

tebalance — Check balance after teffects or stteffects estimation

[Description](#) [Syntax](#) [Remarks and examples](#) [Methods and formulas](#)
[References](#) [Also see](#)

Description

The `tebalance` postestimation commands produce diagnostic statistics, test statistics, and diagnostic plots to assess whether a `teffects` or an `stteffects` command balanced the covariates over treatment levels.

Syntax

```
tebalance subcommand ... [ , options ]
```

<i>subcommand</i>	Description
<code>summarize</code>	compare means and variances in raw and balanced data
<code>overid</code>	overidentification test
<code>density</code>	kernel density plots for raw and balanced data
<code>box</code>	box plots for each treatment level for balanced data

Remarks and examples

[stata.com](#)

This entry provides an overview of the commands in `tebalance`. We recommend that you read this entry before proceeding to [\[TE\] `tebalance summarize`](#), [\[TE\] `tebalance overid`](#), [\[TE\] `tebalance density`](#), or [\[TE\] `tebalance box`](#) for command-specific syntax and details.

A covariate is said to be balanced when its distribution does not vary over treatment levels.

Covariates are balanced in experimental data because treatment assignment is independent of the covariates because of the study design. In contrast, covariates must be balanced by weighting or matching in observational data because treatment assignment is related to the covariates that also affect the outcome of interest.

The estimators implemented in `teffects` and `stteffects` use a model or matching method to make the outcome conditionally independent of the treatment by conditioning on covariates. If this model or matching method is well specified, it should balance the covariates. Balance diagnostic techniques and tests check the specification of the conditioning method used by a `teffects` or an `stteffects` estimator; see [\[TE\] `teffects intro advanced`](#) for an introduction to `teffects`, and [\[TE\] `stteffects intro`](#) for an introduction to `stteffects`.

`tebalance` implements four methods to check for balance after `teffects` and `stteffects`. Which `tebalance` methods are available depends on the `teffects` estimation method, as summarized in the table below.

2 tebalance — Check balance after teffects or stteffects estimation

tebalance		Works after teffects					Works after stteffects	
method	Description	ipw	aipw	ipwra	nnmatch	psmatch	ipw	ipwra
summarize	standardized differences and variance ratios	x	x	x	x	x	x	x
overid	chi-squared test for balance	x	x	x			x	x
density	diagnostic kernel-density plots	x	x	x	x	x	x	x
box	diagnostic box plots				x	x		

`tebalance` `overid` implements a formal test, while the other three methods are exploratory diagnostic techniques. There is no balance check after `teffects ra`, `stteffects ra`, or `stteffects wra`, because they use neither a treatment model nor a matching method.

Austin (2009, 2011) and Guo and Fraser (2015, sec. 5.52) provide introductions to covariate balance. Imai and Ratkovic (2014) derived a test for balance implemented in `tebalance overid`.

The remainder of this entry provides a quick introduction to using `tebalance` to check for balance after `teffects`. See [TE] [stteffects intro](#) for examples after `stteffects`.

► Example 1: Balance after estimators that use weighting

Inverse-probability-weighted (IPW) estimators use a model for the treatment to make the outcome conditionally independent of the treatment. If this model is well specified, it will also balance the covariates.

Using an extract from Cattaneo (2010), we use `teffects ipw` 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 http://www.stata-press.com/data/r15/cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)
. teffects ipw (bweight) (mbsmoke mmarried mage prenatal1 fbaby)

Iteration 0:  EE criterion = 1.873e-22
Iteration 1:  EE criterion = 3.315e-26

Treatment-effects estimation      Number of obs      =      4,642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: logit
```

bweight	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
ATE						
mbsmoke (smoker vs nonsmoker)	-236.1038	23.86187	-9.89	0.000	-282.8722	-189.3354
POmean						
mbsmoke nonsmoker	3402.552	9.539555	356.68	0.000	3383.855	3421.249

Rubin (2008) recommends finding a model that balances the covariates before looking at results for the estimated treatment effect. Thus we do not interpret the above results, and we note that we could pay closer heed to Rubin's recommendation by preceding the `teffects` command with `quietly` to suppress the output.

Imai and Ratkovic (2014) derived a test for balance by viewing the restrictions imposed by balance as overidentifying conditions. This test is implemented in `tebalance overid`, and we use it to test whether the above treatment model balanced the covariates.

```
. tebalance overid

Iteration 0:  criterion = .02146858
Iteration 1:  criterion = .02159149   (backed up)
Iteration 2:  criterion = .02177783
Iteration 3:  criterion = .02260102
Iteration 4:  criterion = .02267956
Iteration 5:  criterion = .02292367
Iteration 6:  criterion = .02431697
Iteration 7:  criterion = .02457043
Iteration 8:  criterion = .02488579
Iteration 9:  criterion = .02529453
Iteration 10: criterion = .02545885
Iteration 11: criterion = .02550248
Iteration 12: criterion = .02552866
Iteration 13: criterion = .02554462
Iteration 14: criterion = .02554512
Iteration 15: criterion = .02554514

Overidentification test for covariate balance
      H0: Covariates are balanced:
      chi2(5)      = 38.1464
      Prob > chi2 = 0.0000
```

We reject the null hypothesis that the treatment model balanced the covariates.

Let's use `tebalance summarize` to see where the problem lies. To get an idea of the extent to which the covariates are unbalanced, we begin by summarizing the covariates by group in the raw data by specifying the `baseline` option.

```
. tebalance summarize, baseline

Covariate balance summary
```

		Raw	Weighted
Number of obs =		4,642	4,642.0
Treated obs =		864	2,315.3
Control obs =		3,778	2,326.7

	Means		Variances	
	Control	Treated	Control	Treated
mmarried	.7514558	.4733796	.1868194	.2495802
mage	26.81048	25.16667	31.87141	28.10429
prenatal1	.8268925	.6898148	.1431792	.2142183
fbaby	.4531498	.3715278	.2478707	.2337654

The output indicates that the covariates may not be balanced in the raw data. For example, the distribution of the mother's age may differ over the treatment groups. We can investigate the differences further with standardized differences and variance ratios. A perfectly balanced covariate has a standardized difference of zero and variance ratio of one. There are no standard errors on these

4 tebalance — Check balance after teffects or stteffects estimation

statistics, so inference is informal. Austin (2009) provides a recent introduction to these diagnostics, although they have been used at least since Rosenbaum and Rubin (1985).

By omitting the baseline option, we obtain these diagnostic statistics for the raw data and the weighted data.

```
. tebalance summarize
```

Covariate balance summary

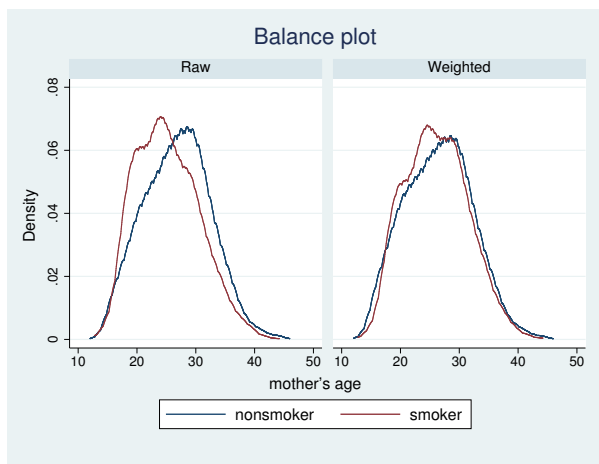
	Raw	Weighted
Number of obs =	4,642	4,642.0
Treated obs =	864	2,315.3
Control obs =	3,778	2,326.7

	Standardized differences		Variance ratio	
	Raw	Weighted	Raw	Weighted
mmarried	-.5953009	-.0105562	1.335944	1.009079
mage	-.300179	-.0672115	.8818025	.8536401
prenatal1	-.3242695	-.0156339	1.496155	1.023424
fbaby	-.1663271	.0257705	.9430944	1.005698

Reviewing the output, we see that for `mmarried`, `prenatal1`, and `fbaby`, our model improved the level of balance. The weighted standardized differences are all close to zero and the variance ratios are all close to one. However, output indicates that `mage` may not be balanced by our model. The weighted standardized difference is close to zero, but the weighted variance ratio still appears to be considerably less than one.

Now, let's use `tebalance density` to look at how the densities of `mage` for treated and control groups differ.

```
. tebalance density mage
```



The plots also indicate a lack of balance in `mage` between the treatment groups.

To try to achieve better balance, we specify a richer model with interactions between mage and the other covariates and look at the resulting standardized differences.

```
. quietly teffects ipw (bweight) (mbsmoke mmarried mage prenatal1 fbaby
> c.mage#(c.mage i.mmarried prenatal1))
```

```
. tebalance summarize
```

```
Covariate balance summary
```

	Raw	Weighted
Number of obs =	4,642	4,642.0
Treated obs =	864	2,329.1
Control obs =	3,778	2,312.9

	Standardized differences		Variance ratio	
	Raw	Weighted	Raw	Weighted
mmarried	-.5953009	.0053497	1.335944	.9953184
mage	-.300179	.0410889	.8818025	1.076571
prenatal1	-.3242695	.0009807	1.496155	.9985165
fbaby	-.1663271	-.0130638	.9430944	.9965406
mage#				
mage	-.3028275	.0477465	.8274389	1.109134
mmarried#				
mage				
married	-.6329701	.0197209	1.157026	1.034108
prenatal1#				
mage				
Yes	-.4053969	.0182109	1.226363	1.032561

6 tebalance — Check balance after teffects or stteffects estimation

The standardized difference and variance ratio results for `mage` look closer to the expected values of zero and one, so we proceed to the formal test.

```
. tebalance overid
Iteration 0: criterion = .0602349
Iteration 1: criterion = .06172749 (backed up)
Iteration 2: criterion = .06428588 (backed up)
Iteration 3: criterion = .06489623 (backed up)
Iteration 4: criterion = .06527284 (backed up)
Iteration 5: criterion = .06643426
Iteration 6: criterion = .07120383
Iteration 7: criterion = .07647097
Iteration 8: criterion = .07674915
Iteration 9: criterion = .07684127
Iteration 10: criterion = .07703321
Iteration 11: criterion = .0776508
Iteration 12: criterion = .07771863
Iteration 13: criterion = .07773156
Iteration 14: criterion = .07773561
Iteration 15: criterion = .07774891
Iteration 16: criterion = .07775314
Iteration 17: criterion = .07775324
Iteration 18: criterion = .07775325
Iteration 19: criterion = .07775325
Iteration 20: criterion = .07775325
Iteration 21: criterion = .07775325
Iteration 22: criterion = .07775325

Overidentification test for covariate balance
H0: Covariates are balanced:
      chi2(8)      = 11.8612
      Prob > chi2  = 0.1575
```

We do not reject the null hypothesis that the specified treatment model balances the covariates. ◀

► Example 2: Balance after estimators that use matching

Instead of weighting, we might want to use a matching estimator. We can select `teffects nnmatch` or `teffects psmatch` for balance and estimation; in this example, we use `teffects nnmatch`.

```
. teffects nnmatch (bweight mmarried mage prenatal1 fbaby)
> (mbsmoke), bias(mage) ematch(mmarried prenatal1 fbaby)

Treatment-effects estimation      Number of obs      =      4,642
Estimator      : nearest-neighbor matching      Matches: requested =      1
Outcome model  : matching                      min =      1
Distance metric: Mahalanobis                  max =      139
```

	bweight	Coef.	AI Robust Std. Err.	z	P> z	[95% Conf. Interval]
ATE						
	mbsmoke (smoker vs nonsmoker)	-240.4589	28.43008	-8.46	0.000	-296.1808 -184.7369

Again we ignore the estimated effect and first check for balance. We begin by reviewing whether the summary statistics indicate good balance.

```
. tebalance summarize
note: refitting the model using the generate() option
```

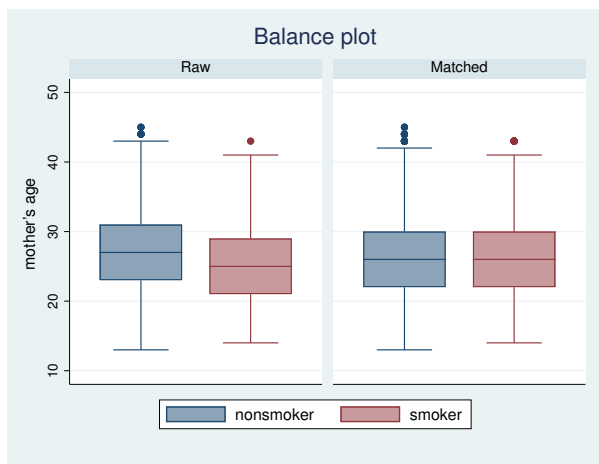
Covariate balance summary

		Raw	Matched	
<hr/>				
Number of obs =		4,642	9,284	
Treated obs =		864	4,642	
Control obs =		3,778	4,642	
<hr/>				
	Standardized differences	Variance ratio		
	Raw	Raw	Matched	
mmarried	-.5953009	-2.42e-16	1.335944	1
mage	-.300179	-.0040826	.8818025	.9815517
prenatal1	-.3242695	-2.78e-16	1.496155	1
fbaby	-.1663271	2.24e-16	.9430944	1

We do not have standard errors on these statistics, so we cannot make any formal conclusions, but the summary statistics appear to indicate much better balance than the IPW results. `tebalance summarize` has to refit the model to recover the matched sample because the `generate()` option was not specified on the `teffects nnmatch` command. The reestimation does not affect the results, although the computation takes longer; see [example 3](#) for details.

Because it is a matching estimator, and not an IPW estimator, we cannot use `tebalance overid` after `teffects nnmatch`. The matching estimators, however, provide diagnostic box plots using `tebalance box` that are not available after the IPW estimators.

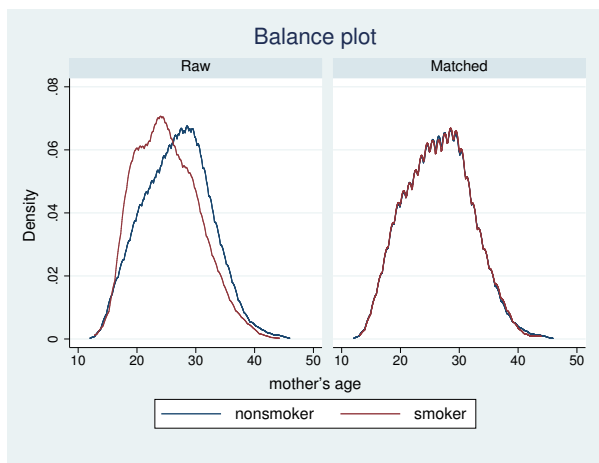
```
. tebalance box mage
note: refitting the model using the generate() option
```



The box plots of the matched data also indicate covariate balance.

Let's look at the kernel density plots using the matched data.

```
. tebalance density mage
note: refitting the model using the generate() option
```



The plots using the matched data appear to be balanced.

◀

□ Technical note

`teffects` implements matching estimators, IPW estimators, regression-adjustment (RA) estimators, and estimators that combine IPW and RA. Matching estimators define a matched sample, and IPW estimators define a weighted sample, each of which can be used to compute covariate balance statistics. RA estimators do not define an adjusted sample that can be used to compute covariate balance statistics, and `tebalance` does not work after `teffects ra`. Only the IPW component of the estimators that combine RA and IPW defines a weighted sample that can be used to compute balance statistics. So, `tebalance` produces the same results after `teffects aipw` or `teffects ipwra` as it does after `teffects ipw`.

□

▷ Example 3: Faster results after a matching estimator

The `tebalance` commands run in [example 2](#) executed more slowly than necessary. `tebalance` issued the note

```
note: refitting the model using the generate() option
```

after the commands

```
. tebalance summarize
```

```
. tebalance box mage
```

and

```
. tebalance density mage
```


After `teffects nnmatch` or `teffects psmatch`, `tebalance` computes the balance statistics on the matched sample defined by the matching estimator. `teffects nnmatch` and `teffects psmatch` leave behind only variables that identify the matched sample when the `generate()` option is specified. Unless the `generate()` option is specified with `teffects nnmatch` or `teffects psmatch`, each `tebalance` command must rerun the estimation command to recover the matched sample.

Typing

```
. teffects nnmatch (bweight mmarried mage fbaby prenatal1)
> (mbsmoke), bias(mage) ematch(mmarried fbaby prenatal1)
> generate(matchv)
```

would generate variables that identify the matched sample that the `tebalance` commands could use. See *Remarks and examples* in [TE] [tebalance box](#), [TE] [tebalance density](#), and [TE] [tebalance summarize](#) for examples using the option `generate()` on `teffects psmatch` to speed up the postestimation computations.

◀

Methods and formulas

Methods and formulas are presented under the following headings:

- [Introduction](#)
- [Matched samples](#)
- [IPW samples](#)
- [Testing the propensity-score model specification](#)

Introduction

For covariate z , we observe values $\{z_1, z_2, \dots, z_N\}$. Define a treatment indicator variable for J treatment levels as $t_i \in \{1, 2, \dots, J\}$, for $i = 1, \dots, N$, and frequency weights as $\{w_1, w_2, \dots, w_N\}$. The sample mean and variance of z for level t are

$$\hat{\mu}_z(t) = \frac{\sum_i^N I(t_i = t) w_i z_i}{N_t} \quad \text{and}$$

$$\hat{\sigma}_z^2(t) = \frac{\sum_i^N I(t_i = t) w_i \{z_i - \hat{\mu}_z(t)\}^2}{N_t - 1}$$

where $N_t = \sum_i^N w_i I(t_i = t)$, and

$$I(t_i = t) = \begin{cases} 1 & \text{if } t_i = t \\ 0 & \text{otherwise} \end{cases}$$

As shown in [Austin \(2011\)](#), the standardized differences for covariate z between level t and the control t_0 are computed as

$$\delta_z(t) = \frac{\hat{\mu}_z(t) - \hat{\mu}_z(t_0)}{\sqrt{\frac{\hat{\sigma}_z^2(t) + \hat{\sigma}_z^2(t_0)}{2}}} \quad (1)$$

The variance ratio is $\rho_z(t) = \{\hat{\sigma}_z^2(t)\} / \{\hat{\sigma}_z^2(t_0)\}$.

Matched samples

We now turn our attention to the matched samples for the potential-outcome mean (POM), average treatment effect (ATE), and average treatment effect on the treated (ATET) estimators. We estimate the covariate for the counter-factual treatment by taking the mean of the matched observations

$$\dot{z}_i = \frac{\sum_{j \in \Omega(i)} w_j z_j}{\sum_{j \in \Omega(i)} w_j}$$

where $\Omega(i) = (k_1, k_2, \dots, k_{m_j})$ is the set of observation indices that are matched to observation i of the opposite treatment condition. The observed covariate and matched covariate pairs, (z_i, \dot{z}_i) , $i = 1, \dots, N$, are used in the box plot (see [G-2] **graph box**) and the kernel density plot (see [R] **kdensity**). The ATET sample is limited to those observations from the conditional treatment, \tilde{t} , and their matched covariate means.

In *Methods and formulas* of [TE] **teffects nmatch**, we define $K_m(i)$ as the number of times observation i is used in a match with observation j of the opposite treatment condition, $i \in \Omega(j)$, weighted by the total number of matches for observation j . Specifically,

$$K_m(i) = \sum_{j=1}^N I\{i \in \Omega(j)\} \frac{w_j}{\sum_{k \in \Omega(j)} w_k}$$

These weights are used in the estimation of the mean and variance for the matched dataset. For the POM and ATE models, the estimated mean and variance are computed as

$$\begin{aligned} \hat{\mu}_{\dot{z}}(t) &= \frac{\sum_i^N I(t_i = t) w_i z_i \{1 + K_m(i)\}}{M_t} \quad \text{and} \\ \hat{\sigma}_{\dot{z}}^2(t) &= \frac{\sum_i^N I(t_i = t) w_i \{1 + K_m(i)\} \{z_i - \hat{\mu}_{\dot{z}}(t)\}^2}{M_t - 1} \end{aligned}$$

where $M_t = \sum_i^N I(t_i = t) w_i \{1 + K_m(i)\}$.

The standardized differences between the control level and all other levels for the matched covariate distribution are computed as in (1), but $\hat{\mu}_{\dot{z}}(t)$ is substituted for $\hat{\mu}_z(t)$ and $\hat{\sigma}_{\dot{z}}^2(t)$ for $\hat{\sigma}_z^2(t)$.

For the ATET model, there is no matched sample for the treatment levels other than the conditional treatment \tilde{t} . The covariate mean and variance for the conditional treatment are the same as that of the raw data, $\mu_z(\tilde{t})$ and $\sigma_z(\tilde{t})$. However, the covariate mean and variance for the sample matched to the conditional treatment, $t \neq \tilde{t}$, are computed as

$$\begin{aligned} \tilde{\mu}_{\dot{z}}(t) &= \frac{\sum_i^N I(t_i = t) w_i z_i K_m(i)}{M_t} \quad \text{and} \\ \tilde{\sigma}_{\dot{z}}^2(t) &= \frac{\sum_i^N I(t_i = t) w_i K_m(i) \{z_i - \tilde{\mu}_{\dot{z}}(t)\}^2}{M_t - 1} \end{aligned}$$

where $M_t = \sum_i^N I(t_i = t) w_i K_m(i)$.

IPW samples

Computation of the inverse-probability weights is discussed in *Methods and formulas* of [TE] **teffects aipw** and in *Methods and formulas* of [TE] **stteffects ipwra**. For the POM and ATE estimators, we defined the normalized IPW weights as $\bar{d}_i(t) = N_t d_i(t) / \sum_i^N d_i(t)$, where $d_i(t) = I(t_i = t) / p(\mathbf{z}_i, t, \hat{\gamma})$ for treatment level t and individual i .

For the ATET estimator, we use the normalized weights $\bar{f}_i = N f_i / \sum_i^N f_i$, where $f_i = p(\mathbf{z}_i, \tilde{t}, \hat{\gamma}) / p(\mathbf{z}_i, t_i, \hat{\gamma})$ are the treatment-adjusted inverse-probability weights, and \tilde{t} is the conditional treatment.

We will simplify notation by defining a single weight

$$\bar{w}_i(t) = \begin{cases} \bar{d}_i(t) & \text{if model is ATE or POM} \\ \bar{f}_i(t) & \text{if model is ATET} \end{cases}$$

The covariate mean and variance for treatment level t are

$$\begin{aligned} \tilde{\mu}_{\tilde{z}}(t) &= \frac{\sum_i^N I(t_i = t) w_i \bar{w}_i x_i}{M_t} & \text{and} \\ \tilde{\sigma}_{\tilde{z}}^2(t) &= \frac{I(t_i = t) w_i \bar{w}_i \{z_i - \tilde{\mu}_{\tilde{z}}(t)\}^2}{M_t - 1} \end{aligned}$$

where $M_t = \sum_i^N I(t_i = t) w_i \bar{w}_i$.

The kernel density is computed by `kdensity` for each covariate conditioned on each treatment level using the raw covariate with `iweights` equal to $w_i \bar{w}_i$.

Testing the propensity-score model specification

We estimate the probability of treatment conditioned on a set of covariates with a propensity-score model. Imai and Ratkovic (2014) derive a test for whether the estimated propensity score balances the covariates. The score equations for parameters of the propensity-score model define an exactly identified generalized method of moments (GMM) estimator. Imai and Ratkovic (2014) use the conditions imposed by mean balance as overidentifying conditions. A standard GMM test for the validity of the overidentifying conditions is then a test for covariate balance. See [R] `gmm` for a discussion of this overidentifying test, which is known as Hansen's J test in the econometrics literature.

Here are the details about the score equations and the overidentifying balance conditions. Recall from *Methods and formulas* of [TE] **teffects aipw** and *Methods and formulas* of [TE] **stteffects ipwra**, we have the first-order condition of the treatment model

$$\frac{1}{N} \sum_{i=1}^N s_{\text{tm},i}(\mathbf{z}_i, \hat{\gamma}) = 0$$

For a two-level treatment-effects model with conditional treatment \tilde{t} and control t_0 , the score is

$$s_{\text{tm},i}(\mathbf{z}_i, \gamma) = \frac{I(t_i = \tilde{t})}{p(\mathbf{z}_i, \tilde{t}, \gamma)} \frac{\partial p(\mathbf{z}_i, t, \gamma)}{\partial \gamma'} - \left\{ \frac{I(t_i = t_0)}{1 - p(\mathbf{z}_i, \tilde{t}, \gamma)} \right\} \frac{\partial p(\mathbf{z}_i, \tilde{t}, \gamma)}{\partial \gamma'} \Bigg|_{\gamma = \hat{\gamma}}$$

The score reduces to

$$\mathbf{s}_{\text{tm},i}(\mathbf{z}_i, \tilde{\gamma}) = \left[\frac{I(t_i = \tilde{t}) - p(\mathbf{z}_i, \tilde{t}, \gamma)}{p(\mathbf{z}_i, \tilde{t}, \gamma) \{1 - p(\mathbf{z}_i, \tilde{t}, \gamma)\}} \right] \frac{\partial p(\mathbf{z}_i, \tilde{t}, \gamma)}{\partial \gamma'} \Bigg|_{\gamma = \tilde{\gamma}}$$

The corresponding covariate balancing moment conditions are

$$\mathbf{w}_{\text{tm},i}(\mathbf{z}_i, \gamma) = \left[\frac{I(t_i = \tilde{t}) - p(\mathbf{z}_i, \tilde{t}, \gamma)}{p(\mathbf{z}_i, \tilde{t}, \gamma) \{1 - p(\mathbf{z}_i, \tilde{t}, \gamma)\}} \right] \mathbf{z}_i$$

for the POM and ATE models. For the ATET model with conditional treatment \tilde{t} , we multiply by $p(\mathbf{z}_i, \tilde{t}, \gamma)$ and scale by $N/N_{\tilde{t}}$:

$$\mathbf{w}_{\text{tm},i}(\mathbf{z}_i, \gamma) = \frac{N}{N_{\tilde{t}}} \left\{ \frac{I(t_i = \tilde{t}) - p(\mathbf{z}_i, \tilde{t}, \gamma)}{1 - p(\mathbf{z}_i, \tilde{t}, \gamma)} \right\} \mathbf{z}_i$$

We stack the moment conditions

$$\begin{aligned} \mathbf{g}_{\text{tm}}(\mathbf{Z}, \gamma) &= \frac{1}{N} \sum_{i=1}^N \begin{Bmatrix} \mathbf{s}_{\text{tm},i}(\mathbf{z}_i, \gamma) \\ \mathbf{w}_{\text{tm},i}(\mathbf{z}_i, \gamma) \end{Bmatrix} \\ &= \frac{1}{N} \sum_{i=1}^N \mathbf{g}_{\text{tm},i}(\mathbf{z}_i, \gamma) \end{aligned}$$

The overidentified GMM estimator is then

$$\tilde{\gamma} = \operatorname{argmin}_{\gamma} N \mathbf{g}_{\text{tm}}(\mathbf{Z}, \gamma)' \mathbf{W}_{\text{tm}}(\mathbf{Z}, \gamma)^{-1} \mathbf{g}_{\text{tm}}(\mathbf{Z}, \gamma) \quad (2)$$

where

$$\mathbf{W}_{\text{tm}}(\mathbf{Z}, \gamma) = \frac{1}{N} \sum_{i=1}^N E_T \{ \mathbf{g}_{\text{tm},i}(\mathbf{z}, \gamma) \mathbf{g}_{\text{tm},i}(\mathbf{z}, \gamma)' \}$$

and the expectation is taken with respect to treatment distribution. The weight matrix $\mathbf{W}_{\text{tm}}(\mathbf{Z}, \gamma)$ is computed explicitly (Imai and Ratkovic 2014), and (2), written as a maximization problem, is solved using `m1`.

Finally, Hansen's J statistic is evaluated at its minimum,

$$J = N \mathbf{g}_{\text{tm}}(\mathbf{Z}, \tilde{\gamma})' \mathbf{W}_{\text{tm}}(\mathbf{Z}, \tilde{\gamma})^{-1} \mathbf{g}_{\text{tm}}(\mathbf{Z}, \tilde{\gamma})$$

and is asymptotically distributed chi-squared with degrees of freedom d ,

$$d = \operatorname{rank} \{ \mathbf{W}_{\text{tm}}(\mathbf{Z}, \tilde{\gamma}) \} - \operatorname{rank} \left[\frac{1}{N} \sum_{i=1}^N E_T \{ \mathbf{s}_{\text{tm},i}(\mathbf{z}_i, \tilde{\gamma}) \mathbf{s}_{\text{tm},i}(\mathbf{z}_i, \tilde{\gamma})' \} \right]$$

References

- Austin, P. C. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine* 28: 3083–3107.
- . 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46: 399–424.
- Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154.
- Guo, S., and M. W. Fraser. 2015. *Propensity Score Analysis: Statistical Methods and Applications*. 2nd ed. Thousand Oaks, CA: Sage.
- Imai, K., and M. Ratkovic. 2014. Covariate balancing propensity score. *Journal of the Royal Statistical Society, Series B* 76: 243–263.
- Rosenbaum, P. R., and D. B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician* 39: 33–38.
- Rubin, D. B. 2008. For objective causal inference, design trumps analysis. *Annals of Applied Statistics* 2: 808–840.

Also see

[TE] [tebalance summarize](#) — Covariate-balance summary statistics

[TE] [tebalance overrides](#) — Test for covariate balance

[TE] [tebalance density](#) — Covariate balance density

[TE] [tebalance box](#) — Covariate balance box

[TE] [teffects](#) — Treatment-effects estimation for observational data

[TE] [stteffects intro](#) — Introduction to treatment effects for observational survival-time data