

Package ‘TwoWayFEWeights’

October 12, 2022

Type Package

Title Estimation of the Weights Attached to the Two-Way Fixed Effects Regressions

Version 0.1.0

Maintainer Shuo Zhang <shuo_zhang@ucsb.edu>

Description

Estimates the weights and measure of robustness to treatment effect heterogeneity attached to two-way fixed effects regressions.

Clément de Chaisemartin, Xavier D’Haultfœuille (2020) <[DOI:10.1257/aer.20181169](https://doi.org/10.1257/aer.20181169)>.

License MIT + file LICENSE

Encoding UTF-8

Depends R(>= 3.4.0)

Imports dplyr (>= 1.0.0), fixest (>= 0.6.0), stringr, stats, estimatr, rlang

Suggests wooldridge

NeedsCompilation no

Author Shuo Zhang [aut, cre],
Clément de Chaisemartin [aut]

Repository CRAN

Date/Publication 2021-05-15 11:40:02 UTC

R topics documented:

twowayfeweights 2

Index 6

twowayfeweights	<i>Estimation of the weights attached to the two-way fixed effects regressions.</i>
-----------------	---

Description

Estimates the weights and measure of robustness to treatment effect heterogeneity attached to two-way fixed effects regressions.

Usage

```
twowayfeweights(df, Y, G, T, D, cmd_type, D0 = NULL, controls = c(), weights = NULL,
  other_treatments = c(), test_random_weights = c())
```

Arguments

df	data frame or data matrix.
Y	the dependent variable in the regression. Y is the level of the outcome if one wants to estimate the weights attached to the fixed-effects regression, and Y is the first difference of the outcome if one wants to estimate the weights attached to the first-difference regression.
G	a variable identifying each group.
T	a variable identifying each period.
D	the treatment variable in the regression. D is the level of the treatment if one wants to estimate the weights attached to the fixed-effects regression, and D is the first difference of the treatment if one wants to estimate the weights attached to the first-difference regression.
cmd_type	a required option that can take four values: feTR, feS, fdTR, fdS. If feTR is specified, the command estimates the weights and sensitivity measures attached to the fixed-effects regression under the common trends assumption. With feS, it estimates the weights and sensitivity measures attached to the fixed-effects regression under common trends and the assumption that groups' treatment effect does not change over time. With fdTR, it estimates the weights and sensitivity measures attached to the first-difference regression under the common trends assumption. Finally, with fdS it estimates the weights and sensitivity measures attached to the first-difference regression under common trends and the assumption that groups' treatment effect does not change over time.
D0	If cmd_type is specified in the option type below, then the command requires a fifth argument, D0. D0 is the mean of the treatment in group g and at period t. It should be non-missing at the first period when a group appears in the data (e.g. at t=1 for the groups that are in the data from the beginning), and for all observations for which the first-difference of the group-level mean outcome and treatment are non missing.

controls	a list of control variables that are included in the regression. Controls should not vary within each group*period cell, because the results in de Chaisemartin & D'Haultfoeuille (2020a) apply to two-way fixed effects regressions with group*period level controls. If a control does vary within a group*period cell, the command will replace it by its average value within each group*period cell, default to <code>c()</code> if not specified.
weights	if the regression is weighted, the weight variable can be specified in <code>weight</code> . If <code>type(fdTR)</code> is specified, then the weight variable should be non-missing at the first period when a group appears in the data (e.g. at <code>t=1</code> for the groups that are in the data from the beginning), and for all observations for which the first-difference of the group-level mean outcome and treatment are non missing.
other_treatments	a list of other treatment variables that are included in the regression. While the results in de Chaisemartin & D'Haultfoeuille (2020a) do not cover two-way fixed effects regressions with several treatments, those in de Chaisemartin & D'Haultfoeuille(2020b) do, so the command follows results from that second paper when <code>other_treatments</code> is specified. This option can only be used when <code>type(feTR)</code> is specified. When it is specified, the command reports the number and sum of positive and negative weights attached to the treatment, but it does not report the summary measures of the regression's robustness to heterogeneous treatment effects, as these summary measures are no longer applicable when the regression has several treatment variables. The command also reports the weights attached to the other treatments. The weights reported by the command are those in Corollary 1 in de Chaisemartin & D'Haultfoeuille (2020b). See de Chaisemartin & D'Haultfoeuille (2020b) for further details.
test_random_weights	weights when this option is specified, the command estimates the correlation between each variable in <code>varlist</code> and the weights. Testing if those correlations significantly differ from zero is a way to assess whether the weights are as good as randomly assigned to groups and time periods.

Details

It estimates the weights attached to the two-way fixed effects regressions studied in de Chaisemartin & D'Haultfoeuille (2020a), as well as summary measures of these regressions' robustness to heterogeneous treatment effects.

Value

`saved results` it saves the results in a dataframe containing 3 variables (Group, Time, Weight). This option allows the user to see the weight attached to each group*time cell. If the `other_treatments` option is specified, the weights attached to the other treatments are also saved.

FAQ

How can one interpret the summary measures of the regression's robustness to heterogeneous treatment effects? When the two-way fixed effects regression has only one treatment variable, the command reports two summary measures of the robustness of the treatment coefficient β to treatment

heterogeneity across groups and over time. The first one is defined in point (i) of Corollary 1 in de Chaisemartin & D'Haultfoeuille (2020a). It corresponds to the minimal value of the standard deviation of the treatment effect across the treated groups and time periods under which beta and the average treatment effect on the treated (ATT) could be of opposite signs. When that number is large, this means that beta and the ATT can only be of opposite signs if there is a lot of treatment effect heterogeneity across groups and time periods. When that number is low, this means that beta and the ATT can be of opposite signs even if there is not a lot of treatment effect heterogeneity across groups and time periods. The second summary measure is defined in point (ii) of Corollary 1 in de Chaisemartin & D'Haultfoeuille (2020a). It corresponds to the minimal value of the standard deviation of the treatment effect across the treated groups and time periods under which beta could be of a different sign than the treatment effect in all the treated group and time periods.

How can I tell if the first summary measure is high or low? Assume that the first summary measure is equal to x . How can you tell if x is a low or a high amount of treatment effect heterogeneity? This is not an easy question to answer, but here is one possibility. Let us assume that you find it a priori reasonable to assume that the treatment effect of every group and time period cannot be larger in absolute value than some real number $B > 0$. If you are trying to assess beta's robustness to heterogeneous effects, beta presumably falls within your range of a priori plausible values for the treatment effect, so it seems fair to argue that B is at least as large as $|\beta|$. Now let us also assume that the treatment effects of the treated groups and time periods are drawn from a uniform distribution. Then, to have that the mean of that distribution is 0 while its standard deviation is x , the treatment effects should be uniformly distributed on the $[-\sqrt{3}x, \sqrt{3}x]$ interval. If $|\beta| \geq \sqrt{3}x$, then uniformly distributed treatment effects with mean 0 and standard deviation x are compatible with your a priori plausible values for the treatment effect, so x may not be an implausibly high amount of treatment effect heterogeneity, and the ATT may be equal to 0. If on the other hand $|\beta| < \sqrt{3}x$, x may or may not be an implausibly high amount of treatment effect heterogeneity, depending on whether $B < \sqrt{3}x$ or $B \geq \sqrt{3}x$. The previous reasoning relies on the assumption that treatment effects follow a uniform distribution. You may find it more reasonable to assume that they are, say, normally distributed. Then you can conduct the following, similar exercise. Let us assume that you find it a priori reasonable to assume that most, say 95%, of the treatment effects are not larger in absolute value than some real number $B > 0$. If the treatment effects of the treated groups and time periods are drawn from a mean 0 and standard deviation x normal distribution, then 95% of them will fall within the $[-1.96x, 1.96x]$ interval. If $B \geq 1.96x$, $N(0, x^2)$ distributed treatment effects don't seem incompatible with your prior, so x may not be an implausibly high amount of treatment effect heterogeneity.

How can I tell if the second summary measure is high or low? Assume that the second summary measure is equal to x . Again, let us assume that you find it a priori reasonable to assume that the treatment effect of every group and time period cannot be larger in absolute value than some real number $B > 0$. Again, it seems fair to argue that B is at least as large as $|\beta|$. To fix ideas, let us assume that $\beta > 0$. Let us also assume that the treatment effects of the treated groups and time periods are drawn from a uniform distribution. Then, one could have that those effects are all negative, with a standard deviation equal to x , for instance if they are uniformly drawn from the $[-2\sqrt{3}x, 0]$ interval. If $|\beta| \geq 2\sqrt{3}x$, then treatment effects distributed on the $[-2\sqrt{3}x, 0]$ interval seem compatible with your a priori plausible range of values for the treatment effect, so x may not be an implausibly high amount of treatment effect heterogeneity. If on the other hand $|\beta| < 2\sqrt{3}x$, x may or may not be an implausibly high amount of treatment effect heterogeneity, depending on whether $B < 2\sqrt{3}x$ or $B \geq 2\sqrt{3}x$. If the treatment effects of the treated groups and time periods are all negative, they cannot follow a normal distribution, so we do not discuss that possibility here.

References

de Chaisemartin, C and D'Haultfoeuille, X (2020a). American Economic Review, vol. 110, no. 9. Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. de Chaisemartin, C and D'Haultfoeuille, X (2020b). Two-way fixed effects regressions with several treatments.

Examples

```
# using the same panel of workers as in Vella and Verbeek (1998)
library("wooldridge")
Y = "lwage"
G = "nr"
T = "year"
D = "union"
controls = c("hours")

twowayfeweights(wagepan, Y, G, T, D, cmd_type = "feTR", controls = controls)
```

Index

twowayfeweights, [2](#)