Description

tebalance box produces box plots that are used to check for balance in matched samples after teffects nnmatch and teffects psmatch.

Quick start

Box plot of the propensity score from the last teffects psmatch command
    tebalance box

Box plot of values of x1 in the treatment and control groups from raw data and the matched sample from the last teffects nnmatch or teffects psmatch command
    tebalance box x1

Menu

Statistics > Treatment effects > Balance > Graphs
Syntax

Box plots for the propensity score

tebalance box [ , options ]

Box plots for a covariate


tebalance box varname [ , options ]

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>boxlook_options</td>
<td>graph box options controlling how the box looks</td>
</tr>
<tr>
<td>legending_options</td>
<td>graph box options controlling how the variables are labeled</td>
</tr>
<tr>
<td>axis_options</td>
<td>graph box options controlling how numerical y axis is labeled</td>
</tr>
<tr>
<td>title_and_other_options</td>
<td>graph box options controlling titles, added text, aspect ratio, etc.</td>
</tr>
<tr>
<td>by_options</td>
<td>suboptions inside by() controlling plots by raw and matched samples</td>
</tr>
</tbody>
</table>

Options

boxlook_options are any of the options documented in boxlook_options in [G-2] graph box.
legending_options are any of the options documented in legending_options in [G-2] graph box.
axis_options are any of the options documented in axis_options in [G-2] graph box.
title_and_other_options are any of the options, except by(), documented in title_and_other_options in [G-2] graph box. tebalance box uses by() to differentiate between raw and matched samples, and some twoway_options will be repeated for by graph and might be better specified as byopts().
by_options are any of the byopts documented in [G-3] by_option. byopts() generally affects the entire graph, and some by_options may be better specified as twoway_options; see [G-3] twoway_options.

Remarks and examples

stata.com

When the distribution of a covariate does not vary over the treatment levels, the covariate is said to be balanced. tebalance box produces box plots of a covariate over treatment levels for the raw data and for the matched sample produced by teffects. If the matched-sample box plots are the same over the treatment levels, the covariate is balanced in the matched sample.

After teffects nnmatch and teffects psmatch,

. tebalance box varname [ , options ]

produces box plots to check whether varname is balanced.
After `teffects psmatch`,

```
    . tebalance box [, options]
```

produces box plots to check whether the propensity score estimated by `teffects` is balanced.

We recommend that you read [TE] `tebalance` before proceeding; it provides an introduction to covariate balance and an overview of the implemented methods.

**Example 1: Checking covariate balance after psmatch**

Using an extract from the data used by Cattaneo (2010), we use `teffects psmatch` to estimate the effect of a mother’s smoking behavior (`mbsmoke`) on the birthweight of her child (`bweight`), controlling for marital status (`mmarried`), the mother’s age (`mage`), whether the mother had a prenatal doctor’s visit in the baby’s first trimester (`prenatal1`), and whether this baby is the mother’s first child (`fbaby`).

```
. use https://www.stata-press.com/data/r16/cattaneo2
. teffects psmatch (bweight) (mbsmoke mmarried mage prenatal1 fbaby),
> generate(matchv)
Treatment-effects estimation
    Number of obs = 4,642
Estimator : propensity-score matching
Matches: requested = 1
Outcome model : matching
    min = 1
Treatment model: logit
    max = 139

| Coef. Std. Err. z    P>|z|    [95% Conf. Interval] |
|----------------------|--------------|---------|-----------|-----------------|
| **ATE**              |              |          |           |                 |
| `bweight`            |              | AI Robust|           |                 |
| `mbsmoke` (smoker vs nonsmoker) | -235.1714 | 27.74409 | -8.48     | 0.000           | -289.5488 -180.794 |
```

We specified the option `generate(matchv)` to speed up the postestimation command that produces density plots, as discussed in example 3 under Remarks and examples of [TE] `tebalance`. We do not interpret the estimated effect produced by this preliminary model but rather check the specification. Now we look at the box plots.
The box plots for the matched sample are very similar. The medians, the 25th percentiles, and the 75th percentiles appear to be the same, although there may be some differences in the tails, the upper adjacent values, the lower adjacent values, and the outliers. Matching on the estimated propensity score appears to have balanced `mage`, except for the tails.

To get an idea of whether `teffects psmatch` balanced all the covariates, we look at the box plots for the estimated propensity score.

```
. tebalance box
```

The box plots indicate that `teffects psmatch` balanced the estimated propensity scores.
Reference


Also see

[TE] tebalance — Check balance after teffects or stteffects estimation
[TE] teffects nnmatch — Nearest-neighbor matching
[TE] teffects overlap — Overlap plots
[TE] teffects psmatch — Propensity-score matching