

Package ‘did’

December 11, 2020

Title Treatment Effects with Multiple Periods and Groups

Version 2.0.0

URL <https://bcallaway11.github.io/did/>,
<https://github.com/bcallaway11/did/>

Description The standard Difference-in-Differences (DID) setup involves two periods and two groups -- a treated group and untreated group. Many applications of DID methods involve more than two periods and have individuals that are treated at different points in time. This package contains tools for computing average treatment effect parameters in Difference in Differences setups with more than two periods and with variation in treatment timing using the methods developed in Callaway and Sant'Anna (2020) <<https://www.ssrn.com/abstract=3148250>>. The main parameters are group-time average treatment effects which are the average treatment effect for a particular group at a particular time. These can be aggregated into a fewer number of treatment effect parameters, and the package deals with the cases where there is selective treatment timing, dynamic treatment effects, calendar time effects, or combinations of these. There are also functions for testing the Difference in Differences assumption, and plotting group-time average treatment effects.

Depends R (>= 3.5),

License GPL-2

Encoding UTF-8

LazyData true

Imports BMisc (>= 1.4.1), Matrix, pbapply, ggplot2, ggpubr, DRDID,
knitr

RoxygenNote 7.1.1

VignetteBuilder knitr

Suggests rmarkdown, plm, here

NeedsCompilation no

Author Brantly Callaway [aut, cre],
Pedro H. C. Sant'Anna [aut]

Maintainer Brantly Callaway <brantly.callaway@uga.edu>

Repository CRAN

Date/Publication 2020-12-11 22:40:02 UTC

R topics documented:

| | |
|-------------------------|-----------|
| aggte | 2 |
| AGGTEobj | 4 |
| att_gt | 5 |
| conditional_did_pretest | 8 |
| did | 11 |
| DIDparams | 11 |
| gdid | 13 |
| gdid.AGGTEobj | 14 |
| gdid.MP | 15 |
| indicator | 16 |
| mboot | 16 |
| MP | 17 |
| mp.spatt | 18 |
| mp.spatt.test | 18 |
| MP.TEST | 19 |
| mpdta | 20 |
| pre_process_did | 20 |
| print.MP | 23 |
| process_attgt | 23 |
| summary.AGGTEobj | 24 |
| summary.MP | 24 |
| summary.MP.TEST | 25 |
| test.mboot | 25 |
| Index | 26 |

aggte

Aggregate Group-Time Average Treatment Effects

Description

A function to take group-time average treatment effects and aggregate them into a smaller number of parameters. There are several possible aggregations including "simple", "dynamic", "group", and "calendar."

Usage

```
aggte(
  MP,
  type = "group",
  balance_e = NULL,
  min_e = -Inf,
  max_e = Inf,
  na.rm = FALSE,
  bstrap = NULL,
  biters = NULL,
```

```

    cband = NULL,
    alp = NULL,
    clustervars = NULL
)

```

Arguments

| | |
|-------------|--|
| MP | an MP object (i.e., the results of the <code>att_gt</code> method) |
| type | Which type of aggregated treatment effect parameter to compute. One option is "simple" (this just computes a weighted average of all group-time average treatment effects with weights proportional to group size). Other options are "dynamic" (this computes average effects across different lengths of exposure to the treatment and is similar to an "event study"; here the overall effect averages the effect of the treatment across all positive lengths of exposure); "group" (this is the default option and computes average treatment effects across different groups; here the overall effect averages the effect across different groups); and "calendar" (this computes average treatment effects across different time periods; here the overall effect averages the effect across each time period). |
| balance_e | If set (and if one computes dynamic effects), it balances the sample with respect to event time. For example, if <code>balance_e=2</code> , <code>aggte</code> will drop groups that are not exposed to treatment for at least three periods. (the initial period when <code>e=0</code> as well as the next two periods when <code>e=1</code> and the <code>e=2</code>). This ensures that the composition of groups does not change when event time changes. |
| min_e | For event studies, this is the smallest event time to compute dynamic effects for. By default, <code>min_e = -Inf</code> so that effects at all lengths of exposure are computed. |
| max_e | For event studies, this is the largest event time to compute dynamic effects for. By default, <code>max_e = Inf</code> so that effects at all lengths of exposure are computed. |
| na.rm | Logical value if we are to remove missing Values from analyses. Defaults is FALSE. |
| bstrap | Boolean for whether or not to compute standard errors using the multiplier bootstrap. If standard errors are clustered, then one must set <code>bstrap=TRUE</code> . Default is value set in the MP object. If <code>bstrap</code> is FALSE, then analytical standard errors are reported. |
| biters | The number of bootstrap iterations to use. The default is the value set in the MP object, and this is only applicable if <code>bstrap=TRUE</code> . |
| cband | Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability $1-\alpha$. In order to compute uniform confidence bands, <code>bstrap</code> must also be set to TRUE. The default is the value set in the MP object |
| alp | the significance level, default is value set in the MP object. |
| clustervars | A vector of variables to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as <code>idname</code> which allows for clustering at the individual level. Default is the variables set in the MP object |

Value

An [AGGTEobj](#) object that holds the results from the aggregation

AGGTEobj

AGGTEobj

Description

Objects of this class hold results on aggregated group-time average treatment effects

An object for holding aggregated treatment effect parameters.

Usage

```
AGGTEobj(
  overall.att = NULL,
  overall.se = NULL,
  type = "simple",
  egt = NULL,
  att.egt = NULL,
  se.egt = NULL,
  crit.val.egt = NULL,
  inf.function = NULL,
  min_e = NULL,
  max_e = NULL,
  balance_e = NULL,
  call = NULL,
  DIDparams = NULL
)
```

Arguments

| | |
|--------------------------|--|
| <code>overall.att</code> | The estimated overall ATT |
| <code>overall.se</code> | Standard error for overall ATT |
| <code>type</code> | Which type of aggregated treatment effect parameter to compute. One option is "simple" (this just computes a weighted average of all group-time average treatment effects with weights proportional to group size). Other options are "dynamic" (this computes average effects across different lengths of exposure to the treatment and is similar to an "event study"; here the overall effect averages the effect of the treatment across all positive lengths of exposure); "group" (this is the default option and computes average treatment effects across different groups; here the overall effect averages the effect across different groups); and "calendar" (this computes average treatment effects across different time periods; here the overall effect averages the effect across each time period). |
| <code>egt</code> | Holds the length of exposure (for dynamic effects), the group (for selective treatment timing), or the time period (for calendar time effects) |

| | |
|--------------|---|
| att.egt | The ATT specific to egt |
| se.egt | The standard error specific to egt |
| crit.val.egt | A critical value for computing uniform confidence bands for dynamic effects, selective treatment timing, or time period effects. |
| inf.function | The influence function of the chosen aggregated parameters |
| min_e | For event studies, this is the smallest event time to compute dynamic effects for. By default, min_e = -Inf so that effects at all lengths of exposure are computed. |
| max_e | For event studies, this is the largest event time to compute dynamic effects for. By default, max_e = Inf so that effects at all lengths of exposure are computed. |
| balance_e | If set (and if one computes dynamic effects), it balances the sample with respect to event time. For example, if balance.e=2, aggte will drop groups that are not exposed to treatment for at least three periods. (the initial period when e=0 as well as the next two periods when e=1 and the e=2). This ensures that the composition of groups does not change when event time changes. |
| call | The function call to aggte |
| DIDparams | A DIDparams object |

Value

an AGGTEobj

att_gt *Group-Time Average Treatment Effects*

Description

att_gt computes average treatment effects in DID setups where there are more than two periods of data and allowing for treatment to occur at different points in time and allowing for treatment effect heterogeneity and dynamics. See Callaway and Sant'Anna (2020) for a detailed description.

Usage

```
att_gt(
  yname,
  tname,
  idname = NULL,
  gname,
  xformula = NULL,
  data,
  panel = TRUE,
  allow_unbalanced_panel = FALSE,
  control_group = c("nevertreated", "notyettreated"),
  anticipation = 0,
  weightsname = NULL,
  alp = 0.05,
```

```

bstrap = TRUE,
cband = TRUE,
biters = 1000,
clustervars = NULL,
est_method = "dr",
print_details = FALSE,
pl = FALSE,
cores = 1
)

```

Arguments

| | |
|------------------------|---|
| yname | The name of the outcome variable |
| tname | The name of the column containing the time periods |
| idname | The individual (cross-sectional unit) id name |
| gname | The name of the variable in data that contains the first period when a particular observation is treated. This should be a positive number for all observations in treated groups. It defines which "group" a unit belongs to. It should be 0 for units in the untreated group. |
| xformula | A formula for the covariates to include in the model. It should be of the form $\sim X1 + X2$. Default is NULL which is equivalent to <code>xformula=~1</code> . This is used to create a matrix of covariates which is then passed to the 2x2 DID estimator chosen in <code>est_method</code> . |
| data | The name of the data.frame that contains the data |
| panel | Whether or not the data is a panel dataset. The panel dataset should be provided in long format – that is, where each row corresponds to a unit observed at a particular point in time. The default is TRUE. When using a panel dataset, the variable <code>idname</code> must be set. When <code>panel=FALSE</code> , the data is treated as repeated cross sections. |
| allow_unbalanced_panel | Whether or not function should "balance" the panel with respect to time and id. The default values if FALSE which means that <code>att_gt</code> will drop all units where data is not observed in all periods. The advantage of this is that the computations are faster (sometimes substantially). |
| control_group | Which units to use the control group. The default is "nevertreated" which sets the control group to be the group of units that never participate in the treatment. This group does not change across groups or time periods. The other option is to set <code>group="notyettreated"</code> . In this case, the control group is set to the group of units that have not yet participated in the treatment in that time period. This includes all never treated units, but it includes additional units that eventually participate in the treatment, but have not participated yet. |
| anticipation | The number of time periods before participating in the treatment where units can anticipate participating in the treatment and therefore it can affect their untreated potential outcomes |
| weightsname | The name of the column containing the sampling weights. If not set, all observations have same weight. |

| | |
|---------------|---|
| alp | the significance level, default is 0.05 |
| bstrap | Boolean for whether or not to compute standard errors using the multiplier bootstrap. If standard errors are clustered, then one must set bstrap=TRUE. Default is TRUE (in addition, cband is also by default TRUE indicating that uniform confidence bands will be returned. If bstrap is FALSE, then analytical standard errors are reported. |
| cband | Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability $1-\text{alp}$. In order to compute uniform confidence bands, bstrap must also be set to TRUE. The default is TRUE. |
| biters | The number of bootstrap iterations to use. The default is 1000, and this is only applicable if bstrap=TRUE. |
| clustervars | A vector of variables to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as idname which allows for clustering at the individual level. |
| est_method | the method to compute group-time average treatment effects. The default is "dr" which uses the doubly robust approach in the DRDID package. Other built-in methods include "ipw" for inverse probability weighting and "reg" for first step regression estimators. The user can also pass their own function for estimating group time average treatment effects. This should be a function $f(Y1, Y0, \text{treat}, \text{covariates})$ where $Y1$ is an $n \times 1$ vector of outcomes in the post-treatment outcomes, $Y0$ is an $n \times 1$ vector of pre-treatment outcomes, treat is a vector indicating whether or not an individual participates in the treatment, and covariates is an $n \times k$ matrix of covariates. The function should return a list that includes ATT (an estimated average treatment effect), and inf.func (an $n \times 1$ influence function). The function can return other things as well, but these are the only two that are required. <code>est_method</code> is only used if <code>covariates</code> are included. |
| print_details | Whether or not to show details/progress of computations. Default is FALSE. |
| pl | Whether or not to use parallel processing (not implemented yet) |
| cores | The number of cores to use for parallel processing (not implemented yet) |

Value

an MP object containing all the results for group-time average treatment effects

References

Callaway, Brantly and Sant'Anna, Pedro H. C. "Difference-in-Differences with Multiple Time Periods" Forthcoming at the Journal of Econometrics <<https://arxiv.org/abs/1803.09015>> (2020).

Examples

```
data(mpdt)

# with covariates
out1 <- att_gt(yname="lemp",
              tname="year",
```

```

        idname="countyreal",
        gname="first.treat",
        xformula=~lpop,
        data=mpdta)
summary(out1)

# without covariates
out2 <- att_gt(yname="lemp",
               tname="year",
               idname="countyreal",
               gname="first.treat",
               xformula=NULL,
               data=mpdta)
summary(out2)

```

conditional_did_pretest

Pre-Test of Conditional Parallel Trends Assumption

Description

An integrated moments test for the conditional parallel trends assumption holding in all pre-treatment time periods for all groups

Usage

```

conditional_did_pretest(
  yname,
  tname,
  idname = NULL,
  gname,
  xformula = NULL,
  data,
  panel = TRUE,
  allow_unbalanced_panel = FALSE,
  control_group = c("nevertreated", "notyettreated"),
  weightsname = NULL,
  alp = 0.05,
  bstrap = TRUE,
  cband = TRUE,
  biters = 1000,
  clustervars = NULL,
  est_method = "ipw",
  print_details = FALSE,
  pl = FALSE,
  cores = 1
)

```


Arguments

| | |
|------------------------|---|
| yname | The name of the outcome variable |
| tname | The name of the column containing the time periods |
| idname | The individual (cross-sectional unit) id name |
| gname | The name of the variable in data that contains the first period when a particular observation is treated. This should be a positive number for all observations in treated groups. It defines which "group" a unit belongs to. It should be 0 for units in the untreated group. |
| xformula | A formula for the covariates to include in the model. It should be of the form $\sim X1 + X2$. Default is NULL which is equivalent to $xformula \sim 1$. This is used to create a matrix of covariates which is then passed to the 2x2 DID estimator chosen in <code>est_method</code> . |
| data | The name of the data.frame that contains the data |
| panel | Whether or not the data is a panel dataset. The panel dataset should be provided in long format – that is, where each row corresponds to a unit observed at a particular point in time. The default is TRUE. When using a panel dataset, the variable <code>idname</code> must be set. When <code>panel=FALSE</code> , the data is treated as repeated cross sections. |
| allow_unbalanced_panel | Whether or not function should "balance" the panel with respect to time and id. The default values if FALSE which means that <code>att_gt</code> will drop all units where data is not observed in all periods. The advantage of this is that the computations are faster (sometimes substantially). |
| control_group | Which units to use the control group. The default is "nevertreated" which sets the control group to be the group of units that never participate in the treatment. This group does not change across groups or time periods. The other option is to set <code>group="notyettreated"</code> . In this case, the control group is set to the group of units that have not yet participated in the treatment in that time period. This includes all never treated units, but it includes additional units that eventually participate in the treatment, but have not participated yet. |
| weightsname | The name of the column containing the sampling weights. If not set, all observations have same weight. |
| alp | the significance level, default is 0.05 |
| bstrap | Boolean for whether or not to compute standard errors using the multiplier bootstrap. If standard errors are clustered, then one must set <code>bstrap=TRUE</code> . Default is TRUE (in addition, <code>cband</code> is also by default TRUE indicating that uniform confidence bands will be returned. If <code>bstrap</code> is FALSE, then analytical standard errors are reported. |
| cband | Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability $1-\alpha p$. In order to compute uniform confidence bands, <code>bstrap</code> must also be set to TRUE. The default is TRUE. |
| biters | The number of bootstrap iterations to use. The default is 1000, and this is only applicable if <code>bstrap=TRUE</code> . |

| | |
|----------------------------|---|
| <code>clustervars</code> | A vector of variables to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as <code>idname</code> which allows for clustering at the individual level. |
| <code>est_method</code> | the method to compute group-time average treatment effects. The default is "dr" which uses the doubly robust approach in the DRDID package. Other built-in methods include "ipw" for inverse probability weighting and "reg" for first step regression estimators. The user can also pass their own function for estimating group time average treatment effects. This should be a function $f(Y1, Y0, treat, covariates)$ where $Y1$ is an $n \times 1$ vector of outcomes in the post-treatment outcomes, $Y0$ is an $n \times 1$ vector of pre-treatment outcomes, <code>treat</code> is a vector indicating whether or not an individual participates in the treatment, and <code>covariates</code> is an $n \times k$ matrix of covariates. The function should return a list that includes ATT (an estimated average treatment effect), and <code>inf.func</code> (an $n \times 1$ influence function). The function can return other things as well, but these are the only two that are required. <code>est_method</code> is only used if <code>covariates</code> are included. |
| <code>print_details</code> | Whether or not to show details/progress of computations. Default is FALSE. |
| <code>pl</code> | Whether or not to use parallel processing (not implemented yet) |
| <code>cores</code> | The number of cores to use for parallel processing (not implemented yet) |

Value

an `MP.TEST` object

References

Callaway, Brantly and Sant'Anna, Pedro H. C. "Difference-in-Differences with Multiple Time Periods and an Application on the Minimum Wage and Employment." Working Paper <<https://arxiv.org/abs/1803.09015v2>> (2018).

Examples

```
## Not run:
data(mpdt)
pre.test <- conditional_did_pretest(yname="lemp",
                                   tname="year",
                                   idname="countyreal",
                                   gname="first.treat",
                                   xformula=~lpop,
                                   data=mpdt)

summary(pre.test)

## End(Not run)
```

| | |
|-----|----------------------------------|
| did | <i>Difference in Differences</i> |
|-----|----------------------------------|

Description

Difference in Differences with multiple periods and variation in treatment timing

| | |
|-----------|------------------|
| DIDparams | <i>DIDparams</i> |
|-----------|------------------|

Description

Object to hold did parameters that are passed across functions

Usage

```
DIDparams(
  yname,
  tname,
  idname = NULL,
  gname,
  xformula = NULL,
  data,
  control_group,
  anticipation = 0,
  weightsname = NULL,
  alp = 0.05,
  bstrap = T,
  biters = 1000,
  clustervars = NULL,
  cband = T,
  print_details = TRUE,
  pl = FALSE,
  cores = 1,
  est_method = "dr",
  panel = TRUE,
  true_repeated_cross_sections,
  n = NULL,
  nG = NULL,
  nT = NULL,
  tlist = NULL,
  glist = NULL,
  call = NULL
)
```

Arguments

| | |
|---------------|---|
| yname | The name of the outcome variable |
| tname | The name of the column containing the time periods |
| idname | The individual (cross-sectional unit) id name |
| gname | The name of the variable in data that contains the first period when a particular observation is treated. This should be a positive number for all observations in treated groups. It defines which "group" a unit belongs to. It should be 0 for units in the untreated group. |
| xformula | A formula for the covariates to include in the model. It should be of the form $\sim X1 + X2$. Default is NULL which is equivalent to <code>xformula=~1</code> . This is used to create a matrix of covariates which is then passed to the 2x2 DID estimator chosen in <code>est_method</code> . |
| data | The name of the data.frame that contains the data |
| control_group | Which units to use the control group. The default is "nevertreated" which sets the control group to be the group of units that never participate in the treatment. This group does not change across groups or time periods. The other option is to set <code>group="notyettreated"</code> . In this case, the control group is set to the group of units that have not yet participated in the treatment in that time period. This includes all never treated units, but it includes additional units that eventually participate in the treatment, but have not participated yet. |
| anticipation | The number of time periods before participating in the treatment where units can anticipate participating in the treatment and therefore it can affect their untreated potential outcomes |
| weightsname | The name of the column containing the sampling weights. If not set, all observations have same weight. |
| alp | the significance level, default is 0.05 |
| bstrap | Boolean for whether or not to compute standard errors using the multiplier bootstrap. If standard errors are clustered, then one must set <code>bstrap=TRUE</code> . Default is TRUE (in addition, <code>cband</code> is also by default TRUE indicating that uniform confidence bands will be returned. If <code>bstrap</code> is FALSE, then analytical standard errors are reported. |
| biters | The number of bootstrap iterations to use. The default is 1000, and this is only applicable if <code>bstrap=TRUE</code> . |
| clustervars | A vector of variables to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as <code>idname</code> which allows for clustering at the individual level. |
| cband | Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability $1-\alpha$. In order to compute uniform confidence bands, <code>bstrap</code> must also be set to TRUE. The default is TRUE. |
| print_details | Whether or not to show details/progress of computations. Default is FALSE. |
| pl | Whether or not to use parallel processing (not implemented yet) |
| cores | The number of cores to use for parallel processing (not implemented yet) |

| | |
|---|--|
| <code>est_method</code> | the method to compute group-time average treatment effects. The default is "dr" which uses the doubly robust approach in the DRDID package. Other built-in methods include "ipw" for inverse probability weighting and "reg" for first step regression estimators. The user can also pass their own function for estimating group time average treatment effects. This should be a function $f(Y1, Y0, treat, covariates)$ where $Y1$ is an $n \times 1$ vector of outcomes in the post-treatment outcomes, $Y0$ is an $n \times 1$ vector of pre-treatment outcomes, $treat$ is a vector indicating whether or not an individual participates in the treatment, and $covariates$ is an $n \times k$ matrix of covariates. The function should return a list that includes <code>ATT</code> (an estimated average treatment effect), and <code>inf.func</code> (an $n \times 1$ influence function). The function can return other things as well, but these are the only two that are required. <code>est_method</code> is only used if <code>covariates</code> are included. |
| <code>panel</code> | Whether or not the data is a panel dataset. The panel dataset should be provided in long format – that is, where each row corresponds to a unit observed at a particular point in time. The default is <code>TRUE</code> . When is using a panel dataset, the variable <code>idname</code> must be set. When <code>panel=FALSE</code> , the data is treated as repeated cross sections. |
| <code>true_repeated_cross_sections</code> | Whether or not the data really is repeated cross sections. (We include this because unbalanced panel code runs through the repeated cross sections code) |
| <code>n</code> | The number of observations. This is equal to the number of units (which may be different from the number of rows in a panel dataset). |
| <code>nG</code> | The number of groups |
| <code>nT</code> | The number of time periods |
| <code>tlist</code> | a vector containing each time period |
| <code>glist</code> | a vector containing each group |
| <code>call</code> | Function call to <code>att_gt</code> |

ggdid

Plot did objects using ggplot2

Description

Function to plot objects from the did package

Usage

```
ggdid(object, ...)
```

Arguments

| | |
|---------------------|---------------------------------------|
| <code>object</code> | either a MP object or AGGTEobj object |
| <code>...</code> | other arguments |

ggdid.AGGTEobj *Plot AGGTEobj objects*

Description

A function to plot AGGTEobj objects

Usage

```
## S3 method for class 'AGGTEobj'  
ggdid(  
  object,  
  ylim = NULL,  
  xlab = NULL,  
  ylab = NULL,  
  title = "",  
  xgap = 1,  
  legend = TRUE,  
  ...  
)
```

Arguments

| | |
|--------|--|
| object | either a MP object or AGGTEobj object |
| ylim | optional y limits for the plot; setting here makes the y limits the same across different plots |
| xlab | optional x-axis label |
| ylab | optional y-axis label |
| title | optional plot title |
| xgap | optional gap between the labels on the x-axis. For example, xgap=3 indicates that the labels should show up for every third value on the x-axis. The default is 1. |
| legend | Whether or not to include a legend (which will indicate color of pre- and post-treatment estimates). Default is TRUE. |
| ... | other arguments |

`ggdid.MP`*Plot MP objects using ggplot2*

Description

A function to plot MP objects

Usage

```
## S3 method for class 'MP'  
ggdid(  
  object,  
  ylim = NULL,  
  xlab = NULL,  
  ylab = NULL,  
  title = "Group",  
  xgap = 1,  
  ncol = 1,  
  legend = TRUE,  
  ...  
)
```

Arguments

| | |
|---------------------|---|
| <code>object</code> | either a MP object or AGGTEobj object |
| <code>ylim</code> | optional y limits for the plot; setting here makes the y limits the same across different plots |
| <code>xlab</code> | optional x-axis label |
| <code>ylab</code> | optional y-axis label |
| <code>title</code> | optional plot title |
| <code>xgap</code> | optional gap between the labels on the x-axis. For example, <code>xgap=3</code> indicates that the labels should show up for every third value on the x-axis. The default is 1. |
| <code>ncol</code> | The number of columns to include in the resulting plot. The default is 1. |
| <code>legend</code> | Whether or not to include a legend (which will indicate color of pre- and post-treatment estimates). Default is TRUE. |
| <code>...</code> | other arguments |

| | |
|-----------|------------------|
| indicator | <i>indicator</i> |
|-----------|------------------|

Description

indicator weighting function

Usage

```
indicator(X, u)
```

Arguments

| | |
|---|--------------------------------------|
| X | matrix of X's from the data |
| u | a particular value to compare X's to |

Value

numeric vector

Examples

```
data(mpdata)
dta <- subset(mpdata, year==2007)
X <- model.matrix(~lpop, data=dta)
X <- indicator(X, X[1,])
```

| | |
|-------|-----------------------------|
| mboot | <i>Multiplier Bootstrap</i> |
|-------|-----------------------------|

Description

A function to take an influence function and use the multiplier bootstrap to compute standard errors and critical values for uniform confidence bands.

Usage

```
mboot(inf.func, DIDparams)
```

Arguments

| | |
|-----------|-----------------------|
| inf.func | an influence function |
| DIDparams | DIDparams object |

Value

| | |
|----------|---|
| | list with elements |
| bres | results from each bootstrap iteration |
| V | variance matrix |
| se | standard errors |
| crit.val | a critical value for computing uniform confidence bands |

MP

MP

Description

Multi-period objects that hold results for group-time average treatment effects

Usage

```
MP(
  group,
  t,
  att,
  V_analytical,
  se,
  c,
  inffunc,
  n = NULL,
  W = NULL,
  Wpval = NULL,
  aggte = NULL,
  alp = 0.05,
  DIDparams = NULL
)
```

Arguments

| | |
|--------------|---|
| group | which group (defined by period first treated) an group-time average treatment effect is for |
| t | which time period a group-time average treatment effect is for |
| att | the group-average treatment effect for group group and time period t |
| V_analytical | Analytical estimator for the asymptotic variance-covariance matrix for group-time average treatment effects |
| se | standard errors for group-time average treatment effects. If bootstrap is set to TRUE, this provides bootstrap-based se. |
| c | simultaneous critical value if one is obtaining simultaneous confidence bands. Otherwise it reports the critical value based on pointwise normal approximation. |

| | |
|------------------------|---|
| <code>inffunc</code> | the influence function for estimating group-time average treatment effects |
| <code>n</code> | the number of observations |
| <code>W</code> | the Wald statistic for pre-testing the common trends assumption |
| <code>Wpval</code> | the p-value of the Wald statistic for pre-testing the common trends assumption |
| <code>aggte</code> | an aggregate treatment effects object |
| <code>alp</code> | the significance level, default is 0.05 |
| <code>DIDparams</code> | a <code>DIDparams</code> object. A way to optionally return the parameters of the call to <code>att_gt</code> or <code>conditional_did_pretest</code> . |

Value

MP object

| | |
|-----------------------|---|
| <code>mp.spatt</code> | <i>Group-Time Average Treatment Effects with Multiple Periods</i> |
|-----------------------|---|

Description

Deprecated function for computing group-time average treatment effects

Usage

`mp.spatt(...)`

Arguments

`...` extra arguments

| | |
|----------------------------|----------------------|
| <code>mp.spatt.test</code> | <i>mp.spatt.test</i> |
|----------------------------|----------------------|

Description

Deprecated function for integrated moments test for conditional parallel trends holding in all pre-treatment time periods across all groups

Usage

`mp.spatt.test(...)`

Arguments

`...` extra arguments

MP.TEST

MP.TEST

Description

An object that holds results from computing pre-test of the conditional parallel trends assumption

Usage

```
MP.TEST(
  CvM = NULL,
  CvMb = NULL,
  CvMcval = NULL,
  CvMpval = NULL,
  KS = NULL,
  KSb = NULL,
  KScval = NULL,
  KSpval = NULL,
  clustervars = NULL,
  xformula = NULL
)
```

Arguments

| | |
|-------------|---|
| CvM | Cramer von Mises test statistic |
| CvMb | a vector of bootstrapped Cramer von Mises test statistics |
| CvMcval | CvM critical value |
| CvMpval | p-value for CvM test |
| KS | Kolmogorov-Smirnov test statistic |
| KSb | a vector of bootstrapped KS test statistics |
| KScval | KS critical value |
| KSpval | p-value for KS test |
| clustervars | vector of which variables were clustered on for the test |
| xformula | formula for the X variables used in the test |

 mpdta

County Teen Employment Dataset

Description

A dataset containing (the log of) teen employment in 500 counties in the U.S. from 2004 to 2007. This is a subset of the dataset used in Callaway and Sant'Anna (2020). See that paper for additional descriptions.

Usage

mpdta

Format

A data frame with 2000 rows and 5 variables:

year the year of the observation

countyreal a unique identifier for a particular county

lpop the log of 1000s of population for the county

lemp the log of teen employment in the county

first.treat the year that the state where the county is located raised its minimum wage, it is set equal to 0 for counties that have minimum wages equal to the federal minimum wage over the entire period.

treat whether or not a particular county is treated in that year

Source

Callaway and Sant'Anna (2020)

 pre_process_did

Process did Function Arguments

Description

Function to process arguments passed to the main methods in the did package as well as conducting some tests to make sure data is in proper format / try to throw helpful error messages.

Usage

```
pre_process_did(
  yname,
  tname,
  idname,
  gname,
  xformula = NULL,
  data,
  panel = TRUE,
  allow_unbalanced_panel,
  control_group = c("nevertreated", "notyettreated"),
  anticipation = 0,
  weightsname = NULL,
  alp = 0.05,
  bstrap = FALSE,
  cband = FALSE,
  biters = 1000,
  clustervars = NULL,
  est_method = "dr",
  print_details = TRUE,
  pl = FALSE,
  cores = 1,
  call = NULL
)
```

Arguments

| | |
|------------------------|--|
| yname | The name of the outcome variable |
| tname | The name of the column containing the time periods |
| idname | The individual (cross-sectional unit) id name |
| gname | The name of the variable in data that contains the first period when a particular observation is treated. This should be a positive number for all observations in treated groups. It defines which "group" a unit belongs to. It should be 0 for units in the untreated group. |
| xformula | A formula for the covariates to include in the model. It should be of the form $\sim X1 + X2$. Default is NULL which is equivalent to $xformula \sim 1$. This is used to create a matrix of covariates which is then passed to the 2x2 DID estimator chosen in <code>est_method</code> . |
| data | The name of the data.frame that contains the data |
| panel | Whether or not the data is a panel dataset. The panel dataset should be provided in long format – that is, where each row corresponds to a unit observed at a particular point in time. The default is TRUE. When using a panel dataset, the variable <code>idname</code> must be set. When <code>panel=FALSE</code> , the data is treated as repeated cross sections. |
| allow_unbalanced_panel | Whether or not function should "balance" the panel with respect to time and id. The default values if FALSE which means that <code>att_gt</code> will drop all units where |

data is not observed in all periods. The advantage of this is that the computations are faster (sometimes substantially).

| | |
|---------------|--|
| control_group | Which units to use the control group. The default is "nevertreated" which sets the control group to be the group of units that never participate in the treatment. This group does not change across groups or time periods. The other option is to set group="notyettreated". In this case, the control group is set to the group of units that have not yet participated in the treatment in that time period. This includes all never treated units, but it includes additional units that eventually participate in the treatment, but have not participated yet. |
| anticipation | The number of time periods before participating in the treatment where units can anticipate participating in the treatment and therefore it can affect their untreated potential outcomes |
| weightsname | The name of the column containing the sampling weights. If not set, all observations have same weight. |
| alp | the significance level, default is 0.05 |
| bstrap | Boolean for whether or not to compute standard errors using the multiplier bootstrap. If standard errors are clustered, then one must set bstrap=TRUE. Default is TRUE (in addition, cband is also by default TRUE indicating that uniform confidence bands will be returned. If bstrap is FALSE, then analytical standard errors are reported. |
| cband | Boolean for whether or not to compute a uniform confidence band that covers all of the group-time average treatment effects with fixed probability $1-\alpha$. In order to compute uniform confidence bands, bstrap must also be set to TRUE. The default is TRUE. |
| biters | The number of bootstrap iterations to use. The default is 1000, and this is only applicable if bstrap=TRUE. |
| clustervars | A vector of variables to cluster on. At most, there can be two variables (otherwise will throw an error) and one of these must be the same as idname which allows for clustering at the individual level. |
| est_method | the method to compute group-time average treatment effects. The default is "dr" which uses the doubly robust approach in the DRDID package. Other built-in methods include "ipw" for inverse probability weighting and "reg" for first step regression estimators. The user can also pass their own function for estimating group time average treatment effects. This should be a function $f(Y1, Y0, treat, covariates)$ where $Y1$ is an $n \times 1$ vector of outcomes in the post-treatment outcomes, $Y0$ is an $n \times 1$ vector of pre-treatment outcomes, $treat$ is a vector indicating whether or not an individual participates in the treatment, and $covariates$ is an $n \times k$ matrix of covariates. The function should return a list that includes ATT (an estimated average treatment effect), and $inf.func$ (an $n \times 1$ influence function). The function can return other things as well, but these are the only two that are required. <code>est_method</code> is only used if <code>covariates</code> are included. |
| print_details | Whether or not to show details/progress of computations. Default is FALSE. |
| pl | Whether or not to use parallel processing (not implemented yet) |
| cores | The number of cores to use for parallel processing (not implemented yet) |
| call | Function call to <code>att_gt</code> |

Value

a [DIDparams](#) object

| | |
|-----------------------|-----------------|
| <code>print.MP</code> | <i>print.MP</i> |
|-----------------------|-----------------|

Description

prints value of a MP object

Usage

```
## S3 method for class 'MP'
print(x, ...)
```

Arguments

| | |
|------------------|-----------------|
| <code>x</code> | a MP object |
| <code>...</code> | extra arguments |

| | |
|----------------------------|--|
| <code>process_attgt</code> | <i>Process Results from compute.att_gt</i> |
|----------------------------|--|

Description

Process Results from [compute.att_gt](#)

Usage

```
process_attgt(attgt.results.list)
```

Arguments

| | |
|---------------------------------|---|
| <code>attgt.results.list</code> | list of results from compute.att_gt |
|---------------------------------|---|

Value

list with elements:

| | |
|-----------------------|---|
| <code>group</code> | which group a set of results belongs to |
| <code>tt</code> | which time period a set of results belongs to |
| <code>att</code> | the group time average treatment effect |
| <code>inf.func</code> | the influence function for that group time average treatment effect |

summary.AGGTEobj *Summary Aggregate Treatment Effect Parameter Objects*

Description

A function to summarize aggregated treatment effect parameters.

Usage

```
## S3 method for class 'AGGTEobj'  
summary(object, ...)
```

Arguments

| | |
|--------|--------------------|
| object | an AGGTEobj object |
| ... | other arguments |

summary.MP *summary.MP*

Description

prints a summary of a MP object

Usage

```
## S3 method for class 'MP'  
summary(object, ...)
```

Arguments

| | |
|--------|-----------------|
| object | an MP object |
| ... | extra arguments |

| | |
|-----------------|------------------------|
| summary.MP.TEST | <i>summary.MP.TEST</i> |
|-----------------|------------------------|

Description

print a summary of test results

Usage

```
## S3 method for class 'MP.TEST'
summary(object, ...)
```

Arguments

| | |
|--------|-------------------|
| object | an MP.TEST object |
| ... | other variables |

| | |
|------------|---|
| test.mboot | <i>Multiplier Bootstrap for Conditional Moment Test</i> |
|------------|---|

Description

A slightly modified multiplier bootstrap procedure for the pre-test of the conditional parallel trends assumption

Usage

```
test.mboot(inf.func, DIDparams, cores = 1)
```

Arguments

| | |
|-----------|---|
| inf.func | an influence function |
| DIDparams | DIDparams object |
| cores | The number of cores to use to bootstrap the test statistic in parallel. Default is cores=1 which corresponds to not running parallel. |

Value

| | |
|----------|--|
| list | |
| bres | CvM test statistics for each bootstrap iteration |
| crit.val | critical value for CvM test statistic |

Index

* datasets

mpdta, 20

aggte, 2

AGGTEobj, 4, 4

att_gt, 5, 18

compute.att_gt, 23

conditional_did_pretest, 8, 18

did, 11

DIDparams, 11, 18, 23

ggdid, 13

ggdid.AGGTEobj, 14

ggdid.MP, 15

indicator, 16

mboot, 16

MP, 7, 17

mp.spatt, 18

mp.spatt.test, 18

MP.TEST, 10, 19

mpdta, 20

pre_process_did, 20

print.MP, 23

process_attgt, 23

summary.AGGTEobj, 24

summary.MP, 24

summary.MP.TEST, 25

test.mboot, 25