Title

stata.com

teffects ipwra — Inverse-probability-weighted regression adjustment

Syntax

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

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

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

tmodel specifies the model for the treatment variable.
For multivariate 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

options	Description
SE/Robust	
`vce(`*vcetype*`)`	*vcetype* may be `robust`, `bootstrap`, or `jackknife`
Reporting	
`level(#)`	set confidence level; default is `level(95)`
`aequations`	display auxiliary-equation results
display_options	control column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
Advanced	
`pstolerance(#)`	set tolerance for overlap assumption
`osample(`*newvar*`)`	*newvar* identifies observations that violate the overlap assumption
`control(#` | *label*`)`	specify the level of *tvar* that is the control
`tlevel(#` | *label*`)`	specify the level of *tvar* that is the treatment
`coeflegend`	display legend instead of statistics

tvar must contain integer values representing the treatment levels.

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**.

`fweights` and `iweights` 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.

Menu

Statistics > Treatment effects > Doubly robust estimators > Regress adjustment and IPW

Description

`teffects ipwra` estimates treatment effects from observational data by using inverse-probability-weighted regression-adjustment (IPWRA) estimators. 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 provide the estimated treatment effects. IPWRA estimators have the double-robust property. `teffects ipwra` offers several choices for the functional forms of the outcome model and of the treatment model.

See [TE] **teffects intro** or [TE] **teffects intro advanced** for more information about these estimators.

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`) and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [R] ***vce_option***.

Reporting

`level(#)`; see [R] **estimation options**.

`aequations` 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: `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(`*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.

`control(#|`*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:

`coeflegend`; see [R] **estimation options**.

Remarks and examples

stata.com

Remarks are presented under the following headings:

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 http://www.stata-press.com/data/r13/cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)
. teffects ipwra (bweight prenatal1 mmarried mage fbaby)
> (mbsmoke mmarried c.mage##c.mage fbaby medu, probit)
Iteration 0:   EE criterion =  9.482e-21
Iteration 1:   EE criterion =  1.200e-25
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: probit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-229.9671	26.62668	-8.64	0.000	-282.1544	-177.7798
POmean						
mbsmoke nonsmoker	3403.336	9.57126	355.58	0.000	3384.576	3422.095

Smoking causes infants' birthweights to be reduced by an average of 230 grams from the average of 3,403 grams for infants of mothers who do not smoke.

◁

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.

Example 2: Displaying the POMs and equations

```
. use http://www.stata-press.com/data/r13/cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)
. teffects ipwra (bweight prenatal1 mmarried mage fbaby)
> (mbsmoke mmarried c.mage##c.mage fbaby medu, probit), pomeans aequations
Iteration 0:   EE criterion =  9.482e-21
Iteration 1:   EE criterion =  1.157e-25
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: probit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
POmeans						
mbsmoke						
nonsmoker	3403.336	9.57126	355.58	0.000	3384.576	3422.095
smoker	3173.369	24.86997	127.60	0.000	3124.624	3222.113
OME0						
prenatal1	67.98549	28.78428	2.36	0.018	11.56933	124.4017
mmarried	155.5893	26.46903	5.88	0.000	103.711	207.4677
mage	2.893051	2.134788	1.36	0.175	-1.291056	7.077158
fbaby	-71.9215	20.39317	-3.53	0.000	-111.8914	-31.95162
_cons	3194.808	55.04911	58.04	0.000	3086.913	3302.702
OME1						
prenatal1	34.76923	43.18534	0.81	0.421	-49.87248	119.4109
mmarried	124.0941	40.29775	3.08	0.002	45.11193	203.0762
mage	-5.068833	5.954425	-0.85	0.395	-16.73929	6.601626
fbaby	39.89692	56.82072	0.70	0.483	-71.46966	151.2635
_cons	3175.551	153.8312	20.64	0.000	2874.047	3477.054
TME1						
mmarried	-.6484821	.0554173	-11.70	0.000	-.757098	-.5398663
mage	.1744327	.0363718	4.80	0.000	.1031452	.2457202
c.mage#						
c.mage	-.0032559	.0006678	-4.88	0.000	-.0045647	-.0019471
fbaby	-.2175962	.0495604	-4.39	0.000	-.3147328	-.1204595
medu	-.0863631	.0100148	-8.62	0.000	-.1059917	-.0667345
_cons	-1.558255	.4639691	-3.36	0.001	-2.467618	-.6488926

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 =  1.312e-10
Iteration 1:   EE criterion =  1.526e-20

Treatment-effects estimation                    Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: heteroskedastic probit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-229.6322	26.33453	-8.72	0.000	-281.2469	-178.0174
POmean						
mbsmoke nonsmoker	3403.74	9.545798	356.57	0.000	3385.03	3422.449
OME0						
prenatal1	64.95125	28.62162	2.27	0.023	8.853899	121.0486
mmarried	154.2297	26.45868	5.83	0.000	102.3717	206.0878
fbaby	-71.61131	20.33774	-3.52	0.000	-111.4725	-31.75006
mage	3.010148	2.133812	1.41	0.158	-1.172047	7.192344
_cons	3195.355	55.05451	58.04	0.000	3087.45	3303.26
OME1						
prenatal1	38.55274	43.57026	0.88	0.376	-46.8434	123.9489
mmarried	126.3377	40.7398	3.10	0.002	46.48922	206.1863
fbaby	45.43547	56.44832	0.80	0.421	-65.20121	156.0721
mage	-6.069913	5.952513	-1.02	0.308	-17.73662	5.596798
_cons	3195.795	152.3979	20.97	0.000	2897.101	3494.49
TME1						
mmarried	-.0295523	.0238304	-1.24	0.215	-.0762589	.0171544
mage	.0157893	.0105203	1.50	0.133	-.0048302	.0364088
c.mage# c.mage	-.0002837	.0001896	-1.50	0.135	-.0006553	.0000878
fbaby	-.0093306	.0079804	-1.17	0.242	-.024972	.0063107
medu	-.0036773	.0030235	-1.22	0.224	-.0096033	.0022487
_cons	-.1822201	.1177454	-1.55	0.122	-.4129969	.0485567
TME1_lnsigma						
mage	-.2211492	.0629973	-3.51	0.000	-.3446217	-.0976767
c.mage# c.mage	.0037613	.0012413	3.03	0.002	.0013285	.0061942

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 weights with regression adjustment

Stored results

`teffects ipwra` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(n`*j*`)`	number of observations for treatment level *j*
`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`, or `poisson`
`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(vcetype)`	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`

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**.

Reference

Cattaneo, M. D. 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154.

Also see

[TE] **teffects postestimation** — Postestimation tools for teffects

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

[TE] **teffects aipw** — Augmented inverse-probability weighting

[R] **heckman** — Heckman selection model

[R] **hetprobit** — Heteroskedastic probit model

[R] **logit** — Logistic regression, reporting coefficients

[R] **mlogit** — Multinomial (polytomous) logistic regression

[R] **poisson** — Poisson regression

[R] **probit** — Probit regression

[R] **regress** — Linear regression

[U] **20 Estimation and postestimation commands**