teffects multivalued — Multivalued treatment effects

Description

This entry discusses the use of `teffects` when the treatment is multivalued. This entry presumes you are already familiar with the potential-outcome framework and the use of the `teffects` commands with binary treatments. See [TE] **teffects intro** or [TE] **teffects intro advanced** for more information.

Remarks and examples

Remarks are presented under the following headings:

Introduction

When the treatment is binary, each subject could either receive the treatment or not receive the treatment. In contrast, multivalued treatments refer to cases in which each subject could receive one of several different treatments or else not receive treatment at all. For example, in testing the efficacy of a drug, a patient could receive a 10 milligram (mg) dose, a 20 mg dose, a 30 mg dose, or no dose at all. We first want to be able to compare a patient receiving the 10 mg dose with a patient receiving no dose, a patient receiving the 20 mg dose with a patient receiving no dose, and a patient receiving the 30 mg dose with a patient receiving no dose. Once we can make those comparisons, we can then, for example, compare the efficacy of a 30 mg dose with that of a 20 mg dose or a 10 mg dose.

To highlight an example in economics, we consider an unemployed person who could participate in a comprehensive skills training program, attend a one-day workshop that helps job seekers write their resumés, or choose not to participate in either. We want to know how effective each of those programs is relative to not participating; once we know that, we can then compare the effectiveness of the comprehensive program with that of the one-day program.

Multivalued treatments increase the number of parameters that must be estimated and complicate the notation. Fortunately, however, using the `teffects` commands is not much more difficult with multivalued treatments than with binary treatments.

You can use `teffects ra`, `teffects ipw`, `teffects ipwra`, and `teffects aipw` to estimate multivalued treatment effects. However, the theory developed in Abadie and Imbens (2006, 2012) has not been extended to handle multivalued treatments, so you cannot use `teffects nnmatch` or `teffects psmatch` in these cases.

Cattaneo (2010), Imbens (2000), and Wooldridge (2010, sec. 21.6.3) discuss aspects of treatment-effect estimation with multivalued treatments.

Parameters and notation

We denote the potential outcome that subject i would obtain if given treatment-level t as y_{ti}, where y_{ti} is the realization of the random variable y_t. Throughout this entry, i subscripts denote realizations of the corresponding unsubscripted random variables. We again let y_0 denote the potential outcome of a subject who did not receive any treatment. To handle the case of multivalued treatments, we extend the definition of the unobservable, individual-level treatment effects to be $y_t - y_0$ for $t \in \{1, \ldots, q\}$.

As in the binary-valued case, we again focus on three parameters of interest: the average treatment effect (ATE), the potential-outcome mean (POM), and the average treatment effect on the treated (ATET).

ATE The ATE is the average effect of giving each individual treatment t instead of treatment 0:

$$\text{ATE}_t = E(y_t - y_0)$$

POM The POM for each treatment level is an average of each potential outcome:

$$\text{POM}_t = E(y_t)$$

ATET The ATET is the average effect among those subjects that receive treatment level $\breve{t}$ of giving each subject treatment $\widetilde{t}$ instead of treatment 0:

$$\text{ATET}_{\widetilde{t},\breve{t}} = E\left\{(y_{\widetilde{t}} - y_0)|t = \breve{t}\right\}$$

The extra notation required to define the ATET in this case indicates the difficulties surrounding this parameter.

Defining the ATET in the multivalued treatment case requires three different treatment levels: $\widetilde{t}$ defines the treatment level of the treated potential outcome; 0 is the treatment level of the control potential outcome; and $t = \breve{t}$ restricts the expectation to include only those individuals who actually receive treatment level $\breve{t}$.

Illustrating multivalued treatments

To illustrate the concept of a potential outcome and the parameters we would like to estimate, we consider the following table:

y	t	y_0	y_1	y_2
−0.50	0	−0.50	1.06	1.93
2.42	1	2.13	2.42	2.43
3.15	2	1.26	2.57	3.15
−0.39	0	−0.39	−0.18	0.52
2.22	2	−0.24	−0.01	2.22

We observe the outcome y as well as the treatment indicator t. There are three levels of treatment: 0, 1, or 2. Ideally, we would observe y_0, y_1, and y_2, but in fact all we have is y. In the first row, the subject received treatment level 0, so $y = y_0$ for that subject. In the last row, the subject received treatment 2, so $y = y_2$. We reiterate that we do not actually observe y_0, y_1, or y_2.

If we did have data on y_0, y_1, and y_2, then we could define subject-level treatment variables $te_1 = y_1 - y_0$ and $te_2 = y_2 - y_0$. Here we would be following the convention of taking treatment level 0 to be the control level. The following table adds these two variables:

y	t	y_0	y_1	y_2	te_1	te_2
−0.50	0	−0.50	1.06	1.93	1.56	2.43
2.42	1	2.13	2.42	2.43	0.29	0.30
3.15	2	1.26	2.57	3.15	1.31	1.89
−0.39	0	−0.39	−0.18	0.52	0.21	0.91
2.22	2	−0.24	−0.01	2.22	0.23	2.46

Once we have te_1 and te_2, obtaining the ATEs is straightforward. The ATE of going from treatment 0 to treatment 1 is simply the mean of the five entries in the column labeled te_1, which here works out to 0.72. Going from treatment level 0 to treatment level 1 causes the outcome to increase an average of 0.72. Similarly, the ATE of going from treatment 0 to treatment 2 is the mean of the entries in the column labeled te_2, which is 1.60. Exposing all subjects to treatment level 2 would cause the outcome to rise by an average of 1.60 relative to the outcome obtained by exposing them to treatment level 0.

The ATET is the average difference in the potential outcomes among those that get a particular treatment level. To compute this, we must specify two treatment levels: the actual treatment level the subjects we are interested in received as well as the treatment level we want to compare them with. For example, suppose we are interested in the ATET of going from treatment 0 to treatment 1 for those who received treatment 0. This ATET is the average of te_1 for those subjects for which $t = 0$. Here that ATET is just $(1.56 + 0.21)/2 \approx 0.89$. If we exposed the subjects who received treatment 0 to treatment 1 instead, the outcome would increase an average of 0.89.

The ATET of going from treatment 0 to treatment 2 for those subjects who received treatment 2 is the mean of te_2 for those subjects for which $t = 2$, which is $(1.89 + 2.46)/2 \approx 2.18$. Receiving treatment 2 increased the outcome of those who received treatment 2 by an average of 2.18 relative to receiving the control.

Examples

In the remainder of this entry, we provide several examples demonstrating how to estimate multivalued treatments using `teffects`.

▷ Example 1: Potential outcomes with four treatment levels

`bdsianesi5.dta` contains an extract of data from Blundell, Dearden, and Sianesi (2005). In this dataset on individuals in the United Kingdom, `wages` records hourly wages in pounds; `ed` records the highest educational degree obtained; `paed` records the highest educational level obtained by each individual's father; `math7` records a score obtained on a standardized math test when the individual was seven; `read7` records a score obtained on a standardized reading test when the individual was seven; and `london` and `eastern` are indicators for whether an individual lives in the expensive area of London or the east. We want to know how the level of education obtained affects a person's wage.

We begin by using `mean` to report the estimated means of wages over the four education levels. The value labels on `mean` are coded as `none` for no degree, `O` for an O-level degree, `A` for an A-level degree, or `H` for a higher-education degree.

```
. use http://www.stata-press.com/data/r15/bdsianesi5
(Excerpt from Blundell, Dearden, & Sianesi (2005) JRSSA 168: 473)
. mean wage, over(ed)
Mean estimation                    Number of obs     =      1,693

          none: ed = none
             O: ed = O
             A: ed = A
             H: ed = H

         Over         Mean   Std. Err.     [95% Conf. Interval]

wage
         none     6.057816    .154332      5.755114    6.360518
            O     7.501648   .1807359      7.147158    7.856137
            A     8.220637   .1540359      7.918516    8.522758
            H     10.87703   .2257888      10.43417    11.31988
```

The output reveals that the estimated mean wage increases as the education level goes from no degree to an O-level degree, to an A-level degree, and to a higher-education degree, as we would expect. Once we control for other characteristics of each individual, do we still observe a positive effect of education on wage?

We use `teffects ra` (see [TE] **teffects ra**) to estimate the ATEs of the different education levels by regression adjustment (RA), controlling for each person's location, math score, and father's education level:

```
. teffects ra (wage london eastern paed math7, poisson) (ed)
Iteration 0:   EE criterion =  1.865e-18
Iteration 1:   EE criterion =  4.077e-30
Treatment-effects estimation                   Number of obs     =      1,693
Estimator      : regression adjustment
Outcome model  : Poisson
Treatment model: none

                             Robust
         wage        Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

ATE
           ed
 (O vs none)     1.181543   .3520371     3.36   0.001     .4915626    1.871523
 (A vs none)     1.743079   .3282152     5.31   0.000     1.099789    2.386369
 (H vs none)     3.972829   .3840024    10.35   0.000     3.220199     4.72546

POmean
           ed
        none     6.525873   .2931933    22.26   0.000     5.951224    7.100521
```

Because wages are necessarily positive, we used the `poisson` option inside the outcome-model specification. The estimated POM of the control level of no degree is 6.53 pounds per hour. The estimated ATE of going from no degree to an O-level degree is 1.18 pounds per hour; the estimated ATE of going from no degree to an A-level degree is 1.74 pounds per hour; and the estimated ATE of going from no degree to a higher-education degree is 3.97 pounds per hour. All of these effects are highly significant.

For comparison purposes, we also use `teffects aipw` (see [TE] **teffects aipw**). We use the same outcome model as before. We use a multinomial logit model to predict education level, using math and reading scores and both the father's and the mother's educational attainment levels as predictors:

```
. teffects aipw (wage london eastern paed math7, poisson)
> (ed math7 read7 maed paed)
Iteration 0:   EE criterion =  1.877e-18
Iteration 1:   EE criterion =  1.029e-30
Treatment-effects estimation                    Number of obs     =      1,693
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: (multinomial) logit
```

wage	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
ed						
(0 vs none)	1.748197	.3911167	4.47	0.000	.9816221	2.514771
(A vs none)	2.363228	.3741584	6.32	0.000	1.629891	3.096565
(H vs none)	4.359777	.4133059	10.55	0.000	3.549712	5.169842
POmean						
ed						
none	5.946184	.3391531	17.53	0.000	5.281456	6.610912

The results indicate slightly higher treatment effects relative to those indicated by `teffects ra`. That is largely because the AIPW estimator predicts a lower no-higher-education POM than the RA estimator. ◁

▷ Example 2: Expressing ATEs as percentages

As in the binary-treatment case, expressing the ATEs as percentages of the POM for the control level often aids interpretation. Here we first use the replay facility of `teffects aipw` along with the `coeflegend` option to see how the parameters are named.

```
. teffects, coeflegend
Treatment-effects estimation                    Number of obs     =      1,693
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: (multinomial) logit
```

wage	Coef.	Legend
ATE		
ed		
(0 vs none)	1.748197	_b[ATE:r1vs0.ed]
(A vs none)	2.363228	_b[ATE:r2vs0.ed]
(H vs none)	4.359777	_b[ATE:r3vs0.ed]
POmean		
ed		
none	5.946184	_b[POmean:0.ed]

Now that we know the names, we can use `nlcom` to obtain the ATEs relative to the base-level POM:

```
. nlcom (_b[ATE:r1vs0.ed] / _b[POmean:0.ed])
>       (_b[ATE:r2vs0.ed] / _b[POmean:0.ed])
>       (_b[ATE:r3vs0.ed] / _b[POmean:0.ed])

       _nl_1:  _b[ATE:r1vs0.ed] / _b[POmean:0.ed]
       _nl_2:  _b[ATE:r2vs0.ed] / _b[POmean:0.ed]
       _nl_3:  _b[ATE:r3vs0.ed] / _b[POmean:0.ed]
```

wage	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.2940031	.0808156	3.64	0.000	.1356075	.4523988
_nl_2	.3974361	.0840545	4.73	0.000	.2326923	.5621799
_nl_3	.7332059	.1068848	6.86	0.000	.5237156	.9426962

Wages are 29% higher when everyone receives an O-level degree than if no one receives a degree. Wages are 40% higher when everyone receives an A-level degree than if no one receives a degree. Wages are 73% higher when everyone receives an H-level degree than if no one receives a degree.

Although impressive, these changes are not presented in the way that is most commonly discussed. (There is a large amount of literature on the treatment effect of getting a higher-education degree.) In particular, we might rather want to know the percentage changes in wages relative to a person with an A-level degree. Next we estimate the ATEs treating an A-level degree as the control level; to do that, we use the `control()` option. We also specify `coeflegend` again because we are more interested in how the parameters are named rather than in their standard errors at this point:

```
. teffects aipw (wage london eastern paed math7, poisson)
> (ed math7 read7 maed paed), control(A) coeflegend

Iteration 0:   EE criterion =  1.870e-18
Iteration 1:   EE criterion =  2.882e-30

Treatment-effects estimation                    Number of obs     =      1,693
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: (multinomial) logit
```

wage	Coef.	Legend
ATE		
ed		
(none vs A)	-2.363228	_b[ATE:r0vs2.ed]
(O vs A)	-.6150312	_b[ATE:r1vs2.ed]
(H vs A)	1.996549	_b[ATE:r3vs2.ed]
POmean		
ed		
A	8.309412	_b[POmean:2.ed]

Now we use `nlcom` to obtain the ATE of obtaining a higher-education degree as a percentage of the expected A-level wage:

```
. nlcom _b[ATE:r3vs2.ed] / _b[POmean:2.ed], noheader
```

wage	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.2402756	.0355404	6.76	0.000	.1706177	.3099335

The average wage increases by 24% when everyone receives an H-level degree relative to when everyone receives an A-level degree.

◁

▷ Example 3: Obtaining ATETs

In the previous example, we showed that on average, a higher-education degree increases a person's wage by 24% relative to someone with only an A-level degree. Sometimes, though, we would rather know how much the higher-education degree increases wages among the people who actually have a higher-education degree. To answer that question, we want to examine the ATET rather than the ATE.

Here we use the IPWRA estimator to obtain our answer. We specify the `control(A)` option so that an A-level education is treated as the basis for comparisons. We specify the `atet` option to obtain ATETs rather than ATEs, and we specify the `tlevel(H)` option to indicate that we want the ATETs to be calculated for the subset of people who actually receive higher-education degrees.

```
. teffects ipwra (wage london eastern paed math7, poisson)
> (ed math7 read7 maed paed), atet control(A) tlevel(H)

Iteration 0:   EE criterion =  2.731e-18
Iteration 1:   EE criterion =  5.636e-31

Treatment-effects estimation                   Number of obs     =      1,693
Estimator      : IPW regression adjustment
Outcome model  : Poisson
Treatment model: (multinomial) logit
```

wage	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
ed						
(none vs A)	-2.87423	.361093	-7.96	0.000	-3.58196	-2.166501
(O vs A)	-.8246604	.3609131	-2.28	0.022	-1.532037	-.1172837
(H vs A)	1.866757	.3277701	5.70	0.000	1.224339	2.509174
POmean						
ed						
A	9.010271	.2503971	35.98	0.000	8.519501	9.50104

The point estimates are similar to the ATEs we obtained above, suggesting that the means of the covariates among those with a higher-education degree are similar to the means for the entire population.

In output not shown to save space, we replayed the previous results with the `coeflegend` option to determine how the parameters are named. Armed with that information, we call `nlcom`:

```
. nlcom _b[ATET:r3vs2.ed] / _b[POmean:2.ed], noheader
```

wage	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.207181	.0407528	5.08	0.000	.127307	.287055

Our estimate of the percentage increase is now noticeably smaller once we restrict ourselves to only those people who actually received a higher-education degree. However, because of the width of the confidence intervals, there is no evidence to suggest that the difference between the estimates is statistically significant.

◁

▷ Example 4: ATEs comparing adjacent treatments

In the first example, we obtained the three ATEs, and they were all expressed relative to the base level of no degree. Now we show how we can express the gains to an O-level degree relative to no degree, the gains to an A-level degree relative to an O-level degree, and the gains to a higher-education degree relative to an A-level degree.

First, we use an AIPW estimator to obtain all the POMs for our example dataset:

```
. teffects aipw (wage london eastern paed math7, poisson)
> (ed math7 read7 maed paed), pom
Iteration 0:   EE criterion =  1.877e-18
Iteration 1:   EE criterion =  1.542e-30

Treatment-effects estimation                    Number of obs     =      1,693
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: (multinomial) logit
```

wage	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
POmeans						
ed						
none	5.946184	.3391531	17.53	0.000	5.281456	6.610912
O	7.694381	.1915192	40.18	0.000	7.31901	8.069752
A	8.309412	.1563348	53.15	0.000	8.003001	8.615823
H	10.30596	.2285837	45.09	0.000	9.857945	10.75398

ATEs are contrasts of POMs, and here we show how to use `contrast` to obtain the estimated ATEs:

```
. contrast r.ed, nowald
Warning: cannot perform check for estimable functions.

Contrasts of marginal linear predictions

Margins      : asbalanced
```

	Contrast	Std. Err.	[95% Conf.	Interval]
POmeans				
ed				
(O vs none)	1.748197	.3911167	.9816221	2.514771
(A vs none)	2.363228	.3741584	1.629891	3.096565
(H vs none)	4.359777	.4133059	3.549712	5.169842

These estimated ATEs match those we obtained in example 2.

Now that we know how to use `contrast` to obtain the ATEs based on the POMs, we can take advantage of `contrast`'s ability to obtain "reverse adjacent" contrasts, which compare each level with the previous level. We use the `ar.` operator with `contrast` to accomplish this:

```
. contrast ar.ed, nowald
Warning: cannot perform check for estimable functions.
Contrasts of marginal linear predictions
Margins      : asbalanced
```

	Contrast	Std. Err.	[95% Conf.	Interval]
POmeans				
ed				
(O vs none)	1.748197	.3911167	.9816221	2.514771
(A vs O)	.6150312	.2432806	.13821	1.091852
(H vs A)	1.996549	.2730712	1.461339	2.531759

These ATEs are for incremental increases. In contrast, the ATEs considered above had a common base.

◁

❑ Technical note

The multivalued treatment AIPW estimators implemented in **teffects aipw** are EIF estimators based on the results of Cattaneo (2010). The results in Cattaneo (2010) are for semiparametric estimators, and we implement parametric versions. Of more practical importance, Cattaneo (2010) contains results for quantile treatment effects that are not implemented in **teffects** but implemented in the user-written **poparms** command discussed in Cattaneo, Drukker, and Holland (2013). See Emsley et al. (2008) for another implementation of the AIPW estimator, and see Frölich and Melly (2010) for other commands that estimate quantile treatment effects.

❑

References

Abadie, A., and G. W. Imbens. 2006. Large sample properties of matching estimators for average treatment effects. *Econometrica* 74: 235–267.

——. 2012. Matching on the estimated propensity score. Harvard University and National Bureau of Economic Research. http://www.nber.org/papers/w15301.

Blundell, R. W., L. Dearden, and B. Sianesi. 2005. Evaluating the effect of education on earnings: Models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society, Series A* 168: 473–512.

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

Cattaneo, M. D., D. M. Drukker, and A. D. Holland. 2013. Estimation of multivalued treatment effects under conditional independence. *Stata Journal* 13: 407–450.

Emsley, R., M. Lunt, A. Pickles, and G. Dunn. 2008. Implementing double-robust estimators of causal effects. *Stata Journal* 8: 334–353.

Frölich, M., and B. Melly. 2010. Estimation of quantile treatment effects with Stata. *Stata Journal* 10: 423–457.

Imbens, G. W. 2000. The role of the propensity score in estimating dose–response functions. *Biometrika* 87: 706–710.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Also see

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

[TE] **teffects intro** — Introduction to treatment effects for observational data

[TE] **teffects intro advanced** — Advanced introduction to treatment effects for observational data