Analyzing the Economic Impacts of Government Expenditure and Income

Mehrad Hajati, Ahnaf Iqbal, Bright Adams

Government expenditure and income plays a pivotal role in shaping a nation. It will influence all sectors within a country. This project aims to explore relationships between government expenditure on education and research to how employment and other factors are affected. Using the World Bank dataset (2000–2019) provided on Kaggle with the following link: https://www.kaggle.com/datasets/robertolofaro/selected-indicators-from-world-bank-20002019

The dataset provides key points on every country in the world and what percentage of GDP was spent in what sectors. It also provides data on the society of a country, for example labor force participation, unemployment rates, trained teachers and high-tech exports.

Our dataset contains many columns (37 in total) which are not related to our research questions so some will be dropped in the pre-processing stage. For each research question, we will use statistical methods (dplyr, corrplot) to determine which variables in our dataset are most crucial. Since our research questions are very focused and specific, finding which variables to disregard will be relatively easy. We will also clean and organize the dataset to address missing or inconsistent data, then normalize variables to ensure comparability across countries and time.

These are our research questions:

1. **What is the relationship between Tax revenue and annual GDP growth?**
    a. Our data set includes the percentage of government income and percentage of that which is from taxes. It also includes the annual GDP growth for each country. Allowing us to explore relationships.
    b. High taxation rates can be an important source of income for a government; but this research question will be focused on if that income is always translated to a faster growth in the economy or not.
2. **What is the relationship between government expenditure on R&D (research and development) and high-tech export?**
    a. Our data set includes the percentage of government expenditure spent on Research and Development. It also includes high-tech exports as a

percentage of total manufactured exports. Allowing us to explore relationships.

b. High-tech exports can be very beneficial for a country and can act as a boost to the economy but is it worth the R&D cost?

3. **What is the relationship between government expenditure on education and unemployment rates?**
    a. Our data set includes the percentage of government expenditure spent on Education. It also includes percentage of unemployment rates for both women and men.
    b. Expenditure on education can be incredibly beneficial for a country as it leads to a smarter population, contributing more to the economy using their studies.

For each of the research questions we will use linear and generalized linear (lm, mgcv. lme4) to identify relationships between the variables involved. We will then conduct hypothesis testing along with ANOVA (t.test, stats::anova, p-values) to gain more information. Following the analysis, we will use visual aids (mainly scatterplots in ggplot) to showcase the relationships during our presentation. The statistical methods mentioned here are the ones we are currently familiar with. However, with the progression of the course material, we expect to have more experience with new methods which we can then apply.

Data Sources

1. https://www.kaggle.com/datasets/robertolofaro/selected-indicators-from-world-bank-20002019