An Empirical Investigation into the Contributing Factors of Income Inequality:

Rory Mc Stay

Trinity College Dublin

mcstayr@tcd.ie

March 19, 2017

Research Objectives

What are the contributing factors to inequality and of those factors, which features most prominently in a regression model?

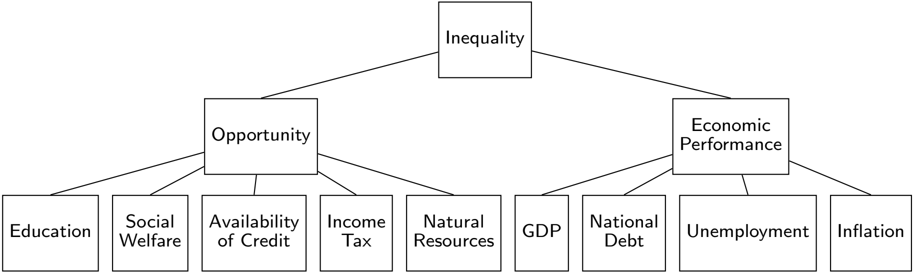
- Explore the relationship of determining characteristics of Inequality in an Economy.
- Determine which characteristics have the greatest effect on inequality in an economy.
- Classification for different degrees or origins of inequality of an economy.

These aims will be pursued through attempts to find simple and multiple regression models with inequality being the dependent variable as well as hierarchial clustering techniques.

Variables of interest

Inequality will be investigated under the Gini coefficient. This offers a description of inequality throughout all levels of income.
The explanatory variables of interest consist of;

1. National Debt/Debt Servicing (%GDP)
2. Social Welfare/Contributions
3. Income Tax Level
4. Natural Resources
5. GDP/Capita
6. Unemploment
7. Inflation
8. Education Expenditure
9. Availability of Credit


Inequality
Opportunity
Economic Performance
Education
Social Welfare
Availability of Credit
Income Tax
Natural Resources
GDP
National Debt
Unemployment
Inflation

Sample Population

Due to availability of accurate statistics, the sample population consists of OECD member states. This sample is not random and each observation meets the following criteria.

Political: stability of institutions guaranteeing democracy, the rule of law, human rights and respect for and protection of minorities;

Economic: a functioning market economy and the capacity to cope with competition and market forces;

Administrative and Institutional Capacity: effectively implement the acquis* and ability to take on the obligations of membership.

Data sources: World Bank, OECD.org
Link to data:
https://drive.google.com/drive/folders/0B_ntIGJEKrTWNmpGUFo2ZUVrVG

Assumptions

- No Autocorrelation - Inequality is determined by variables of an economy and is independent of time. Thus data will be treated as panel data initially.
- Random Sample - Our sample consists of a random sample of functioning econonomies as per the outlined criteria.
- Linear in Parameters - $G_i = \beta_0 + \beta_1 X_i + u_i$
- Zero Conditional Mean - $E[u_i|X] = 0$
- Homoscedascity - $V[u_i|X_i] = \sigma^2$