. Moreover, the authors link the correlation to ideology and political value system, culture and religion.

The external research is from "Gradstein, Mark; Milanovic, Branko; Ying, Yvonne. 2001. Democracy and Income Inequality : An Empirical Analysis" published as a Policy Research Working Paper of World Bank Group. According to and quoting from the paper itself and its introduction on the official website, the motivations and the conclusions of the research goes as follows: "standard political economy theories suggest that democratization has a moderating effect on income inequality, whereas the empirical academia has failed to uncover any such robust relationship. The authors claim that prevailing ideology could be an important determinant of inequality and the democratization effect "works through" ideology. In societies that value equality highly there is less distributional conflict among income groups, thereby democratization may have only a negligible effect on inequality. On the other hand, in societies that value equality less, democratization reduces inequality through redistribution as the poor outvote the rich. Moreover, the authors conducted a cross-country empirical analysis, covering 126 countries in 1960-98, confirms their research hypothesis: ideology, as proxied by a country's dominant religion, seems to be related to inequality. Furthermore, while in Judeo-Christian societies increased democratization tend to result in lower inequality, in Muslim and Confucian societies it has an insignificant effect. They hypothesize that Muslim and Confucian societies rely on informal transfers to reach the desired level of inequality, while Judeo-Christian societies, where family ties are weaker, use more formal means of political action." [31]

If there is sufficient time in my future individual research, I will continue the discussion in this term paper, extend the research further and resort to and rely on more existing historical and recent research and published papers from other scholars and organizations instead of referring to only one research, as means to capture more worthy details under this far-reaching and wide-ranging political and economic topic of much academic value and interests.

Acknowledgements

- (1) Term Paper for Comparative Politics, Pr. Jou Willy
- (2) I would like to thank my seminar professor Jou Willy from the School of Political Science and Economics, Waseda University, for valuable advice on the structure and arrangement of the paper, and I would also like to thank my friend and fellow classmate, Ethan Ding from the Department of Economics, for his help with data analysis and regression.

References

- [1] Linear regression analysis in Excel-Ablebit.com.
- [2] Oecd.org.
- [3] "World Economic Outlook Database". International Monetary Fund. 17 April 2018.
- [4] "Report for Selected Country Groups and Subjects (PPP valuation of country GDP).IMF. Retrieved 9 May 2018.
- [5] "International Organizations." www.un.org.
- [6] Freedomhouse.org.
- [7] "Financial Statements". (PDF) Freedom House. June.30, 2016. Retrieved November 27, 2017.
- [8] 2006 Freedom House Annual Report.

- [9] Freedomhouse.org.
- [10] Freedom in The World 2018 by Freedom House, January 5, 2018.
- [11] Gini, C. (1912), Variability and Mutability.
- [12] Gini, C. (1909). "Concentration and dependency ratios" (in Italian). English translation in Rivista di Politica Economica, 87 (1997), 769–789.
- [13] "Gini Coefficient." United Nations Development program. 2012. Archived from the origin on 12 July 2014.
- [14] Income distribution – Inequality: Income distribution – Inequality – Country tables". OECD. 2012. Archived from the original on 9 November 2014.
- [15] Hillebrand, Evan (June 2009). "Poverty, Growth, and Inequality over the Next 50 Years" (PDF). FAO, United Nations – Economic and Social Development Department. Archived from the original (PDF) on 20 October 2017.
- [16] The Real Wealth of Nations: Pathways to Human Development, 2010 (PDF). United Nations Development Program. 2011. pp. 72–74.
- [17] Yitzhaki, Shlomo (1998). "More than a Dozen Alternative Ways of Spelling Gini" (PDF). Economic Inequality. 8: 13–30.
- [18] Sung, Myung Jae (August 2010). "Population Aging, Mobility of Quarterly Incomes, and Annual Income Inequality: Theoretical Discussion and Empirical Findings".
- [19] Blomquist, N. (1981). "A comparison of distributions of annual and lifetime income: Sweden around 1970". Review of Income and Wealth. 27 (3): 243–264.
- [20] Wikimedia Commons.
- [21] "Distribution of income or consumption" in tables World Development Indicators The World Bank (2014).
- [22] File:Gini since WWII.gif by Citynoise(en.wp)
- [23] Linear regression analysis in Excel-Ablebit.com.
- [24] Income inequality, data.oecd.org.
- [25] Income inequality, data.oecd.org.
- [26] Freedomhouse.org
- [27] Freedomhouse.org
- [28] Own work.
- [29] Own work.
- [30] Own work
- [31] <https://openknowledge.worldbank.org/handle/10986/19685>