tebalance — Check balance after teffects or stteffects estimation

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*]

subcommand	Description
`summarize`	compare means and variances in raw and balanced data
`overid`	overidentification test
`density`	kernel density plots for raw and balanced data
`box`	box plots for each treatment level for balanced data

Remarks and examples

This entry provides an overview of the commands in `tebalance`. We recommend that you read this entry before proceeding to [CAUSAL] **tebalance summarize**, [CAUSAL] **tebalance overid**, [CAUSAL] **tebalance density**, or [CAUSAL] **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 [CAUSAL] **teffects intro advanced** for an introduction to `teffects`, and [CAUSAL] **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.

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	χ^2 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 [CAUSAL] **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 https://www.stata-press.com/data/r19/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	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
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 =  .02267955
Iteration 5:  Criterion =  .02292361
Iteration 6:  Criterion =   .0243172
Iteration 7:  Criterion =  .02457057
Iteration 8:  Criterion =  .02488578
Iteration 9:  Criterion =  .02529419
Iteration 10: Criterion =  .02545882
Iteration 11: Criterion =  .02550251
Iteration 12: Criterion =  .02552869
Iteration 13: Criterion =  .02554463
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 standard-

ized difference of zero and variance ratio of one. There are no standard errors on these 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.

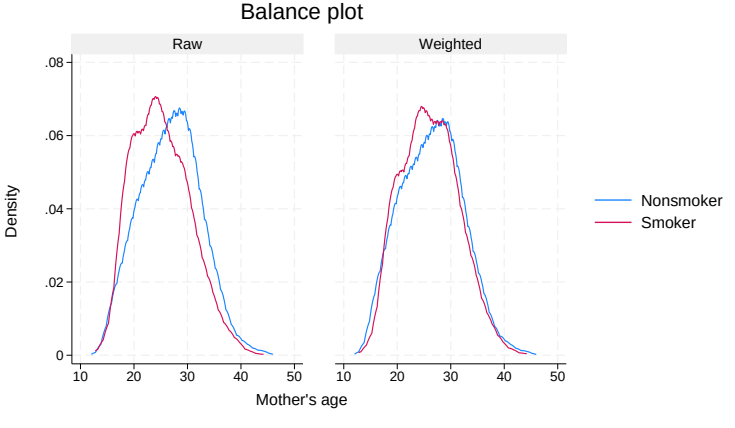
```
. tebalance summarize
Covariate balance summary
                          Raw      Weighted
------------------------------------------
Number of obs =         4,642       4,642.0
Treated obs   =           864       2,315.3
Control obs   =         3,778       2,326.7
------------------------------------------
```

	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

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 =  .06428188  (backed up)
Iteration 3:  Criterion =  .06489132  (backed up)
Iteration 4:  Criterion =  .06527353  (backed up)
Iteration 5:  Criterion =   .0664376
Iteration 6:  Criterion =  .07184528
Iteration 7:  Criterion =   .0762921
Iteration 8:  Criterion =   .0767141
Iteration 9:  Criterion =  .07679677
Iteration 10: Criterion =  .07699122
Iteration 11: Criterion =   .0776527
Iteration 12: Criterion =  .07771774
Iteration 13: Criterion =  .07772609
Iteration 14: Criterion =  .07773294
Iteration 15: Criterion =    .077752
Iteration 16: Criterion =  .07775324
Iteration 17: Criterion =  .07775325
Iteration 18: 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	Coefficient	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
(refitting the model using the generate() option)
Covariate balance summary
```

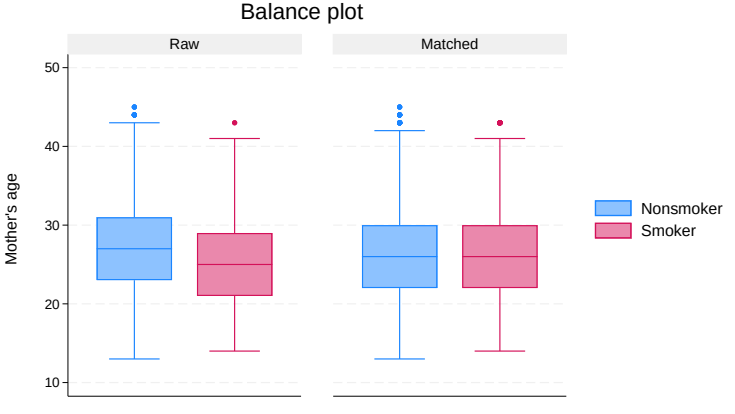
	Raw	Matched
Number of obs =	4,642	9,284
Treated obs =	864	4,642
Control obs =	3,778	4,642

	Standardized differences		Variance ratio	
	Raw	Matched	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.

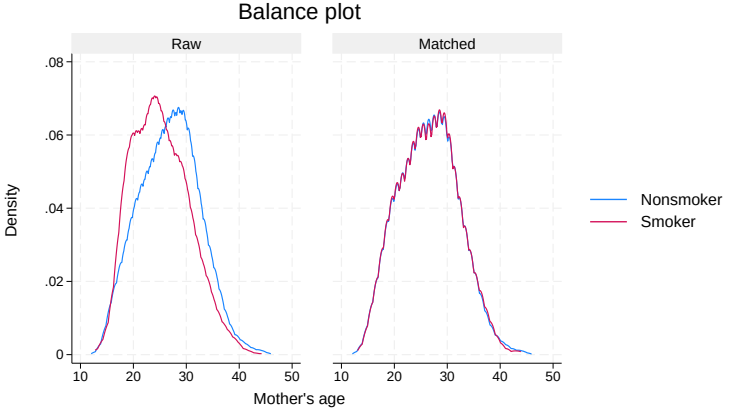
```
. tebalance box mage
(refitting the model using the generate() option)
```



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
(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

```
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 [CAUSAL] **tebalance box**, [CAUSAL] **tebalance density**, and [CAUSAL] **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

For covariate z, we observe values $\{z_1, z_2, \ldots, z_N\}$. Define a treatment indicator variable for J treatment levels as $t_i \in \{1, 2, \ldots, J\}$, for $i = 1, \ldots, N$, and frequency weights as $\{w_1, w_2, \ldots, w_N\}$. 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

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

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 $\Omega(i) = (k_1, k_2, \ldots, k_{m_j})$ is the set of observation indices that are matched to observation i of the opposite treatment condition. The observed covariate and matched covariate pairs, $(z_i, \dot{z}_i), i = 1, \ldots, N$, 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 [CAUSAL] **teffects nnmatch**, 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

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

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

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 $\hat{\sigma}^2_{\dot{z}}(t)$ for $\hat{\sigma}^2_z(t)$.

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

$$\tilde{\mu}_{\dot{z}}(t) = \frac{\sum_i^N I(t_i = t) w_i z_i K_m(i)}{M_t} \quad \text{and}$$

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

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

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

We will simplify notation by defining a single weight

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

The covariate mean and variance for treatment level t are

$$\tilde{\mu}_{\dot{z}}(t) = \frac{\sum_i^N I(t_i = t) w_i \overline{w}_i x_i}{M_t} \qquad \text{and}$$

$$\tilde{\sigma}^2_{\dot{z}}(t) = \frac{I(t_i = t) w_i \overline{w}_i \left\{z_i - \tilde{\mu}_{\dot{z}}(t)\right\}^2}{M_t - 1}$$

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

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

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 [CAUSAL] **teffects aipw** and *Methods and formulas* of [CAUSAL] **stteffects ipwra**, we have the first-order condition of the treatment model

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{s}_{\text{tm},i}(\mathbf{z}_i, \widehat{\boldsymbol{\gamma}}) = 0$$

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

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

The score reduces to

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

The corresponding covariate balancing moment conditions are

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

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

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

We stack the moment conditions

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

The overidentified GMM estimator is then

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

where

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

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

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

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

and is asymptotically distributed χ^2 with degrees of freedom d,

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

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. https://doi.org/10.1002/sim.3697.

———. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46: 399–424. https://doi.org/10.1080/00273171.2011.568786.

Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154. https://doi.org/10.1016/j.jeconom.2009.09.023.

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*, B ser., 76: 243–263. https://doi.org/10.1111/rssb.12027.

Jann, B. 2021. Relative distribution analysis in Stata. *Stata Journal* 21: 885–951.

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. https://doi.org/10.2307/2683903.

Rubin, D. B. 2008. For objective causal inference, design trumps analysis. *Annals of Applied Statistics* 2: 808–840. https://doi.org/10.1214/08-AOAS187.

Also see

[CAUSAL] **stteffects intro** — Introduction to treatment effects for observational survival-time data

[CAUSAL] **tebalance box** — Covariate balance box

[CAUSAL] **tebalance density** — Covariate balance density

[CAUSAL] **tebalance overid** — Test for covariate balance

[CAUSAL] **tebalance summarize** — Covariate-balance summary statistics

[CAUSAL] **teffects** — Treatment-effects estimation for observational data