lateffects — Local average treatment-effect estimation[+]

[+]This command is part of StataNow.

Description

`lateffects` estimates local average treatment effects (LATEs) by using weighting estimators. Outcomes may be continuous, binary, count, or fractional. The treatment is binary. `lateffects` provides three estimators: inverse-probability-weighted regression adjustment (IPWRA), normalized kappa, and normalized covariate balancing.

Quick start

LATE for outcome `y`, with treatment `t`, and instrument propensity score for `z` modeled using `x1` and `x2` via the normalized kappa estimator

```
lateffects kappa (y) (t) (z x1 x2)
```

Same as above, but use a normalized covariate-balancing estimator

```
lateffects balancing (y) (t) (z x1 x2)
```

LATE of treatment `t` via IPWRA estimation using a linear model for outcome `y` on `x1` and `x2`, a logistic model for `t` on `x1` and `w`, and a logistic model for the instrument propensity score of `z` on `x1`, `x2`, and `w`

```
lateffects ipwra (y x1 x2) (t x1 w) (z x1 x2 w)
```

Same as above, but use probit models for the treatment and the instrument propensity score

```
lateffects ipwra (y x1 x2) (t x1 w, probit) (z x1 x2 w, probit)
```

Menu

Statistics > Causal inference/treatment effects > Continuous outcomes > Local average treatment effects

Statistics > Causal inference/treatment effects > Binary outcomes > Local average treatment effects

Statistics > Causal inference/treatment effects > Count outcomes > Local average treatment effects

Statistics > Causal inference/treatment effects > Fractional outcomes > Local average treatment effects

Statistics > Causal inference/treatment effects > Nonnegative outcomes > Local average treatment effects

Syntax

Inverse-probability-weighted regression adjustment

`lateffects ipwra` (*ovar* [*omvarlist* , *omodel* `noconstant`])
(*tvar* [*tmvarlist* , *tmodel* `noconstant`])
(*iv* [*ivvarlist* , *ivmodel* `noconstant`]) [*if*] [*in*] [*weight*] [, *options*]

Normalized kappa

`lateffects kappa` (*ovar*) (*tvar*)
(*iv* [*ivvarlist* , *ivmodel* `noconstant`]) [*if*] [*in*] [*weight*] [, *options*]

Normalized covariate balancing

`lateffects balancing` (*ovar*) (*tvar*)
(*iv* [*ivvarlist* , *ivmodel* `noconstant`]) [*if*] [*in*] [*weight*] [, *options*]

ovar is a binary, count, continuous, fractional, or nonnegative outcome of interest.

omvarlist specifies the covariates in the outcome model.

tvar is a binary variable indicating observed treatment status.

tmvarlist specifies the covariates for the treatment-status model.

iv is a binary instrumental variable indicating treatment assignment.

ivvarlist specifies covariates for the instrument propensity-score model.

omodel	Description
Model	
`linear`	linear outcome model; the default
`logit`	logistic outcome model
`probit`	probit outcome model
`poisson`	count outcome model
`flogit`	fractional logistic outcome model
`fprobit`	fractional probit outcome model

omodel specifies the model for the outcome variable. *omodel* is available only with `ipwra`.

tmodel	Description
Model	
`logit`	logistic treatment-status model; the default
`probit`	probit treatment-status model

tmodel specifies the model for treatment status. *tmodel* is available only with `ipwra`.

ivmodel	Description
Model	
logit	logistic instrument propensity-score model; the default
probit	probit instrument propensity-score model

ivmodel specifies the model for the instrument propensity score.

options	Description
SE/Robust	
vce(*vcetype*)	*vcetype* may be robust, cluster *clustvar*, bootstrap, or jackknife
Reporting	
level(#)	set confidence level; default is level(95)
aequations	display auxiliary-equation results
display_options	control columns and 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
coeflegend	display legend instead of statistics

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

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

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

aweights, 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

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**.

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: noci, nopvalues, 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.

Advanced

pstolerance(#) specifies the tolerance used to check the overlap assumption. The default value is pstolerance(1e-5). lateffects will exit with an error if an observation has an estimated propensity score smaller than # or larger than $1 - \#$.

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

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

coeflegend; see [R] **Estimation options**.

Remarks and examples

A local average treatment effect (LATE) is an average treatment effect for a subpopulation. We would usually prefer to identify an effect for the entire population, but in many instances, this is not feasible. In the LATE framework, we cannot identify a treatment effect for the population because of unobservable differences between treated units and untreated units. Unaccounted for, unobservable differences confound any causal effect we would like to identify (see [CAUSAL] **Intro**). Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996) illustrate how, if there is a binary variable that splits the population into treated and untreated units, as if they were randomly assigned, we may identify a treatment effect for those that comply with the treatment assignment, a LATE. Because the effect is identified only for compliers, sometimes the estimand is referred to as the complier average treatment effect.

In an experimental setting, where units are assigned to treatment or control, “compliance” is readily understood as a unit’s adherence to the treatment status to which it was assigned, for instance, those assigned to be treated who take the medicine and those assigned to be controls who take the placebo. In a context of observational data, however, units often self-select into their treatment status, and the treatment-assignment variable of interest can only encourage, persuade, or otherwise motivate units into a treatment status. Some units, the “always takers”, will opt for the treatment regardless of whether they receive the motivation. Others, the “never takers”, will always opt out of the treatment. “Defiers” will opt out of treatment when motivated and opt in when motivation is absent. Finally, “compliers” opt for the treatment when they are motivated and opt out when they are not; this is the target population of the LATE estimate, which we now define.

In the LATE framework, treatment can be thought of as a potential outcome. To exemplify, let d denote binary treatment and let z be the binary instrument that assigns or motivates units into their treatment status. We can express treatment as

$$d = d(1)z + d(0)(1 - z)$$

That is, we observe $d(1)$ if $z = 1$ and $d(0)$ if $z = 0$. $d(1)$ might be 1 or it might be 0. Realizations of the random variable d for which $d(1) = 1$ when $z = 1$ can arise from two types of behavior. They could be those that always opt into treatment so that $d(1) = 1$ regardless of z. Alternatively, they could be realizations for which $d(1) = 1$ only when $z = 1$. Similarly, realizations for which $d(0) = 0$ when $z = 0$ either never opt into treatment (so that $d(0) = 0$ regardless of z) or opt out only when not assigned to treatment. In this framework, compliers are those that modify their behavior based on treatment assignment. The instrument affects their decisions. For them, $d(1) = 1$ when $z = 1$ and $d(0) = 0$ when $z = 0$, or equivalently, $d(1) > d(0)$.

The LATE estimand, in the absence of covariates, is defined by

$$\tau_{\text{LATE}} = \frac{E(y|z=1) - E(y|z=0)}{E(d=1|z=1) - E(d=1|z=0)}$$

which, under the assumptions in Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996), is proved to be a treatment effect conditional on compliance:

$$\tau_{\text{LATE}} = \frac{E(y|z=1) - E(y|z=0)}{E(d=1|z=1) - E(d=1|z=0)} = E\{y(1) - y(0)\,|d(1) > d(0)\}$$

where y corresponds to the outcome of interest, $y(1)$ is the potential outcome when a unit is treated, and $y(0)$ is the potential outcome when a unit is untreated.

Without covariates, one can estimate the expression above using two-stage least squares with `ivregress`. Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996) illustrate how the LATE estimand derived from a potential-outcomes framework, as described in [CAUSAL] **Intro**, is equivalent to the instrumental-variables regression, which is usually thought of from a simultaneous-equation perspective. In the LATE framework, z satisfies the assumptions of an instrumental variable.

Recent literature (see Słoczyński [2020] and Blandhol et al. [2022]) suggests that, when controlling for covariates, using two-stage least squares does not necessarily lead to an estimate of τ_{LATE}. Słoczyński, Uysal, and Wooldridge (2022, 2025) propose weighting estimators that can be interpreted as estimates of

$$\tau_{\text{LATE}} = \frac{E_x\left\{E(y|z=1,\mathbf{x}) - E(y|z=0,\mathbf{x})\right\}}{E_x\left\{E(d=1|z=1,\mathbf{x}) - E(d=1|z=0,\mathbf{x})\right\}}$$

where $\mathbf{x}$ is a random vector of covariates and E_x denotes an expectation with respect to the covariates. The `lateffects` command provides three estimators that construct weighted estimates for the four conditional expectations above and then average over the data. The weights are constructed so that the instrument, once we condition on the covariates, is as good as if it were randomly assigned.

You can choose from the IPWRA, normalized kappa, or normalized covariate-balancing estimator. IPWRA models the outcome, the treatment, and the instrument. The outcome may be modeled using linear, logistic, probit, Poisson, fractional logistic, or fractional probit regression. The treatment and instrument may be modeled using logistic or probit regression. The normalized kappa estimator models only the instrument using logistic or probit regression to estimate a propensity score for the instrument. The normalized covariate-balancing estimator also models only the instrument using logistic or probit regression, but this estimator forces covariate balance through additional moment conditions. See *Methods and formulas* for a description of these estimators.

For all three estimators, we require conventional identification assumptions for treatment effects: the stable unit treatment value assumption (SUTVA), the unconfoundedness (conditional-independence) assumption, and the overlap assumption. See [CAUSAL] **Intro** for descriptions of these assumptions. Note that, for a LATE, these assumptions are defined with respect to the propensity score of the instrument rather than the propensity score of the treatment indicator. We also require that the instrument affects the outcome only through treatment, an instrumental-variable exclusion restriction. Additionally, we must guarantee that there are no defiers, understood as individuals that would act contrary to their treatment assignment.

Below, we illustrate how to use the `lateffects` command to obtain LATE estimates and how to check for violations of testable assumptions. For more discussion on LATE estimation and theory, see Imbens and Angrist (1994); Angrist, Imbens, and Rubin (1996); Angrist and Pischke (2009, 2015); and Imbens and Rubin (2015) and the references therein.

▷ Example 1: LATE estimation for labor outcomes

We consider the effect of schooling on wages. Schooling is an individual choice that is not independent of unobservable characteristics that affect wages, such as ability. In other words, schooling is endogenous. To address the endogeneity problem, Card (1995) proposes proximity to a four-year college as an instrument. He argues that proximity to a four-year college affects wages only through the effect it has on schooling decisions. For instance, he suggests that living in a college town lowers the cost of attending college, making the decision to pursue additional schooling more likely.

We use data provided by Słoczyński, Uysal, and Wooldridge (2025) to revisit part of the analysis in Card (1995). We study the effect of having some post-secondary education (`somecol`) on the log of hourly wages (`lwage`). In this context, `somecol` is considered the treatment, while living near a four-year college (`nearc4`) is the instrumental variable acting as the treatment assignment. To build some intuition, we first use the normalized kappa estimator without covariates. We type

```
. use https://www.stata-press.com/data/r19/card95
(National Longitudinal Survey of Young Men)
```

In the `lateffects` command, we specify a first set of parentheses with a model for the outcome and a second set with a model for the treatment status. The kappa estimator computes weighted means for both models, where the weights are constructed using instrument propensity scores. In the third set of parentheses, we specify the instrument propensity-score model. Because we have no covariates, the propensity score is just the mean of `nearc4`.

```
. lateffects kappa (lwage) (somecol) (nearc4)
Iteration 0:  EE criterion = 6.298e-30
Iteration 1:  EE criterion = 3.286e-30
Local average treatment effect                        Number of obs = 3,010
Estimator:         Normalized kappa
Outcome model:    Weighted Mean
Treatment model: Weighted Mean
IV pscore model: Logit
```

lwage	Coefficient	Robust std. err.	z	P>|z|	[95% conf. interval]	
LATE						
somecol (Yes vs No)	1.278672	.2203624	5.80	0.000	.8467691	1.710574

In the coefficient table of our output, we see the estimate of the LATE. This is the average treatment effect of pursuing some college education on the log of wages for compliers (for those that will pursue some college education if they live near a four-year college and will not do so otherwise). Thus, for the complier subpopulation, we expect average log wages to be 1.28 higher if everyone has some college education than if no one has college education.

We would obtain the same result using `ivregress`:

```
. ivregress 2sls lwage (somecol=nearc4), vce(robust)
Instrumental-variables 2SLS regression          Number of obs   =      3,010
                                                Wald chi2(1)    =      33.67
                                                Prob > chi2     =     0.0000
                                                Root MSE        =     .69383

------------------------------------------------------------------------------
             |               Robust
       lwage | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
     somecol |   1.278672   .2203624     5.80   0.000     .8467691    1.710574
       _cons |   5.615699   .1117247    50.26   0.000     5.396723    5.834676
------------------------------------------------------------------------------
Endogenous: somecol
Exogenous:  nearc4
```

As was demonstrated in Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996), without covariates, two-stage least squares is LATE.

We now add covariates. In particular, we follow the basic specification in Card (1995). In his specification, he includes binary indicators for race (`black`), living in the south (`south`), living in a standard metropolitan area (`smsa`), and living in a metropolitan area in the initial survey wave (`smsa66`); a categorical variable for the four regions of the United States (`region`); and work experience (`exper`), which enters the model quadratically.

```
. lateffects kappa (lwage) (somecol)
> (nearc4 i.(black south smsa smsa66 region) c.exper##c.exper)
Iteration 0:  EE criterion = 8.456e-20
Iteration 1:  EE criterion = 1.836e-29
Local average treatment effect                         Number of obs = 3,010
Estimator:       Normalized kappa
Outcome model:   Weighted Mean
Treatment model: Weighted Mean
IV pscore model: Logit
---------------------------------------------------------------------------------
                |               Robust
          lwage | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
----------------+----------------------------------------------------------------
LATE            |
        somecol |
     (Yes vs No)|   .3328798   .2237742     1.49   0.137    -.1057095    .7714691
---------------------------------------------------------------------------------
```

The LATE indicates average log wages are 0.33 higher among compliers when they all pursue some college education than when none of them do. If we had used two-stage least squares, the coefficient associated with `somecol` would be 0.69, which, given the presence of covariates, cannot be interpreted as a LATE. Also, the confidence interval for the LATE parameter in the `lateffects` results includes 0, which would not be the case if we had fit the model using `ivregress`.

◁

▷ Example 2: Verifying model assumptions

For us to interpret our estimate as causal, we need to verify that the LATE assumptions hold. Using `latebalance`, we may obtain diagnostics and tests that let us ascertain if the instrument is as good as randomly assigned once we control for the covariates in our model. If this is the case, after weighting with the instrument propensity scores, group characteristics should be equivalent between those assigned to treatment and those assigned to control.

First, we use `latebalance summarize`. The standardized-difference and variance-ratio results are each presented in two columns. The first column corresponds to the raw data, and the second column presents the statistics computed using the instrument propensity-score weights. If the instrument propensity-score weights have balanced the distributions of our covariates, their weighted mean differences should be close to 0 and their weighted variance ratios should be close to 1. The table at the top additionally reports the treated and control sample sizes, which should be similar after weighting.

```
. latebalance summarize
Covariate balance summary
```

Number of observations	Raw	Weighted
Total	3,010	3,010
Assigned to treatment	2,053	1,515.395
Assigned to control	957	1,494.605

	Standardized differences		Variance ratio	
	Raw	Weighted	Raw	Weighted
black				
Yes	-.1586972	.0275062	.8277845	1.037878
south				
Yes	-.484023	-.0130109	.8972617	.995531
smsa				
Yes	.7722047	-.0029765	.5854675	1.002687
smsa66				
Yes	1.079367	-.0064253	.7272248	1.003921
region				
Midwest	.0797295	-.0076143	1.094144	.9917265
South	-.5598127	.002059	.9172754	1.000646
West	.1599008	.0599971	1.501104	1.169778
exper	-.1312633	-.0373534	.9005343	1.038114
exper# exper	-.134808	-.0248842	.7968106	1.008667

The table of diagnostics suggests that the group characteristics have been balanced. For a visual inspection of this balance, we could look at the full distributions of characteristics among controls and treated both in the data and after weighting. We can do this for each covariate. To illustrate with the variable `exper`, we type

```
. latebalance density exper
```

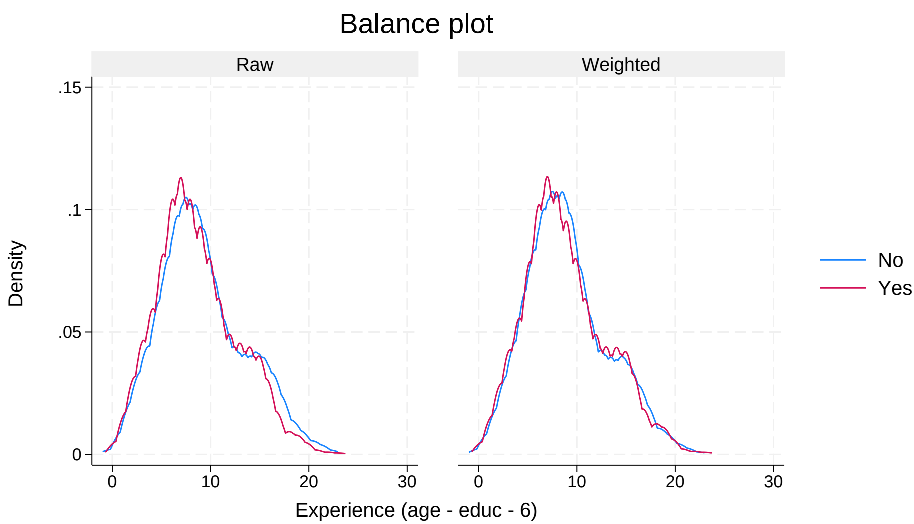

Figure 1.

The distributions of `exper` appear to be more balanced after weighting, as our diagnostics had already suggested.

We can also readily verify the overlap assumption. It states that there is a positive probability of observations being assigned to treatment and control groups once we control for covariates. The `lateoverlap` postestimation command plots estimated densities of the probability of being assigned to treatment or control, allowing us to verify the assumption. A violation of the overlap assumption would be reflected in a plot where observations are bunched toward 0, or toward 1, or toward both extremes, with few treated and control observations sharing the same regions of the support. We type

```
. lateoverlap
```

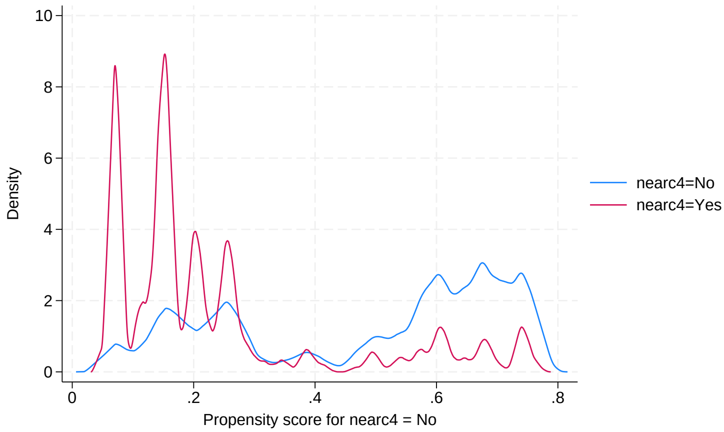

Figure 2.

The graph does not suggest that violations of the overlap assumption are a concern.

◁

Stored results

`lateffects` stores the following in `e()`:

Scalars

`e(N)`	number of observations
`e(rank)`	rank of `e(V)`
`e(N_clust)`	number of clusters
`e(converged)`	`1` if converged, `0` otherwise
`e(late)`	local average treatment-effect estimate

Macros

`e(cmd)`	`lateffects`
`e(cmdline)`	command as typed
`e(depvar)`	name of outcome variable
`e(tvar)`	name of treatment variable
`e(tmodel)`	`logit` or `probit`
`e(omodel)`	`linear`, `logit`, `probit`, `poisson`, `flogit`, or `fprobit`
`e(ivpscmodel)`	`logit` or `probit`
`e(estimator)`	`ipwra`, `kappa`, or `balancing`
`e(wexp)`	weight expression
`e(wtype)`	weight type
`e(title)`	title in estimation output
`e(clustvar)`	name of cluster 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

In addition to the above, the following is stored in r():

Matrices	
r(table)	matrix containing the coefficients with their standard errors, test statistics, p-values, and confidence intervals

Note that results stored in r() are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Below, we provide methods and formulas for the weighting estimators implemented in lateffects. For a detailed discussion of the estimators and the underlying theory, see Słoczyński, Uysal, and Wooldridge (2022, 2025).

The IPWRA, normalized kappa, and normalized covariate-balancing estimators provided by lateffects compute propensity-score weights for cross-sectional data, with observations $i = 1, \dots, N$. The instrument propensity scores estimate the probability of a binary instrument z conditional on covariates. Let $\mathbf{x}_z$ correspond to covariates used to model the instrument propensity scores defined by the expression

$$G\left(\mathbf{x}_{zi}\widehat{\gamma}\right) = \widehat{P}\left(z_i = 1|\mathbf{x}_{zi}\right)$$

$G\left(\mathbf{x}_{zi}\widehat{\gamma}\right)$ is estimated via a probit or logistic regression for the IPWRA and normalized kappa estimators and by using a covariate-balancing propensity score, described below, for the normalized covariate-balancing estimator. $\widehat{\gamma}$ are the parameters fit using these estimators. To simplify notation, we denote $G\left(\mathbf{x}_{zi}\widehat{\gamma}\right) \equiv G_{1i}$. Similarly, $1 - G\left(\mathbf{x}_{zi}\widehat{\gamma}\right) \equiv G_{0i}$.

In what follows, y corresponds to the outcome variable, and d is a binary variable indicating treatment status. A covariate vector used to model the outcome is given by $\mathbf{x}_y$, and a covariate vector used to model treatment status is given by $\mathbf{x}_d$.

The estimators of the LATE defined below are functions of estimates themselves. Thus, they all use gmm to obtain correct standard errors.

Methods and formulas are presented under the following headings:

IPWRA estimator
Normalized kappa estimator
Normalized covariate-balancing estimator

IPWRA estimator

The IPWRA estimator provided by `lateffects ipwra` is computed via the following steps:

1. Compute $\widehat{P}\,(d_i = 1|\mathbf{x}_{di}, z = 0)$ using weights $1/G_{0i}$; denote the estimator $\Lambda(\mathbf{x}_{di}\widehat{\boldsymbol{\delta}}_0) \equiv \Lambda_{0i}$.
2. Compute $\widehat{P}\,(d_i = 1|\mathbf{x}_{di}, z = 1)$ using weights $1/G_{1i}$; denote the estimator $\Lambda(\mathbf{x}_{di}\widehat{\boldsymbol{\delta}}_1) \equiv \Lambda_{1i}$.
3. Compute $\widehat{E}\left(y_i|\mathbf{x}_{yi}, z = 0\right)$ using weights $1/G_{0i}$; denote the estimator $m(\mathbf{x}_{yi}\widehat{\beta}_0) \equiv m_{0i}$.
4. Compute a weighted $\widehat{E}\left(y_i|\mathbf{x}_{yi}, z = 1\right)$ using weights $1/G_{1i}$; denote the estimator $m(\mathbf{x}_{yi}\widehat{\beta}_1) \equiv m_{1i}$.
5. Compute the LATE via $\widehat{\tau}_{\text{LATE}} = \{n^{-1}\sum_{i=1}^{n}(m_{1i} - m_{0i})\}/\{n^{-1}\sum_{i=1}^{n}(\Lambda_{1i} - \Lambda_{0i})\}$.

Normalized kappa estimator

The normalized kappa estimator provided by `lateffects kappa` is given by

$$\widehat{\tau}_{\text{LATE}} = \frac{\left(\sum_{i=1}^{n}\frac{z_i}{G_{1i}}\right)^{-1}\sum_{i=1}^{n}\frac{y_i z_i}{G_{1i}} - \left(\sum_{i=1}^{n}\frac{1-z_i}{G_{0i}}\right)^{-1}\sum_{i=1}^{n}\frac{y_i(1-z_i)}{G_{0i}}}{\left(\sum_{i=1}^{n}\frac{z_i}{G_{1i}}\right)^{-1}\sum_{i=1}^{n}\frac{d_i z_i}{G_{1i}} - \left(\sum_{i=1}^{n}\frac{1-z_i}{G_{0i}}\right)^{-1}\sum_{i=1}^{n}\frac{d_i(1-z_i)}{G_{0i}}}$$

Normalized covariate-balancing estimator

The normalized covariate-balancing estimator provided by `lateffects balancing` is computed using the same formula as the normalized kappa estimator above, but to estimate the instrument propensity scores, the balancing estimator solves the sample analog of the moment condition

$$E\left[\frac{d\mathbf{x}_z}{G(\mathbf{x}_z\boldsymbol{\gamma})} - \frac{(1-d)\,\mathbf{x}_z}{\{1 - G(\mathbf{x}_z\boldsymbol{\gamma})\}}\right] = 0$$

Acknowledgments

We thank Derya Uysal from the Department of Economics at the University of Munich for her presentation on weighting estimators for LATE and their properties at the Stata Economics Virtual Symposium of 2023, Abadie's kappa and weighting estimators of the local average treatment effect. We would also like to thank Tymon Słoczyński from the Department of Economics and International Business School (IBS) at Brandeis University, Derya Uysal, and Jeffrey M. Wooldridge for their `kappalate` command.

References

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91: 444–455. https://doi.org/10.2307/2291629.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

———. 2015. *Mastering Metrics: The Path from Cause to Effect*. Princeton, NJ: Princeton University Press.

Blandhol, C., J. Bonney, M. Mogstad, and A. Torgovitsky. 2022. When is TSLS actually LATE? NBER Working Paper 29709, National Bureau of Economic Research. https://doi.org/10.3386/w29709.

Card, D. 1995. "Using geographic variation in college proximity to estimate the return to schooling". In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, edited by L. N. Christofides, E. K. Grant, and R. Swidinsky, 201–222. Toronto, Canada: University of Toronto Press.

Imbens, G. W., and J. D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62: 467–475. https://doi.org/10.2307/2951620.

Imbens, G. W., and D. B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139025751.

Słoczyński, T. 2020. When should we (not) interpret linear IV estimands as LATE. Working paper, Department of Economics, Brandeis University. https://people.brandeis.edu/~tslocz/Sloczynski_paper_LATE.pdf.

Słoczyński, T., S. D. Uysal, and J. M. Wooldridge. 2022. Doubly robust estimation of local average treatment effects using inverse probability weighted regression adjustment. IZA Discussion Paper 15727, Institute for the Study of Labor (IZA). https://docs.iza.org/dp15727.pdf.

———. 2025. Abadie's kappa and weighting estimators of the local average treatment effect. *Journal of Business and Economic Statistics* 43: 164–177. https://doi.org/10.1080/07350015.2024.2332763.

Also see

[CAUSAL] **lateffects postestimation** — Postestimation tools for lateffects+

[U] 20 Estimation and postestimation commands