help hte

Title

hte — Heterogeneous Treatment Effect Analysis

Syntax

    hte  depvar1 [ depvar2 ... = ] treatvar  indepvars  [if]  [in]  [weight] [,  options ]

    hte  graph  [,  level(#)  graph_options ]

options Description

Main

    alpha(#)  set significance level for balancing tests
    pscore_options  options as described in help pscore (except level())
    join(list)  merge specified strata
    autojoin(#)  merge small strata at low end or high end
    by(groupvar)  repeat analyses for groups defined by groupvar
    separate  construct propensity score strata separately for each by-group
    controls(list)  control variables for within-strata models
    estcom(command)  set estimation command for within strata models; default is regress
    estopts(options)  options to be applied to within-strata models
    nosily  display output from pscore and individual models
    level(#)  set confidence level; default is level(95)
    listwise  use listwise deletion to handle missing values
    casewise  synonym for listwise

Graph

    nograph  suppress graph
    outcomes(numlist)  display results for specified outcomes
    marker_options  change look of markers
    lineopts(options)  change look of fitted lines
    ciopts(options)  change look of confidence intervals
    noci  suppress confidence intervals
    addplot(plot)  add other plots to the generated graph
    twoway_options  any options other than by() documented in [G] twoway_options

fweights, iweights, and pweights are allowed; see help weight.

Description

hte applies pscore (Becker and Ichino 2002) to construct balanced propensity score strata
and, within each stratum, estimates the average treatment effect. hte then tests for
linear trend in treatment effects using variance-weighted least squares. The
stratum-specific treatment effects and the estimated linear trend are displayed in a
twoway graph.
See Brand and Xie (2010) for an application of this procedure.

`hte` specified without arguments redisplays the results from a previous `hte` call (without graph).

`hte graph` redraws the graph based on the results from a previous `hte` call.

**Dependencies**

`hte` requires `pscore` by Becker and Ichino (2002) to be installed on the system. Type

```
. net sj 5-3 st0026 2
```

followed by

```
. net install st0026 2
```

to install the program.

**Options**

### Main

**alpha(#)** sets the significance level for `pscore`'s tests of the balancing property. The default is `alpha(0.01)`.

`pscore options` are any options as described in help `pscore`, except `level()`.

**join(list)** causes the specified strata to be merged together for the treatment effect analysis. The syntax for `list` is

```
numlist [, numlist ...]
```

where `numlist` is a list of consecutive integers that identify the strata to be merged. For example, type `join(1 2)` to merge the first and second stratum. Multiple (disjunctive) `numlist`s may be specified, separated by a comma, in which case multiple merges are applied. After merging, the strata will be renumbered.

**autojoin[(#)]** causes small strata at the low and high end of the propensity score to be merged with subsequent or precedent strata, respectively, so that the number of observations is at least `#` for both the treated and the untreated (`#` defaults to 10). Only one of `join()` and `autojoin()` may be specified.

**by(groupvar)** specifies that the analysis be repeated for each group defined by the values of `groupvar`. The results are plotted in a single graph for all by-groups. Common propensity score strata are used for all groups unless the `separate` option is specified.

**separate** causes the construction of propensity score strata to be repeated for each by-group. The default is to use common strata, that is, to construct the strata once, based on the whole sample including all groups. `separate` has an effect only if `by()` is
specified.

**controls(clist)** specifies control variables to be included in the models used to estimate the within-strata treatment effects. *clist* may be a standard *varlist*, in which case the specified variables are included in each within-strata model. Alternatively, use the following syntax to specify strata-specific sets:

```
[varlist] [numlist1: varlist1] [numlist2: varlist2] [...]
```

*varlist* applies to all strata, *varlist1* applies to the strata specified in *numlist1*, etc.

**estcom(command)** sets the command used to estimate the within strata treatment effects. The default is *regress*.

**estopts(options)** are options to be applied to the models used to estimate the within-strata treatment effects. The options are as described in help *regress* (or as in help *command* where, *command* is the command specified via the *estcom()* option).

**noisily** displays the output from *pscore* and the treatment effect models. *pscore*'s *detail* option implies *noisily*.

**level(#)*** sets the confidence level, as a percentage, for confidence intervals. The default is *level(95)* or as set by *set level*.

**listwise** handles missing values through listwise deletion, meaning that an observation is excluded from all computations if any of the specified variables is missing for that observation. By default, *hte* uses all available observations to compute the propensity strata without regard to whether values for the outcome variables (the *depvars*) or the variables specified in *controls()* are missing.

**casewise** is a synonym for *listwise*.

**nograph** suppresses the graph.

**outcomes(numlist)** causes results to be plotted for the specified outcomes only. Use this option to select results in case of multiple outcome variables or *by()*-groups. Use numbers 1, 2, 3, etc. to refer to the different outcomes. By-groups are ordered within variables if both multiple outcome variables and the *by()* option are specified. That is, number 1 refers to *depvar1* and the first by-group, number 2 refers to *depvar1* and the second by-group, etc.

**marker_options** affect the rendition of the plotted markers, including their shape, size, color, and outline; see [G] marker_options. In case of multiple outcomes you can usually specify lists of elements to be applied to the different outcomes. For example, type *msymbol(D T)* to use diamonds for the first outcome and triangles for the second outcome.

**lineopts(cline_options)** affects the rendition of the plotted lines; see [G] cline_options. In case of multiple outcomes you can usually specify lists of elements to be applied to the different outcomes. For example, type *lineopts(lcolor(blue red))* to use blue
line color for the first outcome and red line color for the second outcome.

ciopts(options) affects the rendition of the capped confidence spikes for the within strata treatment effects; see \[G\] graph twoway rcap.

oci suppresses the confidence intervals for the within strata treatment effects. Confidence intervals are only displayed on plots containing a single outcome.

addplot(plot) provides a way to add other plots to the generated graph; see \[G\] addplot_option.

twoway_options are any options other than by() documented in \[G\] twoway_options.

Examples

Treatment effect of college on wages:

. sysuse nlsw88
. generate sq_exp = ttl_exp^2
. hte wage collgrad ttl_exp sq_exp tenure south smma

Add control variables to within strata treatment effect estimation:

. hte wage collgrad ttl_exp sq_exp tenure south smma, control(ttl_exp sq_exp)

Separate results by union:

. hte wage collgrad ttl_exp sq_exp tenure south smma, by(union)

Redraw graph for second group (union=1):

. hte_graph, outcome(2)

Saved results

hte saves the following in e():

Scalars

   e(N)        number of observations
   e(neq)      number of equations (outcomes)

Macros

   e(cmd)       hte
   e(estcom)    estimation command as specified by estcom()
   e(depvar)    name(s) of dependent variable(s)
   e(treatvar)  name of treatment variable
   e(indepvars) name(s) of independent variables
   e(controls)  expanded controls option
   e(trend)     note about linear fit
   e(byvar)     name of by() variable
   e(depvar#)   outcome #: name of dependent variable
   e(by#)       outcome #: by-group
\textbf{e(trend#)} outcome #: note about linear fit
\textbf{e(wtype)} weight type
\textbf{e(wexp)} weight expression
\textbf{e(properties)} b

\textbf{Matrices}
\textbf{e(b)} results vector
\textbf{e(se)} standard errors
\textbf{e(obs)} number of observations per stratum
\textbf{e(block)} strata ranks
\textbf{e(lfit)} linear fit of treatment effect by strata rank

\textbf{Functions}
\textbf{e(sample)} estimation sample

\textbf{References}


\textbf{Authors}
Ben Jann, ETH Zurich, jannb@ethz.ch

Jennie E. Brand, University of Calofornia - Los Angeles, brand@soc.ucla.edu

Yu Xie, University of Michigan, yuxie@isr.umich.edu

Thanks for citing this software as follows:


\textbf{Also see}

Online: help for \texttt{regress}, \texttt{vwls}; \texttt{pscore} (if installed)