Economics 361
Final Exam: Empirical Exercise

Jun Ishii [*]
Department of Economics
Amherst College

Fall 2022

The data for this portion of the Final Exam may be downloaded from

- https://databank.worldbank.org/source/world-development-indicators

The website is the public portal providing access to the World Bank's "World Development Indicators" (WDI) database. The WDI database allows one to construct a macroeconomic panel dataset (country, year) involving most industrialized countries in the post-WWII years.

Later in this document, I provide instructions on how you may download your own customized dataset from the WDI portal. I have also, on our Moodle page, posted a dataset I had extracted from WDI some years ago. You may choose to use the provided data but I would recommend exploring the customized option first as the provided data is highly limited: the extract contains a small selection of the variables available from WDI and just for 30 OCED countries from 1998 through 2008 (the ten years leading up to the Financial Crisis).

Due date for the Assignment will be announced in class

---
[*]Office: Converse Hall 315 Phone: (413) 542-2901 E-mail: jishii@amherst.edu

Overall Instructions for the Assignment

The WDI data can be used to explore a wide variety of issues concerning the economy and residents of the countries studied by the World Bank. Your task, generally, is to use the data to conduct some kind of **econo**metric study. Your study must have both a **point estimation** and **hypothesis testing** component.

In 2 to 5 typed (10 to 12 font) single spaced pages (excluding any tables, graphs)

- Motivate your chosen econometric study

- Explain the steps involved in your chosen study

- Discuss the strengths and weaknesses of your chosen study

- Discuss your results

Additionally, you must provide a STATA log-file that documents the STATA commands and results underlying your econometric study (not part of the 5 page limit). If you choose to use some other econometrics/statistics software package (e.g. R), you must provide a document analogous to the STATA log-file. Also, if you downloaded your own customized WDI dataset, you will need to provide me with a **copy of the dataset used in your study**.

The **more challenging** the study, the **more lenient** will be my grading of the execution of the study. The point of this assignment is for you to show me that you can apply/implement the concepts taught in this course. Also, I will grade based on the merit of your study and your discussion, not the actual results. No need for a sensational result to get a good grade.

You may discuss your study with your fellow classmates, but nobody outside of the class (faculty or student). No two students may turn in the same study (excepting pure coincidence – something that is unlikely for this open-ended assignment).

Some general advice:

- Browse the data! (in STATA, use Data Editor or Data Browser)

- Keep a log-file running, even when you are just playing around with the data

- Some of you may be interested in the STATA command xtset

  - But to use xtset, you will have to convert the country variable into a numeric (e.g. assign "AUS" to country code "1")

- Commit to a particular study; do not waffle too much! Spend your time refining/explaining your chosen study

Instructions for Downloading WDI Data of Your Own

[**1**] On the menu to the left, click the tab `Layout`

- For `Time` choose the option `Row`
- For `Series` choose the option `Column`
- For `Country` choose the option `Row`

After you have made the above changes, click the tab `Variables`

[**2a**] The `Database` should already be set at `World Development Indicators`. So, we should be able to proceed onto `Country` selection. You are welcome to use any reasonably set of countries but my **strong** recommendation would be for you to limit your selection to the OECD countries (currently 36), as these are the countries for which there are likely to be more complete data

Australia	Austria	Belgium	Canada	Chile	Czech Republic
Denmark	Estonia	Finland	France	Germany	Greece
Hungary	Iceland	Ireland	Israel	Italy	Japan
Korea, Republic of	Latvia	Lithuania	Luxembourg	Mexico	Netherlands
New Zealand	Norway	Poland	Portugal	Slovak Republic	Slovenia
Spain	Sweden	Switzerland	Turkey	United Kingdom	United States

[**2b**] Once the countries have been selected, proceed onto `Series` selection. Here you will be selecting the variables (as in $\{Y, X\}$) you will be downloading for your empirical assignment. To select a series, click on the empty box to the left. To learn more about the series, click on the "(i)" information icon. Note: the more specific/detailed the series, the more likely the data is available only for a highly selected set of countries and/or years.

[**2c**] Once the series have been selected, proceed onto `Time` selection. Here, you will be selecting the years for which you will be downloading the selected series for the selected countries. You may choose all but note that, for many series, there will be no available data for the many of these years. For many series, data will only be available from roughly 2000 - 2017.

[**3a**] After you are done with `Database`, `Country`, `Series`, and `Time` selection, click the `Apply Changes` button the left side of the screen (labeled "Preview").

[**3b**] Once your request has been processed, click on the tab to the upper left corner `Download options`. Select the option that works best for you; I strongly recommend either `Excel` or `CSV`. This will allow you to download the requested data in the requested format.

[**4**] The downloaded dataset should be of the following format:

- Each row should represent data for one observation (year, country)
- The first four columns should be: time, time code, country name, country code. These four columns identify the observation

    I recommend deleting the columns "time code" and "country name" as they are redundant

- The remaining columns should be for each of the series you had selected

    If a column is filled by ".." then the series is unavailable for that (year, country). STATA estimation commands (e.g. `regress`) automatically drops observations with missing data (for the relevant variables).

- If the download format is `Excel`, there should be two sheets – one containing the actual data and the other the description of the downloaded series

    You will want to save the sheet containing just the data separately – makes it easier for STATA/R to read the data. I would recommend saving that data sheet as a comma-delimited (CSV) file. You can find the option under "Save As" in Excel/Google Spreadsheet.

Instructions for the Provided Data

On our Moodle page, I provide the following

- wb_total_F2022.dta is the STATA dataset version of the data, which may be read directly into STATA using the "Open" command under the File Menu

- wb_total_F2022.csv is the comma-delimited version of the data, which may be read into STATA using the infile command

- wb_total_F2022.xls is the Microsoft Excel version of the data

The .dta version of the data is the most convenient to use; the variables have already be named and given labels in that version.

There are 56 variables. The file wb_F2020_codebook.pdf lists the variables and provides a short description for each. The variables vary by (country, year) except for: (countryname, countrycode, easebusiness, rigidemp) which are time invariant and year which is country invariant.

There are 30 countries x 10 years = 300 total observations. The 30 countries are the 30 high income OECD countries and the 10 years are 1999-2008.

Data may be **missing** for some variables and observations. To see how much, use the STATA command summarize or the Data Browser. STATA estimation commands (e.g. regress) automatically drops observations with missing data (for the relevant variables).

The data was obtained from the World Bank's World Development Indicators (WDI), online version. It is one of many online databases to which Amherst College subscribes. You can read more about the database at

- https://datatopics.worldbank.org/world-development-indicators/