**Lines 9-67 cover the creation of the dataset**

**If preferred you can go directly to line 70**

**You will, however, need to install synth (see line 19), synth_runner (see line 22) and grclleg (see line 33)**

*Data sources: Life expectancy: Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 29/03/2017).*


*Combining datasets

*First import txt files from HMD into stata, keep years 1960 to 2003. The countries included (as per Comparative Politics and the Synthetic Control Method are Australia

*Austria, Belgium, Denmark, France, Greece, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Switzerland, UK, USA, West Germany. Greece only starts 1981 in HMD so we exclude)*

*If not already installed you will need to install the following

`scc install synth, replace all`


`net install synth_runner, from( "https://raw.github.com/bquistorff/synth_runner/master/") replace`

`dsconcat - Roger Newson, Imperial College London, UK.
ssc install dsconcat, replace`

ssc install kountry, replace`

`grclleg Program by Vince Wiggins, StataCorp <vwiggins@stata.com>,
net install grclleg.pkg, replace`

*assumes hmd files (for year on year life expectancy for above countries) in a directory on their own
cd E0per

*imports text files and saves them as stata files
local list : dir "." files ".txt", respectcase
foreach file of local list {
import delimited 'file', varnames(3) delimiters(" ", collapse) clear
gen filename="file"
gen country=substr(filename,1,3)
levelsof country
keep year country total male female
drop if year < 1960
drop if year > 2003
save `r(levels)', replace
}
*combines the imported stata files from the last step (uses user written programme dsconcat so if you don't have this use the * out command to install from ssc).
local list2 : dir "." files ".dta", respectcase
dsconcat `list2`

*next merge the files using a user written programme - kountry- again download this if haven't already

kountry country, from(iso3c)
save "germany\le", replace
cd "germany"
use repgermany, clear
kountry country, from(other)
drop if country=="Greece"
merge 1:1 NAMES_STD year using le
drop _merge male female
encode country, gen(country2)
label var total "Life expectancy"
save analysis, replace

*************************************************************************
* Analysis - start here
*************************************************************************
use analysis, clear
*Step 1 - requires no syntax as it concerns theoretical understanding*
*Step 2 - Identification of potential control units - remaining blinded to data
post implementation*
keep if year < 1991
*exclusions - keep Austria, Japan, Netherlands, Switzerland and USA as well as West
Germany as these used in GDP study.
*compares West German trend to the mean of the rest of the 15
egen m_total = mean(total) if country2!=16, by(year)
egen m_gdp = mean(gdp) if country2!=16, by(year)
*pooled after exclusions
egen m_total_ex = mean(total) if inlist(country2,2,7,8,13,15), by(year)
egen m_gdp_ex = mean(gdp) if inlist(country2,2,7,8,13,15), by(year)
*Comparison of average GDP and Life Expectancy trends between West Germany and 5
country/15 country pools*
line m_total year if country2==2, lpattern(dash) lcolor(black) ||
line m_total_ex year if country2==2, lpattern(dash_dot) lcolor(black) ||
line total year if country2==16, name(match_le, replace)
*title("Life expectancy") xline(1990, lcolor(gs8)) ///
legend(label(1 "15 country pool") label(2 "5 country pool") label(3 "West Germany"))
lcolor(black)
line m_gdp year if country2==2, lpattern(dash) lcolor(black) ||
line m_gdp_ex year if country2==2, lpattern(dash_dot) lcolor(black) ||
line gdp year if country2==16 &
year, name(match_gdp, replace)
*title("GDP per capita") xline(1990, lcolor(gs8)) ///
legend(label(1 "15 country pool") label(2 "5 country pool") label(3 "West Germany"))
lcolor(black)
graph leg match_le match_gdp, xcommon
*Figure generated by line 98 (not shown in article) shows better GDP fit for 5
country pool so that is used in the rest of the analysis*
keep if inlist(country2,2,7,8,13,15,16)
*Step 3 - Develop the synthetic control country - a synthetic control West Germany
tsset country2 year
**This approach uses the final gdp observation in the pre-implementation period as
predictor variable**
**Lines 112-124 create the top half of Figure 1**
sort year country2
matrix W = e(W_weights)
svmat W
bysort country2: egen weight = max(W2)
egen m_gdp_temp = total(gdp*weight) if country2!=16 & weight > 0, by(year)
egen m_total_temp = total(total*weight) if country2!=16 & weight > 0, by(year)
**This approach uses averages of GDP over five-year periods in the pre-implementation period as predictor variables**

*Figures 126-142 create the bottom half of Figure 1**

**Line 145 creates Figure 1**

grclegend match_le_1 match_gdp_1 match_le_2 match_gdp_2, xcommon

**Step 4 - Run outcome analysis**

use analysis, clear

set more off

keep if inlist(country2,2,7,8,13,15,16)

**Lines 156-172 create Figure 2**

tset country2 year


drop W weight m_gdp_temp m_total_temp
**Step 6 - Run robustness checks**

**Lines 176-186 create Figure 3**

tempfile keepfile

```
, trunit(16) trperiod(1990) keep('keepfile')

merge 1:1 country2 year using "`keepfile'", nogenenerate

gen double total_synth = total-effect

line effect year if country2 == 2, lcolor(gs8) || ///
line effect year if country2 == 7, lcolor(gs8) || ///
line effect year if country2 == 8, lcolor(gs8) || ///
line effect year if country2 == 13, lcolor(gs8) || ///
line effect year if country2 == 15, lcolor(gs8) || ///
line effect year if country2 == 16, lcolor(gs8) legend(off) linewidth(thick) xline(1990, lcolor(gs8)) yline(0, lcolor(gs8)) ytitle(Life expectancy difference)
```

*RMSPE

**Lines 190-191 generate the data for Table 4**

gen ratio_rmspe = post_rmspe / pre_rmspe
tabstat pre_rmspe post_rmspe ratio_rmspe, by(country2) nototal

**Further sensitivity analysis - not discussed in article**

*Do exclusions matter?

use analysis, clear

drop more off

tset country2 year

sort year country2

matrix W = e(W_weights)

svmat W

degen weight = max(W2)

degen m_gdp_temp = total(gdp*weight) if country2!=16 & weight > 0, by(year)
degen m temp = total(total*weight) if country2!=16 & weight > 0, by(year)
sort country2 year

legend(label(1 "Synthetic West Germany") label(2 "West Germany") name(match_gdp, replace) ytitle("GDP per capita") xline(1990, lcolor(gs8)) lcolor(black)

matrix list e(V_matrix)

tempfile keepfile

```
, trunit(16) trperiod(1990) keep('keepfile')

merge 1:1 country2 year using "`keepfile'", nogenenerate

gen double total_synth = total-effect

line effect year if country2 == 1, lcolor(gs8) || ///
line effect year if country2 == 2, lcolor(gs8) || ///
line effect year if country2 == 3, lcolor(gs8) || ///
line effect year if country2 == 4, lcolor(gs8) || ///
line effect year if country2 == 5, lcolor(gs8) || ///
line effect year if country2 == 6, lcolor(gs8) || ///
line effect year if country2 == 7, lcolor(gs8) || ///
line effect year if country2 == 8, lcolor(gs8) || ///
line effect year if country2 == 9, lcolor(gs8) || ///
line effect year if country2 == 10, lcolor(gs8) || ///
line effect year if country2 == 11, lcolor(gs8) || ///
line effect year if country2 == 12, lcolor(gs8) || ///
line effect year if country2 == 13, lcolor(gs8) || ///
line effect year if country2 == 14, lcolor(gs8) || ///
line effect year if country2 == 15, lcolor(gs8) || ///
line effect year if country2 == 16, lcolor(gs8) legend(off) linewidth(thick) xline(1990) yline(0) ytitle(Effe
*RMPSE

``` stata
238  gen ratio_rmspe = post_rmspe / pre_rmspe
239  tabstat pre_rmspe post_rmspe ratio_rmspe, by(country2) nototal
241
242
243
244
245```
