**teffects ipwra — Inverse-probability-weighted regression adjustment**

**Description**

teffects ipwra estimates the average treatment effect (ATE), the average treatment effect on the treated (ATET), and the potential-outcome means (POMs) from observational data by inverse-probability-weighted regression adjustment (IPWRA). IPWRA estimators use weighted regression coefficients to compute averages of treatment-level predicted outcomes, where the weights are the estimated inverse probabilities of treatment. The contrasts of these averages estimate the treatment effects. IPWRA estimators have the double-robust property. teffects ipwra accepts a continuous, binary, count, fractional, or nonnegative outcome and allows a multivalued treatment.

See [TE] teffects intro or [TE] teffects intro advanced for more information about estimating treatment effects from observational data.

**Quick start**

ATE of binary treatment treat2 estimated by IPWRA using a linear model for outcome y1 on x1 and x2 and a logistic model for treat2 on x1 and w

```
teffects ipwra (y1 x1 x2) (treat2 x1 w)
```

As above, but estimate the ATET

```
teffects ipwra (y1 x1 x2) (treat2 x1 w), atet
```

Probit model for binary outcome y3

```
teffects ipwra (y3 x1 x2, probit) (treat2 x1 w)
```

As above, but use a heteroskedastic probit model for y3 and a probit model for treat2

```
teffects ipwra (y3 x1 x2, hetprobit(x1 x2)) (treat2 x1 w, probit)
```

As above, but use a fractional heteroskedastic probit model for y4 and a probit model for treat2

```
teffects ipwra (y4 x1 x2, fhetprobit(x1 x2)) (treat2 x1 w, probit)
```

ATE for each level of a three-valued treatment treat3

```
teffects ipwra (y1 x1 x2) (treat3 x1 w)
```

As above, and specify that treat3 = 3 is the control level using the value label “MyControl” for 3

```
teffects ipwra (y1 x1 x2) (treat3 x1 w), control(MyControl)
```

**Menu**

Statistics > Treatment effects > Continuous outcomes > Regression adjustment with IPW

Statistics > Treatment effects > Binary outcomes > Regression adjustment with IPW

Statistics > Treatment effects > Count outcomes > Regression adjustment with IPW

Statistics > Treatment effects > Fractional outcomes > Regression adjustment with IPW

Statistics > Treatment effects > Nonnegative outcomes > Regression adjustment with IPW
Syntax

tffects ipwra (ovar omvarlist [ , omodel noconstant ])
(tvar tmvarlist [ , tmodel noconstant ]) [ if ] [ in ] [ weight ]
[ , stat options ]

.ovar is a binary, count, continuous, fractional, or nonnegative outcome of interest.
.omvarlist specifies the covariates in the outcome model.
.tvar must contain integer values representing the treatment levels.
.tmvarlist specifies the covariates in the treatment-assignment model.

omodel Description
---
Model
linear linear outcome model; the default
logit logistic outcome model
probit probit outcome model
hetprobit(varlist) heteroskedastic probit outcome model
poisson exponential outcome model
flogit fractional logistic outcome model
fprobit fractional probit outcome model
fhetprobit(varlist) fractional heteroskedastic probit outcome model

.omodel specifies the model for the outcome variable.

tmodel Description
---
Model
logit logistic treatment model; the default
probit probit treatment model
hetprobit(varlist) heteroskedastic probit treatment model

t.model specifies the model for the treatment variable.
For multivalued treatments, only logit is available and multinomial logit is used.

stat Description
---
Stat
ate estimate average treatment effect in population; the default
atet estimate average treatment effect on the treated
pomeans estimate potential-outcome means
# teffects ipwra — Inverse-probability-weighted regression adjustment

## Options

<table>
<thead>
<tr>
<th>SE/Robust</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>vce(vcetype)</strong></td>
<td>vcetype may be robust, cluster clustvar, bootstrap, or jackknife</td>
</tr>
</tbody>
</table>

### Reporting

<table>
<thead>
<tr>
<th>Reporting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>level(#)</strong></td>
<td>set confidence level; default is level(95)</td>
</tr>
<tr>
<td><strong>aequations</strong></td>
<td>display auxiliary-equation results</td>
</tr>
<tr>
<td><strong>display_options</strong></td>
<td>control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling</td>
</tr>
</tbody>
</table>

### Maximization

<table>
<thead>
<tr>
<th>Maximization</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>maximize_options</strong></td>
<td>control the maximization process; seldom used</td>
</tr>
</tbody>
</table>

### Advanced

<table>
<thead>
<tr>
<th>Advanced</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pstolerance(#)</strong></td>
<td>set tolerance for overlap assumption</td>
</tr>
<tr>
<td><strong>osample(newvar)</strong></td>
<td>newvar identifies observations that violate the overlap assumption</td>
</tr>
<tr>
<td>**control(#</td>
<td>label)**</td>
</tr>
<tr>
<td>**tlevel(#</td>
<td>label)**</td>
</tr>
<tr>
<td><strong>coeflegend</strong></td>
<td>display legend instead of statistics</td>
</tr>
</tbody>
</table>

*omvarlist* and *tmvarlist* may contain factor variables; see [U] 11.4.3 Factor variables.

bootstrap, by, jackknife, and statsby are allowed; see [U] 11.1.10 Prefix commands.

Weights are not allowed with the bootstrap prefix; see [R] bootstrap.

fweights, iweights, and pweights are allowed; see [U] 11.1.6 weight.

coeflegend does not appear in the dialog box.

See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.

## Options

### Model

- **noconstant**: see [R] Estimation options.

### Stat

- *stat* is one of three statistics: *ate*, *atet*, or *pomeans*. *ate* is the default.
  - **ate**: specifies that the average treatment effect be estimated.
  - **atet**: specifies that the average treatment effect on the treated be estimated.
  - **pomeans**: specifies that the potential-outcome means for each treatment level be estimated.

### SE/Robust

- **vce(vcetype)** specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (robust), that allow for intragroup correlation (cluster clustvar), and that use bootstrap or jackknife methods (bootstrap, jackknife); see [R] vce_option.
Reporting

level(##); see [R] Estimation options.

eaequations specifies that the results for the outcome-model or the treatment-model parameters be displayed. By default, the results for these auxiliary parameters are not displayed.

display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels, allbaselevels, nofvlvalue, fwrap(##), fwrapon(style), cformat(%fmt), pformat(%fmt), sformat(%fmt), and nolstretch; see [R] Estimation options.

Maximization

maximize_options: iterate(#), [no]log, and from(init_specs); see [R] Maximize. These options are seldom used.

init_specs is one of

matname [ , skip copy]

# [ , # ...], copy

Advanced

pstolerance(#) specifies the tolerance used to check the overlap assumption. The default value is pstolerance(1e-5). teffects will exit with an error if an observation has an estimated propensity score smaller than that specified by pstolerance().

osample(newvar) specifies that indicator variable newvar be created to identify observations that violate the overlap assumption.

ccontrol(# | label) specifies the level of tvar that is the control. The default is the first treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. control() may not be specified with statistic pomeans. control() and tlevel() may not specify the same treatment level.

tlevel(# | label) specifies the level of tvar that is the treatment for the statistic atet. The default is the second treatment level. You may specify the numeric level # (a nonnegative integer) or the label associated with the numeric level. tlevel() may only be specified with statistic atet. tlevel() and control() may not specify the same treatment level.

The following option is available with teffects ipwra but is not shown in the dialog box:

coefflegend; see [R] Estimation options.

Remarks and examples

Remarks are presented under the following headings:

Overview

Video example

Overview

IPWRA estimators use probability weights to obtain outcome-regression parameters that account for the missing-data problem arising from the fact that each subject is observed in only one of the potential outcomes. The adjusted outcome-regression parameters are used to compute averages of treatment-level predicted outcomes. The contrasts of these averages provide estimates of the treatment effects.
IPWRA estimators use a model to predict treatment status, and they use another model to predict outcomes. Because IPWRA estimators have the double-robust property, only one of the two models must be correctly specified for the IPWRA estimator to be consistent.

IPWRA estimators use a three-step approach to estimating treatment effects:

1. They estimate the parameters of the treatment model and compute inverse-probability weights.
2. Using the estimated inverse-probability weights, they fit weighted regression models of the outcome for each treatment level and obtain the treatment-specific predicted outcomes for each subject.
3. They compute the means of the treatment-specific predicted outcomes. The contrasts of these averages provide the estimates of the ATEs. By restricting the computations of the means to the subset of treated subjects, we can obtain the ATETs.

These steps produce consistent estimates of the effect parameters because the treatment is assumed to be independent of the potential outcomes after conditioning on the covariates. The overlap assumption ensures that predicted inverse-probability weights do not get too large. The standard errors reported by `teffects ipwra` correct for the three-step process. See [TE] `teffects intro` or [TE] `teffects intro advanced` for more information about this estimator.

We will illustrate the use of `teffects ipwra` by using data from a study of the effect of a mother’s smoking status during pregnancy (`mbsmoke`) on infant birthweight (`bweight`) as reported by Cattaneo (2010). This dataset also contains information about each mother’s age (`mage`), education level (`medu`), marital status (`mmarried`), whether the first prenatal exam occurred in the first trimester (`prenatal1`), and whether this baby was the mother’s first birth (`fbaby`).

Example 1: Estimating the ATE

We begin by using `teffects ipwra` to estimate the average treatment effect of smoking on birthweight. We will use a probit model to predict treatment status as a function of `mmarried`, `mage`, and `fbaby`; to maximize the predictive power of this model, we use factor-variable notation to incorporate quadratic effects of the mother’s age, the only continuous covariate in our model. We will use linear regression (the default) to model birthweight, using `prenatal1`, `mmarried`, `mage`, and `fbaby` as explanatory variables. We type
. use https://www.stata-press.com/data/r16/cattaneo2
. teffects ipwra (bweight prenatal1 mmarried mage fbaby)
   > (mbsmoke mmarried c.mage##c.mage fbaby medu, probit)
Iteration 0:  EE criterion = 9.885e-21
Iteration 1:  EE criterion = 7.847e-26
Treatment-effects estimation
Number of obs = 4,642
Estimator : IPW regression adjustment
Outcome model : linear
Treatment model: probit

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
<th></th>
<th></th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
</tr>
<tr>
<td>ATE</td>
<td>mbsmoke (smoker vs nonsmoker)</td>
<td>-229.9671</td>
<td>26.62668</td>
<td>-8.64</td>
</tr>
<tr>
<td>POmean</td>
<td>mbsmoke nonsmoker</td>
<td>3403.336</td>
<td>9.57126</td>
<td>355.58</td>
</tr>
</tbody>
</table>

The average birthweight if all mothers were to smoke would be 230 grams less than the average of 3,403 grams that would occur if none of the mothers had smoked.

By default, *teffects ipwra* displays the ATE and untreated POM. We can specify the *pomeans* option to display both the treated and untreated POMs, and we can use the *aequations* option to display the regression model coefficients used to predict the POMs as well as the coefficients from the model used to predict treatment.
As is well known, the standard probit model assumes that the error terms in the latent-utility framework are homoskedastic; the model is not robust to departures from this assumption. An alternative is to use the heteroskedastic probit model, which explicitly models the error variance as a function of a set of variables.

Example 3: Heteroskedastic probit treatment model

Here we use the variables as before, but we use a heteroskedastic probit model to predict treatment status, modeling the heteroskedasticity as a quadratic function of the mother’s age:
. teffects ipwra (bweight prenatal1 mmarried fbaby c.mage),
> (mbsmoke mmarried c.mage##c.mage fbaby medu, hetprobit(c.mage##c.mage)),
> aequations
Iteration 0:   EE criterion = 4.443e-09
Iteration 1:   EE criterion = 4.325e-18
Treatment-effects estimation  Number of obs   =   4,642
Estimator : IPW regression adjustment
Outcome model : linear
Treatment model: heteroskedastic probit

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
</tr>
<tr>
<td>ATE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nonsmoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-229.6322</td>
<td>26.33452</td>
<td>-8.72</td>
<td>0.000</td>
</tr>
<tr>
<td>POmean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nonsmoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3403.74</td>
<td>9.545798</td>
<td>356.57</td>
<td>0.000</td>
</tr>
<tr>
<td>OME0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prenatal1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmarried</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64.95125</td>
<td>28.62159</td>
<td>2.27</td>
<td>0.023</td>
</tr>
<tr>
<td>fbaby</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-71.61131</td>
<td>20.33774</td>
<td>-3.52</td>
<td>0.000</td>
</tr>
<tr>
<td>mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.010148</td>
<td>2.133812</td>
<td>1.41</td>
<td>0.158</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3195.355</td>
<td>55.05451</td>
<td>58.04</td>
<td>0.000</td>
</tr>
<tr>
<td>OME1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prenatal1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmarried</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>38.55274</td>
<td>43.57024</td>
<td>0.88</td>
<td>0.376</td>
</tr>
<tr>
<td>fbaby</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-6.069913</td>
<td>5.95251</td>
<td>-1.02</td>
<td>0.308</td>
</tr>
<tr>
<td>mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.029552</td>
<td>0.036895</td>
<td>-1.24</td>
<td>0.216</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3195.795</td>
<td>152.3979</td>
<td>20.97</td>
<td>0.000</td>
</tr>
<tr>
<td>TME1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mmarried</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.093306</td>
<td>0.080003</td>
<td>-1.17</td>
<td>0.244</td>
</tr>
<tr>
<td>mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.003677</td>
<td>0.003017</td>
<td>-1.21</td>
<td>0.225</td>
</tr>
<tr>
<td>c.mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0002837</td>
<td>0.0001901</td>
<td>-1.49</td>
<td>0.136</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.1822201</td>
<td>0.1180587</td>
<td>-1.54</td>
<td>0.123</td>
</tr>
<tr>
<td>TME1_lnsigma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.2211492</td>
<td>0.0631504</td>
<td>-3.50</td>
<td>0.000</td>
</tr>
<tr>
<td>c.mage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0037613</td>
<td>0.0012437</td>
<td>3.02</td>
<td>0.002</td>
</tr>
<tr>
<td>c.mage#</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The estimated ATE and base-level POM are essentially the same as those produced by the model that used a homoskedastic probit.
Video example

Treatment effects: Inverse-probability-weighted regression adjustment

Stored results

teffects ipwra stores the following in e():

Scalars

- e(N) number of observations
- e(nj) number of observations for treatment level j
- e(N_clust) number of clusters
- e(k_eq) number of equations in e(b)
- e(k_levels) number of levels in treatment variable
- e(treated) level of treatment variable defined as treated
- e(control) level of treatment variable defined as control
- e(converged) 1 if converged, 0 otherwise

Macros

- e(cmd) teffects
- e(cmdline) command as typed
- e(depvar) name of outcome variable
- e(tvar) name of treatment variable
- e(subcmd) ipwra
- e(tmodel) logit, probit, or hetprobit
- e(omodel) linear, logit, probit, hetprobit, poisson, flogit, fpobit, or fhetprobit
- e(stat) statistic estimated, ate, atet, or pomeans
- e(wtype) weight type
- e(wexp) weight expression
- e(title) title in estimation output
- e(tlevels) levels of treatment variable
- e(vce) vcetype specified in vce()
- e(vctype) title used to label Std. Err.
- e(properties) b V
- e(estat_cmd) program used to implement estat
- e(predict) program used to implement predict
- e(marginsnotok) predictions disallowed by margins
- e(asbalanced) factor variables fvset as asbalanced
- e(asobserved) factor variables fvset as asobserved

Matrices

- e(b) coefficient vector
- e(V) variance–covariance matrix of the estimators

Functions

- e(sample) marks estimation sample

Methods and formulas

teffects ipwra implements a smooth treatment-effects estimator. All smooth treatment-effects estimators are documented in Methods and formulas of [TE] teffects aipw.

References


Also see

[TE] teffects postestimation — Postestimation tools for teffects
[TE] teffects — Treatment-effects estimation for observational data
[TE] teffects aipw — Augmented inverse-probability weighting
[U] 20 Estimation and postestimation commands