Documentation of the EPPE Term Paper project

Ricardo Duque Gabriel

15 February 2019

CONTENTS

CHAPTER

ONE

INTRODUCTION

Documentation on the rationale, Waf, and more background is at http://hmgaudecker.github.io/econ-project-templates/

The Python version of this specific project uses a modified version of Guilherme Nazareth synth package - https://github.com/gnazareths/synth.

1.1 Project paths

A variety of project paths are defined in the top-level wscript file. These are exported to header files in other languages. So in case you require different paths (e.g. if you have many different datasets, you may want to have one path to each of them), adjust them in the top-level wscript file.

The following is taken from the top-level wscript file. Modify any project-wide path settings there.

```
def set_project_paths(ctx):
    """Return a dictionary with project paths represented by Waf nodes."""

    pp = OrderedDict()
    pp["PROJECT_ROOT"] = "."
    pp["IN_DATA"] = "src/original_data/"
    pp["IN_MODEL_CODE"] = "src/model_code"
    pp["IN_MODEL_SPECS"] = "src/model_specs"
    pp["LIBRARY"] = "src/library"
    pp["BLD"] = ""
    pp["OUT_DATA"] = f"{out}/out/data"
    pp["OUT_ANALYSIS"] = f"{out}/out/analysis"
    pp["OUT_FINAL"] = f"{out}/out/final"
    pp["OUT_FIGURES"] = f"{out}/out/figures"
    # OUT_MODEL_SPECS is only required for using Stata with JSON and
    # can be safely deleted otherwise
    pp["OUT_MODEL_SPECS"] = f"{out}/src/model_specs"
    pp["OUT_TABLES"] = f"{out}/out/tables"
```

As should be evident from the similarity of the names, the paths follow the steps of the analysis in the `src` directory:

1. **data_management → OUT_DATA**
2. **analysis → OUT_ANALYSIS**

3. **final → OUT_FINAL, OUT_FIGURES, OUT_TABLES**

These will re-appear in automatically generated header files by calling the `write_project_paths` task generator (just use an output file with the correct extension for the language you need – `.py`, `.r`, `.m`, `.do`)

By default, these header files are generated in the top-level build directory, i.e. `bld`. The Python version defines a dictionary `project_paths` and a couple of convencience functions documented below. You can access these by adding a line:

```
from bld.project_paths import XXX
```

at the top of you Python-scripts. Here is the documentation of the module:

bld.project_paths

Define a dictionary *project_paths* with path definitions for the entire project.

This module is automatically generated by Waf, never change it!

If paths need adjustment, change them in the root wscript file.

project_paths_join(*key*, **args*)

Given input of a *key* in the *project_paths* dictionary and a number of path arguments *args*, return the joined path constructed by:

```
os.path.join(project_paths[key], *args)
```

project_paths_join_latex(*key*, **args*)

Given input of a *key* in the *project_paths* dictionary and a number of path arguments *args*, return the joined path constructed by:

```
os.path.join(project_paths[key], *args)
```

and backslashes replaced by forward slashes.

CHAPTER

TWO

SYNTH PACKAGE

Documentation of the code in *src.library*.

2.1 Synthetic Control Method

It has two main parts. In the first it is done the data preparation part. In the second, the most important, is where the synthetic control group is constructed as explained in the research paper.

dataprep (*foo*, *predictors*, *treated_unit*, *control_units*, *index_variable*, *measured_variable*, *Weights*, *time_variable*, *predict_time*, *optimize_time*, *plot_time*, *function='mean'*)

Given input of a dataframe, the predictors and outcome variables, the treated and control units, the initial guess for the weights, the time variable used, and intervals of periods in which it should predict, optimize and plot; this function will return an error if at least one of the inputs is wrongly specified:

synth_tables (*foo*, *predictors*, *treated_unit*, *control_units*, *index_variable*, *measured_variable*, *Weights*, *time_variable*, *predict_time*, *optimize_time*, *plot_time*, *function='mean'*)

Given input of a dataframe, the predictors and outcome variables, the treated and control units, the initial guess for the weights, the time variable used, and intervals of periods in which it should predict, optimize and plot; this function will return a dataframe with the optimal weights and will plot the final result of both doppelganger and original country's series for the outcome variable:

```
synth_tables(dataframe,
             predictors list,
             treated unit,
             control units,
             index variable,
             outcome variable,
             initial guess for weights (np.ndarray),
             time variable,
             predict period,
             optimize period,
             plot period,
             function="mean")
```

CHAPTER

THREE

CASE STUDIES FOR JAPAN AND UNITED KINGDOM

Documentation of the code in *src.final*.

3.1 Synthetic Control Method for Japan

Created on Sun Feb 3 14:08:09 2019 @author: Ricardo Duque Gabriel

The file Japan_SCM.py has the code for the creation of a SCG for Japanese GDP series using 31 OECD countries and saves the plot of both synthtetic and original series.

3.2 Synthetic Control Method for UK

Created on Mon Feb 11 14:08:09 2019 @author: Ricardo Duque Gabriel

The file GBR_SCM.py has the code for the creation of a SCG for Japanese GDP series using 31 OECD countries and saves the plot of both synthtetic and original series.

3.3 Weight Tables for Japan

Created on Wed Feb 13 9:33:18 2019 @author: Ricardo Duque Gabriel

Code for the creation of a Table in latex with the composition (in weights) of the doppelganger of Japan.

3.4 Weight Tables for UK

Created on Wed Feb 13 9:33:18 2019 @author: Ricardo Duque Gabriel

Code for the creation of a Table in latex with the composition (in weights) of the doppelganger of the United Kingdom.

CHAPTER

FOUR

RESEARCH PAPER / PRESENTATIONS

Purpose of the different files (rename them to your liking):

- `research_paper.tex` contains the actual paper.
- `research_pres_30min.tex` contains a typical conference presentation.
- `research_pres_90min.tex` contains a full-length seminar presentation (add by yourself).
- `formulas` contains short files with the LaTeX formulas – put these into a library for re-use in paper and presentations.

PYTHON MODULE INDEX

INDEX