Third Assignment

Christopher Gandrud, Hertie School of Governance, Spring 2016

Emilia Sicari & Rafael Lopez V.

March 25, 2016

Contents

1 Research question and project description

In our collaborative research project, we seeks to provide an answer to the following question: **How the rise in the rise in inequality and economic growth influences the purchase of luxury goods, particularly cars, and the usage of public transportation systems in Singapore, from 1995 to 2014**. Therefore, we collected data on economic growth, income inequality, usage of public transports and purchase of cars covering the time span of 29 years, from 1995 until 2014 As suggested by our research question, economic growth, income inequality and usage of public transports are the explanatory variables, while purchase of cars is the dependent variable. The reason why we chose cars as example of luxury goods showing social status, is that in Singapore purchasing of cars is particularly expensive, due to high taxation and a certificate of car entitlement, whose cost can be higher than 70.000 dollars. For more details about the research proposal and case justification see ResearchProposal.

2 Processing data

2.1 Data sources and data gathering

The data that we need for our empirical analysis are to be retrieved from different sources:

- IMF Cross Country Macroeconomic Statistics open data, containing cross-country marcoeconomic data.From this source we downloaded the dataset on Singapore GDP per Capita at Current Prices measured in national currency from 1995 until 2014. Data were downloaded in csv format from Quandl, a website providing high-quality financial and economic data in different formats to facilitate data analysis.
- Knoema, a knowledge platform connecting data with analytical and presentation tools, in order to allow users to access, present and share data-driven content. From The World Top Incomes Database - providing access to data on the distribution of top incomes in more than twenty five countries across the globe - we downloaded data on the top 10% average income and bottom 90% average income in Singapore, to be compared in order to understand how many times the top 10% average income is higher than the bottom 90%. Since data were available only until 2009, the values of bottom 90% and top 10% average income between 2010 and 2014 were forcated with a linear regression. The reason why we did not gather data from the database Clio Infra available on Quandl as we had defined in our ResearchProposal , is that it did not provide sufficient data for the time span we are considering.
- PUBLIC TRANSPORTS
- PURCHASE OF CARS

2.2 Cleaning and merging of datasets

After importing data on R and cleaning them, we merged the datasets. The final dataset has nineteen observations, and eleven variables, observed throughout the years 1995 - 2014.

3 Descriptive and inferential statistics

References