Package 'robomit'

July 23, 2025

Version 1.0.7

Date 2025-01-12

Title Robustness Checks for Omitted Variable Bias

Maintainer Sergei Schaub <sergei.schaub@agroscope.admin.ch>

Description Robustness checks for omitted variable bias. The package includes robustness checks proposed by Oster (2019). The 'robomit' package computes i) the bias-adjusted treatment correlation or effect and ii) the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result based on the framework by Oster (2019). The code is based on the 'psacalc' command in 'Stata'. Additionally, 'robomit' offers a set of sensitivity analysis and visualization functions. See Oster, E. 2019. <doi:10.1080/07350015.2016.1227711>. Additionally, see Diegert, P., Masten, M. A., & Poirier, A. (2022) for a recent discussion of the topic: <doi:10.48550/arXiv.2206.02303>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.2

Imports plm, dplyr, ggplot2, broom, tidyr, tibble, stats,

Suggests testthat

NeedsCompilation no

Author Sergei Schaub [aut, cre] (ORCID:
<https://orcid.org/0000-0001-8477-3737>),
ETH Zurich [cph]

Repository CRAN

Date/Publication 2025-01-31 17:20:02 UTC

Contents

---

o_beta   *beta**

Description

Estimates beta*, i.e., the bias-adjusted treatment effect (or correlation) (following Oster 2019). The code is based on the psacalc command in Stata.

Usage

```
o_beta(y, x, con, w = NULL, id = "none", time = "none", delta = 1,
R2max, type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta* should be estimated (default is delta = 1).
R2max	Maximum R-square for which beta* should be estimated.
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates beta*, i.e., the bias-adjusted treatment effect (or correlation).

Value

Returns tibble object, which includes beta* and various other information.

References

Oster, E. (2019) Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate beta*
o_beta(y = "mpg",          # dependent variable
       x = "wt",           # independent treatment variable
       con = "hp + qsec",  # related control variables
       delta = 1,          # delta
       R2max = 0.9,        # maximum R-square
       type = "lm",        # model type
       data = data_oster)  # dataset
```

---

o_beta_boot   *Bootstrapped beta*s*

Description

Estimates bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019).

Usage

```
o_beta_boot(y, x, con, w = NULL, id = "none", time = "none", delta = 1,
R2max, sim, obs, rep, type, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".

w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta*s should be estimated (default is delta = 1).
R2max	Maximum R-square for which beta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement.
type	Model type (either *lm* or *plm*; as string).
useed	User defined seed.
data	Dataset.

Details

Estimates bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes bootstrapped beta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate bootstrapped beta*s
o_beta_boot(y = "mpg",          # dependent variable
            x = "wt",           # independent treatment variable
            con = "hp + qsec",  # related control variables
```

```
          delta = 1,           # delta
          R2max = 0.9,         # maximum R-square
          sim = 100,           # number of simulations
          obs = 30,            # draws per simulation
          rep = FALSE,         # bootstrapping with or without replacement
          type = "lm",         # model type
          useed = 123,         # seed
          data = data_oster)   # dataset
```

---

o_beta_boot_inf	*Bootstrapped mean beta* and confidence intervals*

Description

Provides the mean and confidence intervals of estimated bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019).

Usage

```
o_beta_boot_inf(y, x, con, w = NULL, id = "none", time = "none",
delta = 1, R2max, sim, obs, rep, CI, type, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta*s should be estimated (default is delta = 1).
R2max	Maximum R-square for which beta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement
CI	Confidence intervals, indicated as vector. Can be and/or 90, 95, 99.
type	Model type (either *lm* or *plm*; as string).
useed	User defined seed.
data	Dataset.

Details

Provides the mean and confidence intervals of estimated bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes the mean and confidence intervals of estimated bootstrapped beta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# compute the mean and confidence intervals of estimated bootstrapped beta*s
o_beta_boot_inf(y = "mpg",          # dependent variable
                x = "wt",           # independent treatment variable
                con = "hp + qsec",  # related control variables
                delta = 1,          # delta
                R2max = 0.9,        # maximum R-square
                sim = 100,          # number of simulations
                obs = 30,           # draws per simulation
                rep = FALSE,        # bootstrapping with or without replacement
                CI = c(90,95,99),   # confidence intervals
                type = "lm",        # model type
                useed = 123,        # seed
                data = data_oster)  # dataset
```

---

o_beta_boot_viz   *Visualization of bootstrapped beta*s*

Description

Estimates and visualizes bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019).

Usage

```
o_beta_boot_viz(y, x, con, w = NULL, id = "none", time = "none",
delta = 1, R2max, sim, obs, rep, CI, type, norm = TRUE, bin,
col = c("#08306b","#4292c6","#c6dbef"), nL = TRUE, mL = TRUE, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta*s should be estimated (default is delta = 1).
R2max	Maximum R-square for which beta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement
CI	Confidence intervals, indicated as vector. Can be and/or 90, 95, 99.
type	Model type (either *lm* or *plm*; as string).
norm	Option to include a normal distribution in the plot (default is norm = TURE).
bin	Number of bins used in the histogram.
col	Colors used to indicate different confidence interval levels (indicated as vector). Needs to be the same length as the variable CI. The default is a blue color range.
nL	Option to include a red vertical line at 0 (default is nL = TRUE).
mL	Option to include a vertical line at mean of all beta*s (default is mL = TRUE).
useed	User defined seed.
data	Dataset.

Details

Estimates and visualizes bootstrapped beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns ggplot2 object, which depicts the bootstrapped beta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate and visualize bootstrapped beta*s
o_beta_boot_viz(y = "mpg",          # dependent variable
                x = "wt",           # independent treatment variable
                con = "hp + qsec",  # related control variables
                delta = 1,          # delta
                R2max = 0.9,        # maximum R-square
                sim = 100,          # number of simulations
                obs = 30,           # draws per simulation
                rep = FALSE,        # bootstrapping with or without replacement
                CI = c(90,95,99),   # confidence intervals
                type = "lm",        # model type
                norm = TRUE,        # normal distribution
                bin = 200,          # number of bins
                useed = 123,        # seed
                data = data_oster)  # dataset
```

---

o_beta_rsq   *beta*s over a range of maximum R-squares*

Description

Estimates beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019) over a range of maximum R-squares.

Usage

```
o_beta_rsq(y, x, con, w = NULL, id = "none", time = "none", delta = 1,
type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta*s should be estimated (default is delta = 1).
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019) over a range of maximum R-squares. The range of maximum R-squares starts from the R-square of the controlled model rounded up to the next 1/100 to 1. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes beta*s over a range of maximum R-squares.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate delta*s over a range of maximum R-squares
o_beta_rsq(y = "mpg",          # dependent variable
           x = "wt",           # independent treatment variable
           con = "hp + qsec",  # related control variables
           delta = 1,          # delta
```

```
         type = "lm",          # model type
         data = data_oster)    # dataset
```

---

o_beta_rsq_viz   *Visualization of beta*s over a range of maximum R-squares*

Description

Estimates and visualizes beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019) over a range of maximum R-squares.

Usage

```
o_beta_rsq_viz(y, x, con, w = NULL, id = "none", time = "none", delta = 1,
type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
delta	delta for which beta*s should be estimated (default is delta = 1).
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates and visualizes beta*s, i.e., the bias-adjusted treatment effects (or correlations) (following Oster 2019) over a range of maximum R-squares. The range of maximum R-squares starts from the R-square of the controlled model rounded up to the next 1/100 to 1. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns ggplot2 object, which depicts beta*s over a range of maximum R-squares.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate and visualize beta*s over a range of maximum R-squares
o_beta_rsq_viz(y = "mpg",         # dependent variable
               x = "wt",          # independent treatment variable
               con = "hp + qsec",   # related control variables
               delta = 1,         # delta
               type = "lm",       # model type
               data = data_oster)   # dataset
```

---

o_delta   *delta**

Description

Estimates delta*, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019). The code is based on the psacalc command in Stata.

Usage

```
o_delta(y, x, con, w = NULL, id = "none", time = "none", beta = 0, R2max,
type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta* should be estimated (default is beta = 0).

R2max	Maximum R-square for which delta* should be estimated.
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates delta*, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019). The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes delta* and various other information.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate delta*
o_delta(y = "mpg",          # dependent variable
        x = "wt",           # independent treatment variable
        con = "hp + qsec",  # related control variables
        beta = 0,           # beta
        R2max = 0.9,        # maximum R-square
        type = "lm",        # model type
        data = data_oster)  # dataset
```

---

o_delta_boot   *Bootstrapped delta*s*

Description

Estimates bootstrapped delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019).

Usage

```
o_delta_boot(y, x, con, w = NULL, id = "none", time = "none", beta = 0, R2max,
sim, obs, rep, type, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta*s should be estimated (default is beta = 0).
R2max	Maximum R-square for which delta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement.
type	Model type (either *lm* or *plm*; as string).
useed	User defined seed.
data	Dataset.

Details

Estimates bootstrapped delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes bootstrapped delta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate bootstrapped delta*s
o_delta_boot(y = "mpg",        # dependent variable
             x = "wt",         # independent treatment variable
             con = "hp + qsec", # related control variables
             beta = 0,         # beta
             R2max = 0.9,      # maximum R-square
             sim = 100,        # number of simulations
             obs = 30,         # draws per simulation
             rep = FALSE,      # bootstrapping with or without replacement
             type = "lm",      # model type
             useed = 123,      # seed
             data = data_oster) # dataset
```

---

o_delta_boot_inf   *Bootstrapped mean delta* and confidence intervals*

Description

Provides the mean and confidence intervals of bootstrapped delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019).

Usage

```
o_delta_boot_inf(y, x, con, w = NULL, id = "none", time = "none",
beta = 0, R2max, sim, obs, rep, CI, type, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.

time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta*s should be estimated (default is beta = 0)..
R2max	Maximum R-square for which delta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement
CI	Confidence intervals, indicated as vector. Can be and/or 90, 95, 99.
type	Model type (either *lm* or *plm*; as string).
useed	User defined seed.
data	Dataset.

Details

Provides the mean and confidence intervals of bootstrapped delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes the mean and confidence intervals of bootstrapped delta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# compute the mean and confidence intervals of estimated bootstrapped delta*s
o_delta_boot_inf(y = "mpg",            # dependent variable
               x = "wt",               # independent treatment variable
               con = "hp + qsec",      # related control variables
               beta = 0,               # beta
               R2max = 0.9,            # maximum R-square
               sim = 100,              # number of simulations
```

```
            obs = 30,              # draws per simulation
            rep = FALSE,           # bootstrapping with or without replacement
            CI = c(90,95,99),      # confidence intervals
            type = "lm",           # model type
            useed = 123,           # seed
            data = data_oster)     # dataset
```

---

o_delta_boot_viz   *Visualization of bootstrapped delta*s*

Description

Estimates and visualizes bootstrapped delta*s, i.e., the degree of selection on unobservables relative
to observables (with respect to the treatment variable) that would be necessary to eliminate the result
(following Oster 2019).

Usage

```
o_delta_boot_viz(y, x, con, w = NULL, id = "none", time = "none",
beta = 0, R2max, sim, obs, rep, CI, type, norm = TRUE, bin,
col = c("#08306b","#4292c6","#c6dbef"), nL = TRUE, mL = TRUE, useed = NA, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta*s should be estimated (default is beta = 0).
R2max	Maximum R-square for which delta*s should be estimated.
sim	Number of simulations.
obs	Number of draws per simulation.
rep	Bootstrapping either with (= TRUE) or without (= FALSE) replacement
CI	Confidence intervals, indicated as vector. Can be and/or 90, 95, 99.
type	Model type (either *lm* or *plm*; as string).
norm	Option to include a normal distribution in the plot (default is norm = TURE).
bin	Number of bins used in the histogram.

col	Colors used to indicate different confidence interval levels (indicated as vector). Needs to be the same length as the variable CI. The default is a blue color range.
nL	Option to include a red vertical line at 0 (default is nL = TRUE).
mL	Option to include a vertical line at beta* mean (default is mL = TRUE).
useed	User defined seed.
data	Dataset.

Details

Estimates and visualizes bootstrapped delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019). Bootstrapping can either be done with or without replacement. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns ggplot2 object, which depicts the bootstrapped delta*s.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate and visualize bootstrapped delta*s
o_delta_boot_viz(y = "mpg",          # dependent variable
                 x = "wt",           # independent treatment variable
                 con = "hp + qsec",  # related control variables
                 beta = 0,           # beta
                 R2max = 0.9,        # maximum R-square
                 sim = 100,          # number of simulations
                 obs = 30,           # draws per simulation
                 rep = FALSE,        # bootstrapping with or without replacement
                 CI = c(90,95,99),   # confidence intervals
                 type = "lm",        # model type
                 norm = TRUE,        # normal distribution
                 bin = 200,          # number of bins
                 useed = 123,        # seed
                 data = data_oster)  # dataset
```

---

o_delta_rsq   *delta*s over a range of maximum R-squares*

Description

Estimates delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019) over a range of maximum R-squares following Oster (2019).

Usage

```
o_delta_rsq(y, x, con, w = NULL, id = "none", time = "none", beta = 0,
type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta*s should be estimated (default is beta = 0).
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019) over a range of maximum R-squares. The range of maximum R-squares starts from the R-square of the controlled model rounded up to the next 1/100 to 1. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns tibble object, which includes delta*s over a range of maximum R-squares.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate delta*s over a range of maximum R-squares
o_delta_rsq(y = "mpg",          # dependent variable
            x = "wt",           # independent treatment variable
            con = "hp + qsec",  # related control variables
            beta = 0,           # beta
            type = "lm",        # model type
            data = data_oster)  # dataset
```

o_delta_rsq_viz   *Visualization of delta*s over a range of maximum R-squares*

Description

Estimates and visualizes delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019) over a range of maximum R-squares.

Usage

```
o_delta_rsq_viz(y, x, con, w = NULL, id = "none", time = "none", beta = 0,
type, data)
```

Arguments

y	Name of the dependent variable (as string).
x	Name of the independent treatment variable (i.e., variable of interest; as string).
con	Name of related control variables. Provided as string in the format: "w + z +...".
w	weights (only for weighted estimations). Warning: For weighted panel models R can report different R-square than Stata, leading deviation between R and Stata results.
id	Name of the individual id variable (e.g. firm or farm; as string). Only applicable for fixed effect panel models.
time	Name of the time id variable (e.g. year or month; as string). Only applicable for fixed effect panel models.
beta	beta for which delta*s should be estimated (default is beta = 0).
type	Model type (either *lm* or *plm*; as string).
data	Dataset.

Details

Estimates and visualizes delta*s, i.e., the degree of selection on unobservables relative to observables (with respect to the treatment variable) that would be necessary to eliminate the result (following Oster 2019) over a range of maximum R-squares. The range of maximum R-squares starts from the R-square of the controlled model rounded up to the next 1/100 to 1. The function supports linear cross-sectional (see *lm* objects in R) and fixed effect panel (see *plm* objects in R) models.

Value

Returns ggplot2 object, which depicts delta*s over a range of maximum R-squares.

References

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. Journal of Business & Economic Statistics, 37, 187-204.

Examples

```
# load data, e.g. the in-build mtcars dataset
data("mtcars")
data_oster <- mtcars

# preview of data
head(data_oster)

# load robomit
require(robomit)

# estimate and visualize delta*s over a range of maximum R-squares
o_delta_rsq_viz(y = "mpg",          # dependent variable
                x = "wt",           # independent treatment variable
                con = "hp + qsec",  # related control variables
                beta = 0,           # beta
                type = "lm",        # model type
                data = data_oster)  # dataset
```

Index