R Notebook

```
### Code for "Democratization and Economic Output in Sub-Saharan Africa"
### Daniel De Kadt and Stephen B. Wittels

## A note to users:
## Use setwd() to set the working directory to the location where data files are saved.
## Figures are programmed to be saved automatically to the working directory.
## They are named according to their figure number in the paper.
## The three tables in the paper are saved as the objects "mali.weights," "panel.estimates," and
## "moderators." Code to print these objects to the console is included at the end.


## Options and Libraries
options(scipen = 6, digits = 3)

# Install necessary libraries
if (!require("pacman")) install.packages("pacman")
```

```
## Loading required package: pacman
```

```
pacman::p_load(foreign,
  Synth,
  xtable,
  rgenoud,
  reshape2,
  quadprog,
  ucminf,
  Rcgmin,
  Rvmmin,
  minqa,
  Rcpp,
  ggplot2,
  plyr,
  grid,
  lme4,
  janitor,
  dplyr,
  CausalImpact  # For use in the extension
)
```

```
## Data
load("afripanel_wdk_final.RData")
a <- read.csv("conditioning_variables1.csv")
panel.reg <- read.dta("panel.reg1.dta")

not_any_na <- function(x) all(!is.na(x))
```

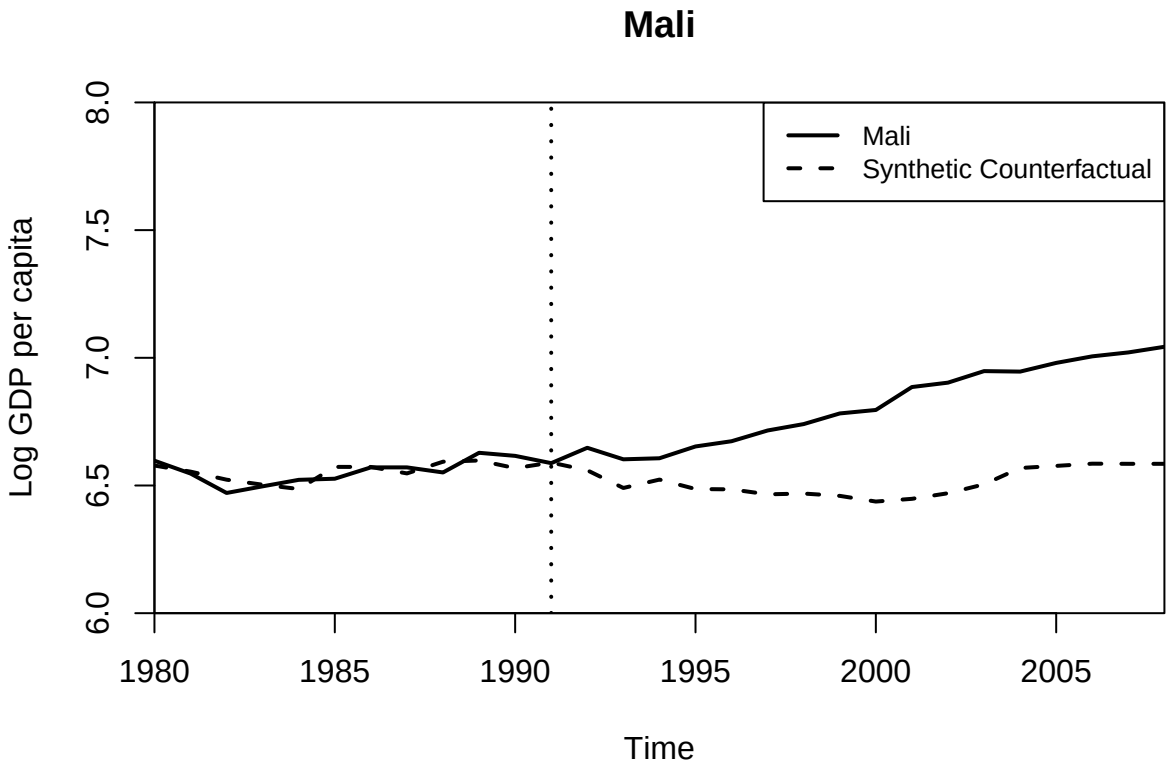
Replication

```
# Replication function
replicate <- function(
  unitID,
  fullname,
  begin,
  end,
  tr2,
  final,
  low,
  high
){

  data <- afripanel[afripanel$WBCode==unitID | afripanel$cont_dem_ind==1,]

  controls <- unique(data$WBCode[data$WBCode!=unitID&data$WBCode!="ETH"&data$WBCode!="SDN"])

  prep <- dataprep(
    foo=data,
    predictors=c(
      "lngdpmadlag",
      "lngdpmadlag2",
      "lngdpmadlag3",
      "lngdpmadlag4",
      "lnpop",
      "ki",
      "openk",
      "civwar",
      "civwarend",
      "pwt_xrate",
      "pwt_xrate_lag1",
      "pwt_xrate_lag2",
      "pwt_xrate_lag3",
      "eximdiff",
      "eximdiff_lag1",
      "eximdiff_lag2"
    ),
    dependent="lngdpmad",
    unit.variable="wbcode2",
    time.variable="year",
    treatment.identifier=unitID,
    controls.identifier=controls,
    time.predictors.prior=c(begin:end),
    time.optimize.ssr=c(begin:tr2),
    time.plot=c(begin:final),
    unit.names.variable="WBCode"
  )

  out <- synth(prep)

  path.plot(synth.res=out, dataprep.res=prep,
            Ylab="Log GDP per capita", Legend=c(fullname, "Synthetic Counterfactual"), tr.intake=tr2,
```

```
            Ylim=c(low,high) , Main=fullname
  )

}
```

```
## Figure 2 Replication
replicate("MLI", "Mali", 1980, 1990, 1991, 2008, 6, 8)
```

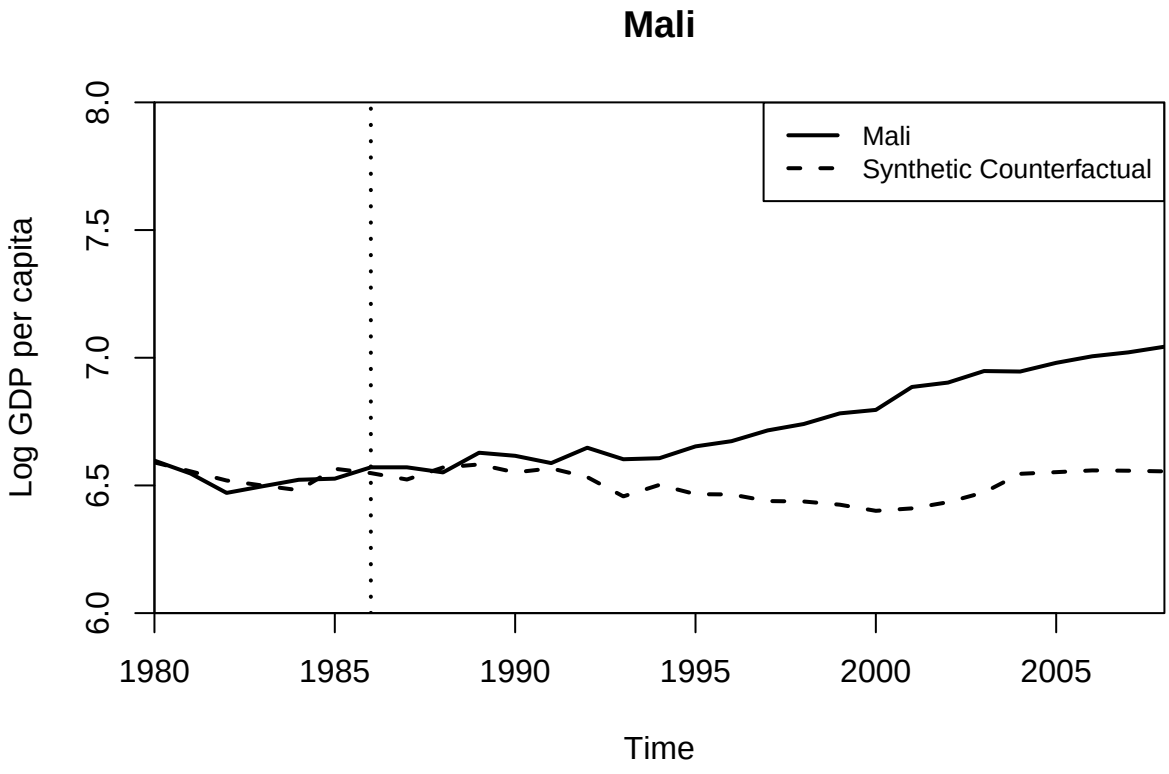
```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00101
##
## solution.v:
##  0.0836 0.123 0.125 0.156 0.14 0.0000143 0.127 0.0354 0.0425 0.0252 0.027 0.0301 0.033 0.000281 0.01
##
## solution.w:
##  0.000000127 0.235 0.115 0.000000639 0.0000537 0.0000719 0.000000269 0.0000137 0.000000612 0.0000007
```




```
## Figure 2 Replication in-time placebo
replicate("MLI", "Mali", 1980, 1985, 1986, 2008, 6, 8)
```

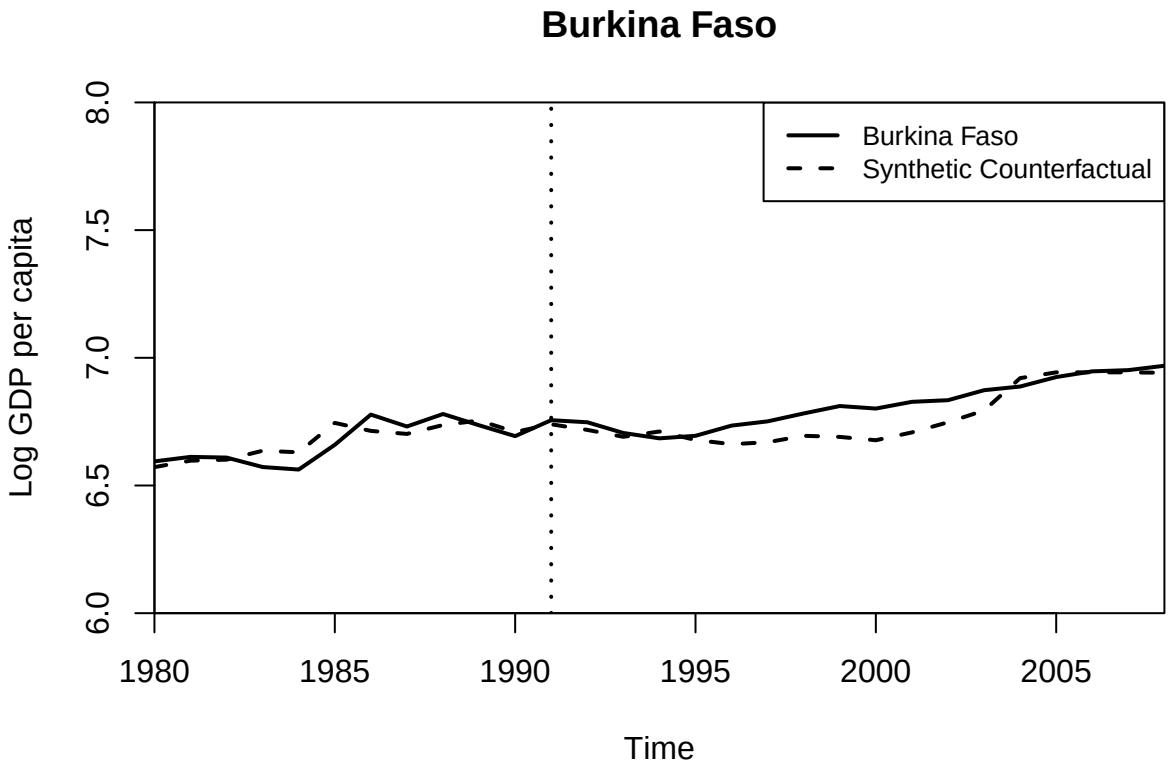
```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
```

```
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00088
##
## solution.v:
##  0.14 0.178 0.159 0.133 0.0884 0.00105 0.152 0.0021 0.00664 0.0194 0.0176 0.0165 0.0152 0.0025 0.009
##
## solution.w:
##  0.00000575 0.219 0.00292 0.0000537 0.0000306 0.0000303 0.000000535 0.000000384 0.00000939 0.0000176
```



```r
## Figure 2 Replication in-space placebo
replicate("BFA", "Burkina Faso", 1980, 1990, 1991, 2008, 6, 8)
```

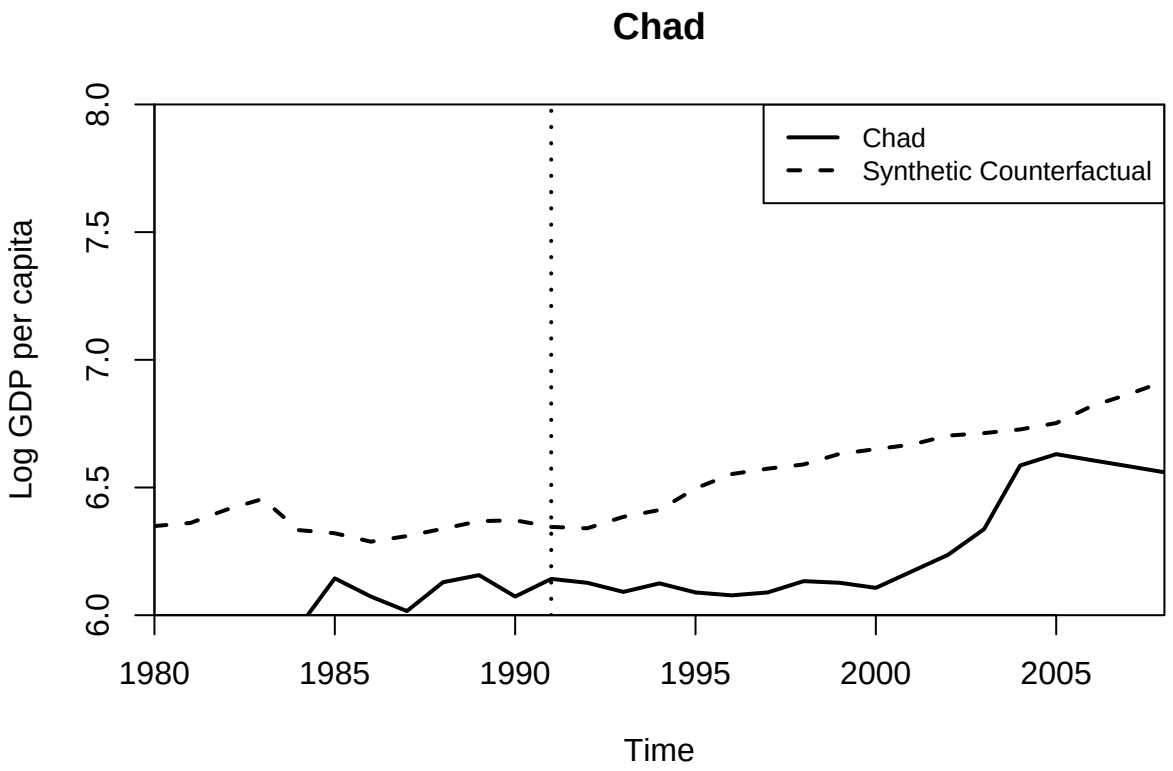
```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00204
##
## solution.v:
##  0.121 0.0888 0.0517 0.015 0.00501 0.388 0.015 0.00861 0.298 0.000331 0.0042 0.000772 0.0029 0.000018
##
## solution.w:
##  0.00000648 0.18 0.0000262 0.0000122 0.000018 0.0000118 0.00496 0.0000341 0.0000255 0.00262 0.000137
```




```
## Figure 2 Replication in-space placebo
replicate("TCD", "Chad", 1980, 1990, 1991, 2008, 6, 8)
```

```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
```

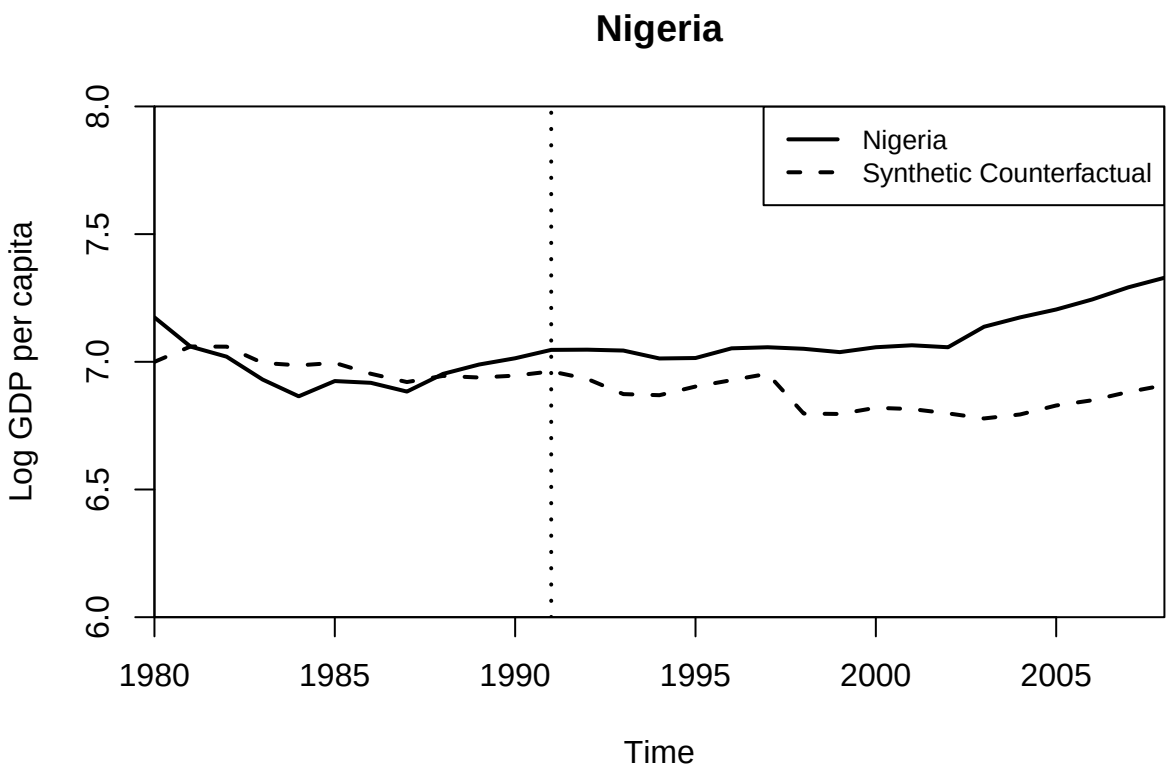
```
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.137
##
## solution.v:
##  0.195 0.196 0.185 0.163 0.00632 0.152 0.0817 0.000169 0.00515 0.00198 0.000662 0.00208 0.00694 0.00
##
## solution.w:
##  0.0000000374 0.000000468 0.000000213 0.000000295 0.0000000067 0.0000000067 2e-10 0.0000000194 0.0000
```




```
## Figure 2 Replication in-space placebo
replicate("NGA", "Nigeria", 1980, 1990, 1991, 2008, 6, 8)
```

```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
```

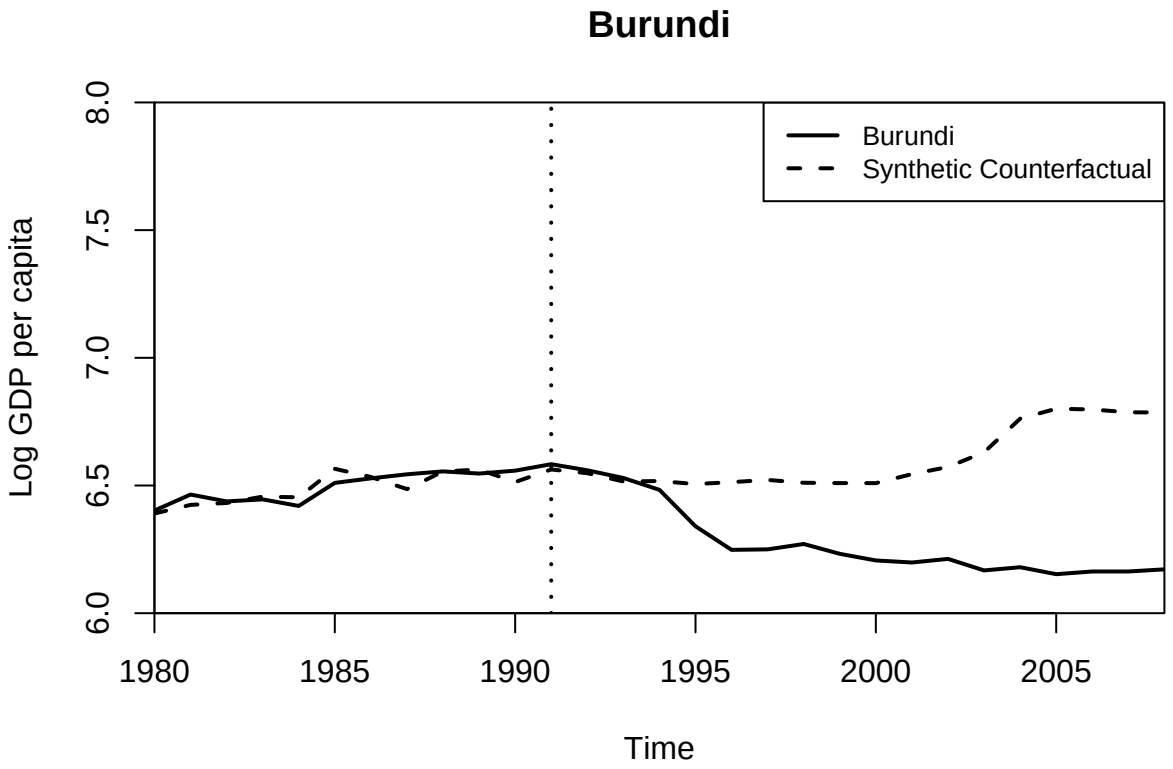
```
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00607
##
## solution.v:
##  0.000759 0.00277 0.00137 0.000229 0.000000116 0.000000127 0.000000245 0.0000000019 0.000000124 0.013
##
## solution.w:
##  0.137 0.0182 0.0172 0.0183 0.019 0.0138 0.137 0.467 0.024 0.0192 0.018 0.0182 0.00818 0.0176 0.0183
```




```
## Figure 2 Replication in-space placebo
replicate("BDI", "Burundi", 1980, 1990, 1991, 2008, 6, 8)
```

```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 14 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
```

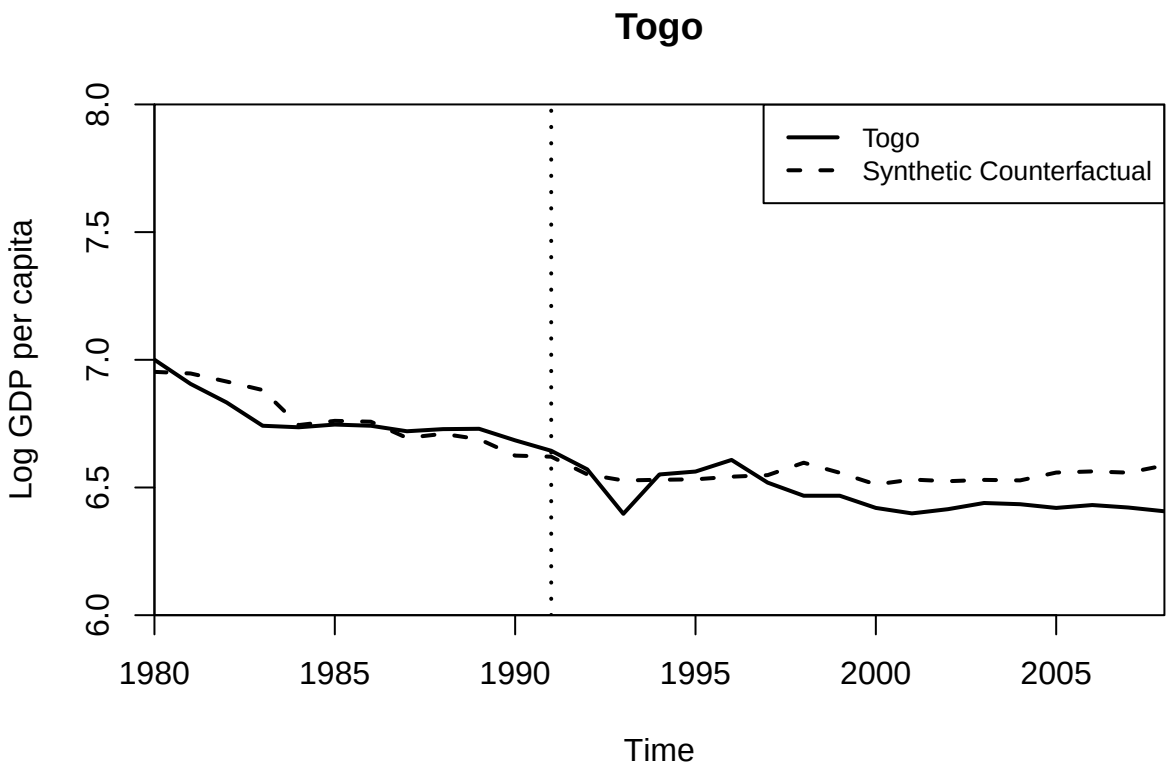
```
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00101
##
## solution.v:
##  0.0105 0.00106 0.000221 0.000231 0.0176 0.0309 0.0000000683 0.000000568 0.000000264 0.0141 0.0919 0
##
## solution.w:
##  0.00138 0.159 0.000637 0.00174 0.108 0.00158 0.114 0.00143 0.00186 0.0241 0.0707 0.00207 0.00213 0.5
```




```
## Figure 2 Replication in-space placebo
replicate("TGO", "Togo", 1980, 1990, 1991, 2008, 6, 8)
```

```
##
##  Missing data- treated unit; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data- treated unit; predictor: eximdiff_lag2 ; for period: 1981
```

```
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag1 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 2 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1980
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
##  Missing data - control unit: 11 ; predictor: eximdiff_lag2 ; for period: 1981
##  We ignore (na.rm = TRUE) all missing values for predictors.op.
##
## X1, X0, Z1, Z0 all come directly from dataprep object.
##
##
## ****************
##  searching for synthetic control unit
##
##
## ****************
## ****************
## ****************
##
## MSPE (LOSS V): 0.00313
##
## solution.v:
##  0.0239 0.00267 0.00904 0.00021 0.000000499 0.00000214 0.000000173 0.031 0.000000331 0.0126 0.0919 0
##
## solution.w:
##  0.000138 0.000551 0.0000373 0.0000122 0.245 0.00972 0.0536 0.000294 0.000279 0.000449 0.578 0.00028
```



Extensions

Google Extension (CausalImpact)

```
# Replication function from Google Extension
show_impact_n <- function(
  Country,
  begin,
  end,
  treatYear
){
  data <- afripanel[which(afripanel$Country == Country), ]
  predictors=c(
      "lngdpmadlag",
      "lngdpmadlag2",
      "lngdpmadlag3",
      "lngdpmadlag4",
      "lnpop",
      "ki",
      "openk",
      "civwar",
      "civwarend",
      "pwt_xrate",
      "pwt_xrate_lag1",
      "pwt_xrate_lag2",
      "pwt_xrate_lag3",
      "eximdiff",
      "eximdiff_lag1",
```

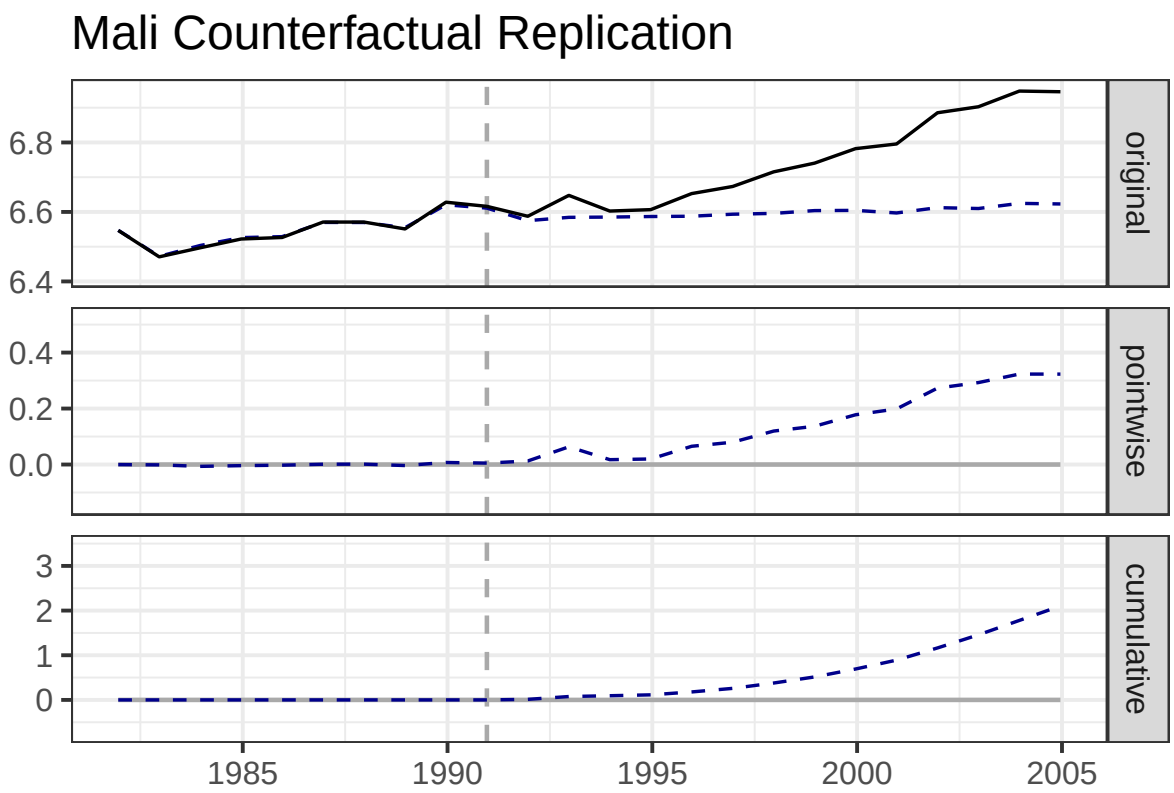
```
      "eximdiff_lag2",
      "wbank",
      "wbank_lag1",
      "wbank_lag2"
  )

  outcome <- 'lngdpmad'
  time.points <- as.Date(as.character(data$year), "%Y")

  data <- data[, c(outcome, predictors)]
  data<-data[!is.na(data[outcome]),]
  data <- data %>% select_if(not_any_na)


  data <- zoo(data, time.points)
  data <- data[index(data) > as.Date(begin, '%Y') & index(data) < as.Date(end, '%Y')]

  nextYear <- as.Date(as.character(as.numeric(treatYear) + 1), "%Y")
  treatYear <- as.Date(treatYear, "%Y")

  start_date <- start(data)
  end_date <- end(data)

  pre.period <- as.Date(c(start_date, treatYear))
  post.period <- as.Date(c(nextYear, end_date))

  impact <- CausalImpact(data,
                         pre.period,
                         post.period,
                         model.args = list(
                           niter = 1000,
                           nseasons = 52)
                         )
  return(impact)
}
```

Graphs:

```
impact <- show_impact_n('Mali', '1980', '2005', '1990')
plot(impact)+ggtitle("Mali Counterfactual Replication")
```

```
## Warning: Removed 24 rows containing missing values (geom_path).
```

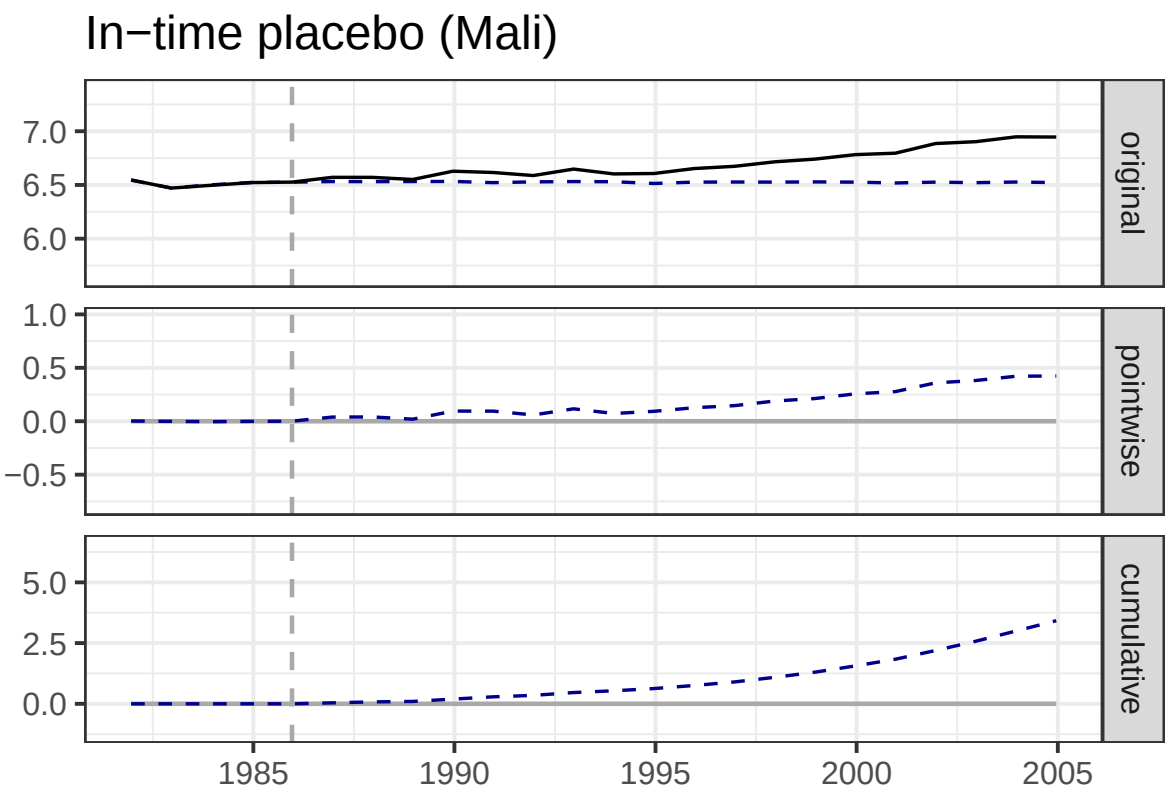
```
## Warning: Removed 48 rows containing missing values (geom_path).
```




```
impact <- show_impact_n('Mali', '1980', '2005', '1985')
plot(impact)+ggtitle("In-time placebo (Mali)")
```

```
## Warning: Removed 24 rows containing missing values (geom_path).
```

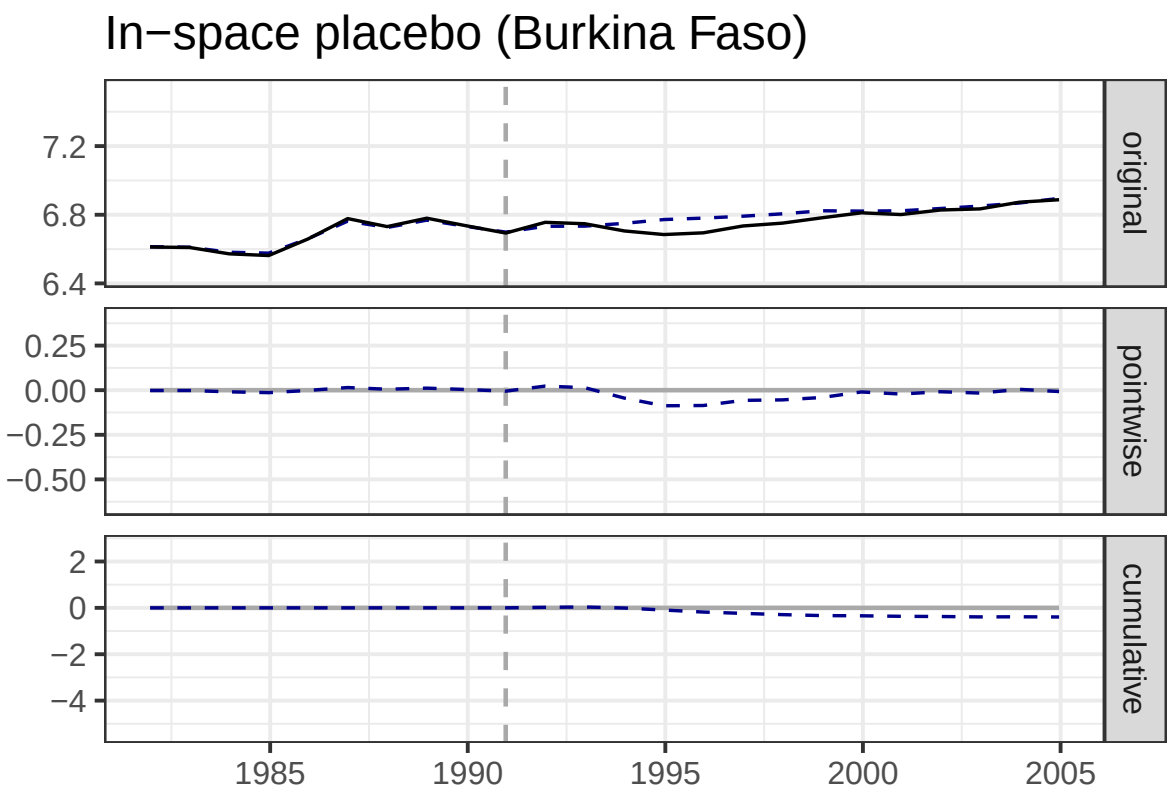
```
## Warning: Removed 48 rows containing missing values (geom_path).
```



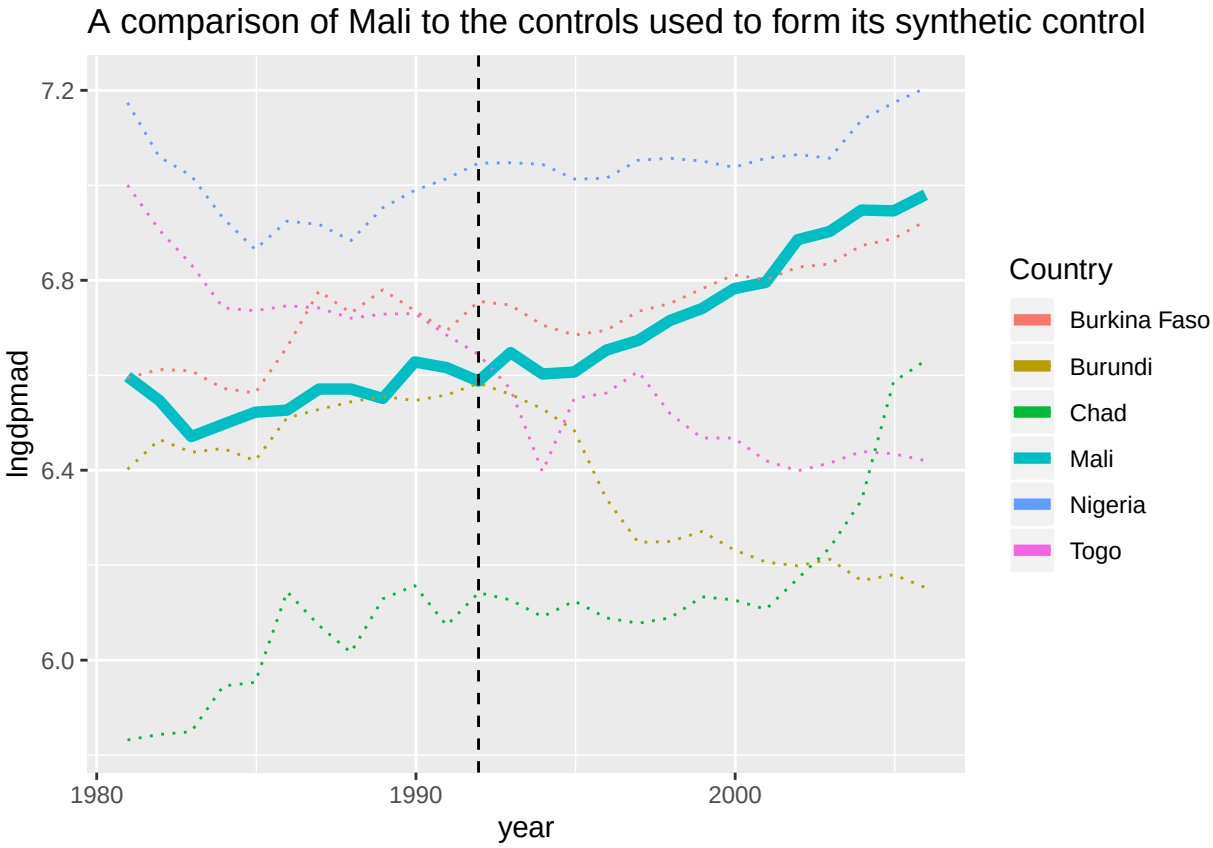

```
impact <- show_impact_n('Burkina Faso', '1980', '2005', '1990')
plot(impact)+ggtitle("In-space placebo (Burkina Faso)")
```

```
## Warning: Removed 24 rows containing missing values (geom_path).
```

```
## Warning: Removed 48 rows containing missing values (geom_path).
```




```
foo <- afripanel
foo$year <- as.Date(as.character(foo$year), '%Y')

foo <- foo[which(foo$year >= as.Date('1980', '%Y') & foo$year <= as.Date('2005', '%Y')),]

mali_data <- foo[which(foo$Country == 'Mali'), ]

controls <- c('Burkina Faso', 'Chad', 'Nigeria', 'Togo', 'Burundi')

control_data <- foo[foo$Country %in% controls,]

ggplot(mali_data, aes(year, lngdpmad, color=Country)) +
  geom_line(size = 2) +
  geom_line(data = control_data, linetype = 'dotted') +
  geom_vline(xintercept = as.Date('1991', '%Y'), linetype='dashed') +
  ggtitle('A comparison of Mali to the controls used to form its synthetic control')
```


A comparison of Mali to the controls used to form its synthetic control
7.2
6.8
6.4
6.0
lngdpmad
1980
1990
2000
year
Country
Burkina Faso
Burundi
Chad
Mali
Nigeria
Togo