Estimating average treatment effects from observational data using teffects

David M. Drukker

Director of Econometrics
Stata

2013 Italian Stata Users Group meeting
October 14, 2013

Outline

A question

- Will a mother hurt her child by smoking while she is pregnant?
  - Too vague
- Will a mother reduce the birthweight of her child by smoking while she is pregnant?
  - Less interesting, but more specific
  - There might even be data to help us answer this question
  - The data will be observational, not experimental

Potential outcomes

- Potential outcomes are the data that we wish we had to estimate causal treatment effects
- Suppose that we could see
  1. the birthweight of a child born to each mother when she smoked while pregnant, and
  2. the birthweight of a child born to each mother when she did not smoke while pregnant

  For example, we wish we had data like

```
. list mother_id bw_smoke bw_nosmoke in 1/5, abbreviate(10)

     +-------------------------------------+
     | mother_id   bw_smoke     bw_nosmoke |
     |-------------------------------------|
  1. |         1       3183           3509 |
  2. |         2       3060           3316 |
  3. |         3       3165           3474 |
  4. |         4       3176           3495 |
  5. |         5       3241           3413 |
     +-------------------------------------+
```

- There are two treatment levels, the mother smokes and the mother does not smoke
  - For each treatment level, there is an outcome (a baby's birthweight) that would be observed if the mother got that treatment level

Average treatment effect

- If we had data on each potential outcome, the sample-average treatment effect would be the sample average of bw_smoke minus bw_nosmoke

```
. mean bw_smoke bw_nosmoke
Mean estimation                    Number of obs     =     4642
```

	Mean	Std. Err.	[95% Conf.	Interval]
bw_smoke	3171.72	.9088219	3169.938	3173.501
bw_nosmoke	3402.599	1.529189	3399.601	3405.597

```
. lincom _b[bw_smoke] - _b[bw_nosmoke]
 ( 1)  bw_smoke - bw_nosmoke = 0
```

Mean	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	-230.8791	1.222589	-188.84	0.000	-233.276	-228.4823

- In population terms, the average treatment effect is

$$ATE = \mathbf{E}[bw_{smoke} - bw_{nosmoke}] = \mathbf{E}[bw_{smoke}] - \mathbf{E}[bw_{nosmoke}]$$

Missing data

- The "fundamental problem of causal inference" (Holland (1986)) is that we only observe one of the potential outcomes
  - The other potential outcome is missing

1. We only see bw_{smoke} for mothers who smoked
2. We only see $bw_{nosmoke}$ for mothers who did not smoked

- We can use the tricks of missing-data analyis to estimate treatment effects
- For more about potential outcomes Rubin (1974), Holland (1986), Heckman (1997), Imbens (2004), (Cameron and Trivedi, 2005, chapter 2.7), Imbens and Wooldridge (2009), and (Wooldridge, 2010, chapter 21)

Random-assignment case

- Many questions require using observational data, because experimental data would be unethical
    - We could not ask a random selection of mothers to smoke while pregnant
- The random-assignment methods used with experimental data are useful, because observational-data methods build on them
- When the treatment is randomly assigned, the potential outcomes are independent of the treatment
- If smoking were randomly assigned to mothers, the missing potential outcome would be missing completely at random
    1. The average birthweight of babies born to mothers who smoked would be a good estimator for mean of the smoking potential outcome of all mothers in the population
    2. The average birthweight of babies born to mothers who did not smoke would be a good estimator for mean of the not-smoking potential outcome of all mothers in the population
    3. The difference in the two averages computed from

Difference in means

```
. regress bweight ibn.mbsmoke, noconstant
      Source |       SS       df       MS              Number of obs =    4642
-------------+------------------------------           F(  2,  4640) =81131.59
       Model |  5.2512e+10     2  2.6256e+10           Prob > F      =  0.0000
    Residual |  1.5016e+09  4640  323622.478           R-squared     =  0.9722
-------------+------------------------------           Adj R-squared =  0.9722
       Total |  5.4014e+10  4642  11635851.6           Root MSE      =  568.88
------------------------------------------------------------------------------
     bweight |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     mbsmoke |
  nonsmoker  |   3412.912   9.255254   368.75   0.000     3394.767    3431.056
     smoker  |    3137.66   19.35363   162.12   0.000     3099.717    3175.602
------------------------------------------------------------------------------
. contrast r.mbsmoke, nowald
Contrasts of marginal linear predictions
Margins      : asbalanced
------------------------------------------------------------
                        |   Contrast   Std. Err.     [95% Conf. Interval]
------------------------+-----------------------------------
                mbsmoke |
(smoker vs nonsmoker)   |  -275.2519    21.4528    -317.3096   -233.1942
------------------------------------------------------------
```

As good as random

- Instead of assuming that the treatment is randomly assigned, we will now assume that the after conditioning on covariates the treatment is as good as randomly assigned
- Formally, this assumption is known as conditional independence
- Even more formally, we only need conditional mean independence which says that after conditioning on covariates, the treatment does not affect the means of the potential outcomes

Assumptions used with observational data

- The assumptions we need vary over estimator and effect parameter, but some version of the following assumptions are required.

CMI The conditional mean-independence CMI assumption restricts the dependence between the treatment model and the potential outcomes

Overlap The overlap assumption ensures that each individual could get any treatment level

IID The independent-and-identically-distributed (IID) sampling assumption ensures that the potential outcomes and treatment status of each individual are unrelated to the potential outcomes and treatment statuses of all the other individuals in the population

The overlap assumption

- The overlap assumption requires that each individual has a positive probability of receiving each treatment level.
- Formally, the overlap assumption requires that for each possible $\mathbf{x}_i$ in the population and each treatment level t, $0 < \mathbf{P}(t_i = t|\mathbf{x}) < 1$.

The IID assumption

- We also make the standard assumption that we have an independently and identically distributed (IID) sample from the population
- In potential-outcome models, IID sampling implies that the potential outcomes and treatment status of each individual are unrelated to the potential outcomes and treatment statuses of all the other individuals in the population
  - IID sampling rules out interactions among the individuals
  - For instance, models of vaccinations in epidemiology and spatially-dependent outcomes in economics violate the independence assumption

Some references for assumptions

- Versions of the CMI assumption are also known as unconfoundedness and selection-on-observables in the literature; see Rosenbaum and Rubin (1983), Heckman (1997), Heckman and Navarro-Lozano (2004), (Cameron and Trivedi, 2005, section 25.2.1), (Tsiatis, 2006, section 13.3), (Angrist and Pischke, 2009, chapter 3), Imbens and Wooldridge (2009), and (Wooldridge, 2010, section 21.3)
- Rosenbaum and Rubin (1983) call the combination of conditional independence and overlap assumptions strong ignorability; see also (Abadie and Imbens, 2006, pp 237-238) and Imbens and Wooldridge (2009).
- The IID assumption is a part of what is known as the stable unit treatment value assumption (SUTVA); see (Wooldridge, 2010, p.905) and Imbens and Wooldridge (2009)

Choice of auxiliary model

- Recall that the potential-outcomes framework formulates the estimation of the ATE as a missing-data problem
- We use the parameters of an auxiliary model to solve the missing-data problem

Model		Estimator
outcome	$\rightarrow$	Regression adjustment (RA)
treatment	$\rightarrow$	Inverse-probability weighted (IPW)
outcome and treatment	$\rightarrow$	Augmented IPW (AIPW)
outcome and treatment	$\rightarrow$	IPW RA (IPWRA)
outcome (nonparametrically)	$\rightarrow$	Nearest-neighbor matching (NNMATCH)
treatment	$\rightarrow$	Propensity-score matching (PSMATCH)

Regression adjustment estimators

- Regression adjustment (RA) estimators:
  - RA estimators run separate regressions for each treatment level, then
    - use means of predicted outcomes for each treatment level to estimate each POM
    - use differences of POMs, or conditional on the treated POMs, to estimate ATEs or ATETs
  - Formally, the CMI assumption implies that we can we can estimate $\mathbf{E}(y_t|\mathbf{x}_i)$ directly from the observations for which person i gets treatment t
    - y_t is the potential outcome for treatment level t
    - Averages of predicted $\mathbf{E}(y_t|\mathbf{x}_i)$ yield estimates of the POM $\mathbf{E}[y_t]$
- See (Cameron and Trivedi, 2005, chapter 25), (Wooldridge, 2010, chapter 21), and (Vittinghoff et al., 2012, chapter 9)

RA example I

```
. use cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)
. teffects ra (bweight mmarried prenatal1 fbaby medu) (mbsmoke)
Iteration 0:   EE criterion =  2.336e-23
Iteration 1:   EE criterion =  5.702e-26
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : linear
Treatment model: none
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE mbsmoke (smoker vs nonsmoker)	-230.9541	24.34012	-9.49	0.000	-278.6599	-183.2484
POmean mbsmoke nonsmoker	3402.548	9.546721	356.41	0.000	3383.836	3421.259

- RA with linear regression to model outcome

RA example II

```
. teffects ra (bweight mmarried prenatal1 fbaby medu, poisson) (mbsmoke)
Iteration 0:   EE criterion =  3.924e-17
Iteration 1:   EE criterion =  2.605e-24
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : Poisson
Treatment model: none
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -230.7723   24.41324    -9.45   0.000    -278.6213   -182.9232
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
  nonsmoker  |   3402.497   9.547989   356.36   0.000     3383.783    3421.211
------------------------------------------------------------------------------
```

- RA with exponential conditional mean to model outcome

RA other models

- `teffects ra` can also model the outcome using probit, logit, or heteroskedastic probit

Inverse-probability-weighted estimators

- Inverse-probability-weighted (IPW) estimators:
  - IPW estimators weight observations on the outcome variable by the inverse of the probability that it is observed to account for the missingness process
  - Observations that are not likely to contain missing data get a weight close to one; observations that are likely to contain missing data get a weight larger than one, potentially much larger
  - IPW estimators model the probability of treatment without any assumptions about the functional form for the outcome model
  - In contrast, RA estimators model the outcome without any assumptions about the functional form for the probability of treatment model
- See Horvitz and Thompson (1952) Robins and Rotnitzky (1995), Robins et al. (1994), Robins et al. (1995), Imbens (2000), Wooldridge (2002), Hirano et al. (2003), (Tsiatis, 2006, chapter 6), Wooldridge (2007) and (Wooldridge, 2010, chapters 19 and 21)

IPW example I

```
. teffects ipw (bweight ) (mbsmoke mmarried prenatal1 fbaby medu)
Iteration 0:   EE criterion =  1.704e-23
Iteration 1:   EE criterion =  4.483e-27
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-231.1516	24.03183	-9.62	0.000	-278.2531	-184.0501
POmean						
mbsmoke nonsmoker	3402.219	9.589812	354.77	0.000	3383.423	3421.015

- IPW with logit to model treatment

IPW example II

```
. teffects ipw (bweight) (mbsmoke mmarried prenatal1 fbaby medu, hetprobit(medu
> ))
Iteration 0:   EE criterion =  7.158e-16
Iteration 1:   EE criterion =  7.930e-26
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: heteroskedastic probit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -217.7521    28.5796    -7.62   0.000    -273.7671   -161.7371
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3401.788   9.570692   355.44   0.000      3383.03    3420.546
------------------------------------------------------------------------------
```

- IPW with heteroskedastic probit to model treatment
- Could have used probit to model the treatment

Augmented IPW estimators

- Augmented IPW (AIPW) estimators
  - Augmented-inverse-probability-weighted (AIPW) estimators model both the outcome and the treatment probability
  - The estimating equation that combines both models is essentially an IPW estimating equation with an augmentation term
  - AIPW estimator have the double-robust property
    - only one of the two models must be correctly specified to consistently estimate the treatment effects
  - AIPW estimators can be more efficient than IPW or RA estimators
- See Robins and Rotnitzky (1995), Robins et al. (1995), Lunceford and Davidian (2004), Bang and Robins (2005), (Tsiatis, 2006, chapter 13), Cattaneo (2010), Cattaneo et al. (2013)

AIPW example I

```
. teffects aipw (bweight mmarried prenatal1 fbaby medu)   ///
>         (mbsmoke mmarried prenatal1 fbaby medu)
Iteration 0:   EE criterion =  4.031e-23
Iteration 1:   EE criterion =  2.196e-26
Treatment-effects estimation                   Number of obs      =       4642
Estimator      : augmented IPW
Outcome model  : linear by ML
Treatment model: logit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -229.7809   24.96839    -9.20   0.000     -278.718   -180.8437
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
  nonsmoker  |   3403.122   9.564165   355.82   0.000     3384.376    3421.867
------------------------------------------------------------------------------
```

- AIPW with linear model for outcome and logit for treatment

AIPW example II

```
. teffects aipw (bweight mmarried prenatal1 fbaby medu, poisson)  ///
>         (mbsmoke mmarried prenatal1 fbaby medu, hetprobit(medu))
Iteration 0:   EE criterion =  7.551e-16
Iteration 1:   EE criterion =  1.312e-24
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: heteroskedastic probit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |   -220.496   28.30292    -7.79   0.000    -275.9687   -165.0233
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3402.429   9.557345   356.00   0.000     3383.697    3421.161
------------------------------------------------------------------------------
```

- AIPW with exponential conditional mean model for outcome and heteroskedastic probit for treatment
- Could have used linear, poisson, logit, probit, or heteroskedastic probit to model the outcome and probit, logit, or heteroskedastic logit to model the treatment

- IPWRA estimators combine models for the outcome and the treatment
- IPWRA estimators are double-robust
- IPWRA use the inverse of the estimated treatment-probability weights to estimate missing-data-corrected regression coefficients that are subsequently used to compute the POMs
    - The ATE is estimated by a difference in the estimated POMs
- See Wooldridge (2007) and (Wooldridge, 2010, section 21.3.4)

IPWRA example I

```
. teffects ipwra (bweight mmarried prenatal1 fbaby medu)  ///
>          (mbsmoke mmarried prenatal1 fbaby medu)
Iteration 0:   EE criterion =  9.630e-22
Iteration 1:   EE criterion =  1.298e-25
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -227.4408   25.62591    -8.88   0.000    -277.6667    -177.215
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3403.027    9.56025   355.96   0.000     3384.289    3421.765
------------------------------------------------------------------------------
```

- IPWRA with linear model for outcome and logit for treatment

IPWRA example II

```
. teffects ipwra (bweight mmarried prenatal1 fbaby medu, poisson) ///
>         (mbsmoke mmarried prenatal1 fbaby medu, hetprobit(medu))
Iteration 0:   EE criterion =  7.496e-16
Iteration 1:   EE criterion =  4.003e-24
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : Poisson
Treatment model: heteroskedastic probit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -221.2331   27.66194    -8.00   0.000    -275.4495   -167.0166
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3402.416   9.558767   355.95   0.000     3383.682    3421.151
------------------------------------------------------------------------------
```

- IPWRA with exponential conditional mean model for outcome and heteroskedastic probit for treatment
- Could have used linear, poisson, logit, probit, or heteroskedastic probit to model the outcome and probit, logit, or heteroskedastic logit to model the treatment

Matching estimators

- Matching estimators use an average of the outcomes of the nearest individuals to impute the missing potential outcome for each sampled individual
- The difference between the observed outcome and the imputed potential outcome is essentially an estimate of the expected individual-level treatment effect conditional on the covariates
- These estimated expected individual-level treatment effects are averaged to estimate the ATE

Nearest-neighbor matching

- Nearest-neighbor matching (NNM) determines "nearest" using a weighted function of the covariates for each observation
- NNM is nonparametric
    - No explicit functional form for either the outcome model or the treatment model is specified
    - The estimator needs more data to get to the true value than an estimator that imposes a functional form
        - The NNM estimator converges to the true value at a rate slower than the parametric rate, when matching on more than one continuous covariate
    - `teffects nnmatch` uses bias-correction to fix this problem

Nearest-neighbor matching II

- See Abadie and Imbens (2006) and Abadie and Imbens (2011) for formal results, rates of convergence, and the details of the bias-correction methods
- Rubin (1973), Rubin (1977), Quade (1982) did early work on matching estimators with formal results in Abadie and Imbens (2006) and Abadie and Imbens (2011)
- `teffect nnmatch` is based on the results in Abadie and Imbens (2006) and Abadie and Imbens (2011) and a previous implementation in Abadie et al. (2004)

NNM example

```
. teffects nnmatch (bweight mmarried prenatal1 fbaby medu) (mbsmoke)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : nearest-neighbor matching     Matches: requested =         1
Outcome model  : matching                                     min =         1
Distance metric: Mahalanobis                                  max =       645
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -220.5255    28.0835    -7.85   0.000    -275.5681   -165.4828
------------------------------------------------------------------------------
```

Propensity-score matching

- Propensity-score matching (PSM) determines "nearest" using the estimated treatment probabilities, which are known as the propensity scores
  - PSM is implemented in `teffects psmatch`
- PSM provides an alternative to bias-correction because it matches on a single continuous covariate, the estimated treatment probabilities
- Abadie and Imbens (2012) derived the standard errors that account for the error in estimating the propensity scores

PSM example I

```
. teffects psmatch (bweight) (mbsmoke mmarried prenatal1 fbaby medu)
Treatment-effects estimation                   Number of obs      =       4642
Estimator      : propensity-score matching     Matches: requested =          1
Outcome model  : matching                                     min =          1
Treatment model: logit                                        max =        645
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
          vs |
  nonsmoker) |  -217.3852   28.98542    -7.50   0.000    -274.1956   -160.5748
------------------------------------------------------------------------------
```

- Used logit for propensity score
- Other choices were probit or heteroskedastic probit

PSMATCH example I

```
. teffects psmatch (bweight) (mbsmoke mmarried prenatal1 fbaby medu)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : propensity-score matching     Matches: requested =         1
Outcome model  : matching                                     min =         1
Treatment model: logit                                        max =       645
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -217.3852   28.98542    -7.50   0.000    -274.1956   -160.5748
------------------------------------------------------------------------------
```

- Used heteroskedastic probit for propensity score
- Other choices were logit or probit

Now what?

- Go to `http://www.stata.com/manuals13/te.pdf` entry `teffects intro advanced` for more information and lots of links to literature and examples

What are QTE

- Quantile treatment effects (QTE) are differences in the quantiles of the marginal potential outcome distributions
  - $q_1(\tau) = F_{y_1}^{-1}(\tau)$ is the τ(th) quantile of the distribution of the treated potential outcome y_1
  - $q_0(\tau) = F_{y_0}^{-1}(\tau)$ is the τ(th) quantile of the distribution of the control potential outcome y_0
  - $q_1(\tau)$ and $q_0(\tau)$ are quantiles of the marginal distributions of the potential outcomes
  - $QTE = q_1(\tau) - q_0(\tau)$, the QTE is the difference in the marginal quantiles
    - The distributions are marginalized over the distributions of the covariates
    - $F_{y_j}(y) = \mathbf{E}_{\mathbf{x}}[F_{y_j|\mathbf{x}}(y|\mathbf{x})]$
    - Keep in mind that $q_j(\tau) = F_{y_j}^{-1}(\tau) \neq \mathbf{E}[q_j(\tau|\mathbf{x})]$, where $q_j(\tau|\mathbf{x})$ is condition-on-$\mathbf{x}$ quantile of the potential-outcome distribution

QTE can differ over τ

- Suppose that robust babies, those born at the .80 quantile, would not be measurably harmed by the mother smoking a few cigarettes
- Further suppose that at-risk babies, those born at the .20 quantile, could be seriously harmed by the mother smoking a few cigarettes
- ATE and ATET cannot investigate this type of hypothesis
- QTE can investigate this type of hypothesis

poparms estimates QTEs

- `poparms` is a user written command documented in Cattaneo et al. (2013)
- `poparms` estimates mean and quantiles of the potential-outcome distributions
  - `poparms` implements an IPW and an AIPW derived in Cattaneo (2010)
  - Cattaneo (2010) and Cattaneo et al. (2013) call the AIPW estimator an efficient-influence function (EIF) estimator because EIF theory is what produces the augmentation term

poparms

- `poparms` installation

```
. findit poparms
. net install st0303, replace
checking st0303 consistency and verifying not already installed...
all files already exist and are up to date.
. help poparms
```

poparms example

- poparms estimates

```
. clear all
. use cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)
. poparms (mbsmoke mmarried fbaby medu mage c.medu#c.medu c.mage#c.mage)  ///
>         (bweight prenatal1 fbaby medu mage),                             ///
>         quantiles(.2 .8)
Treatment Mean and Quantiles Estimation          Number of obs    =       4642
(efficient influence function)
------------------------------------------------------------------------------
             |               bootstrap
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
mean         |
     mbsmoke |
   nonsmoker |    3403.35   9.696517   350.99   0.000     3384.346    3422.355
      smoker |   3183.081   27.67854   115.00   0.000     3128.832     3237.33
-------------+----------------------------------------------------------------
q20          |
     mbsmoke |
   nonsmoker |       3000   13.34484   224.81   0.000     2973.845    3026.155
      smoker |       2778   31.33055    88.67   0.000     2716.593    2839.407
-------------+----------------------------------------------------------------
q80          |
     mbsmoke |
   nonsmoker |       3840    9.76136   393.39   0.000     3820.868    3859.132
      smoker |       3625   28.05127   129.23   0.000     3570.021    3679.979
------------------------------------------------------------------------------
```

poparms example

- poparms estimates

```
. poparms, coeflegend
Treatment Mean and Quantiles Estimation          Number of obs    =       4642
(efficient influence function)
------------------------------------------------------------------------------
     bweight |      Coef.  Legend
-------------+----------------------------------------------------------------
mean         |
     mbsmoke |
  nonsmoker  |    3403.35  _b[mean:0bn.mbsmoke]
     smoker  |   3183.081  _b[mean:1.mbsmoke]
-------------+----------------------------------------------------------------
q20          |
     mbsmoke |
  nonsmoker  |       3000  _b[q20:0bn.mbsmoke]
     smoker  |       2778  _b[q20:1.mbsmoke]
-------------+----------------------------------------------------------------
q80          |
     mbsmoke |
  nonsmoker  |       3840  _b[q80:0bn.mbsmoke]
     smoker  |       3625  _b[q80:1.mbsmoke]
------------------------------------------------------------------------------
```

poparms example

- poparms estimates

```
. lincom _b[mean:1.mbsmoke] - _b[mean:0.mbsmoke]
 ( 1)  - [mean]0bn.mbsmoke + [mean]1.mbsmoke = 0
```

bweight	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	-220.2692	29.24745	-7.53	0.000	-277.5931	-162.9452

```
. lincom _b[q20:1.mbsmoke]  - _b[q20:0.mbsmoke]
 ( 1)  - [q20]0bn.mbsmoke + [q20]1.mbsmoke = 0
```

bweight	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	-222	34.18932	-6.49	0.000	-289.0098	-154.9902

```
. lincom _b[q80:1.mbsmoke]  - _b[q80:0.mbsmoke]
 ( 1)  - [q80]0bn.mbsmoke + [q80]1.mbsmoke = 0
```

bweight	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
(1)	-215	29.51215	-7.29	0.000	-272.8427	-157.1573

Abadie, Alberto, David M. Drukker, Jane Leber Herr, and Guido W. Imbens. 2004. "Implementing matching estimators for average treatment effects in Stata," *Stata journal*, 4(3), 290–311.

Abadie, Alberto and Guido Imbens. 2012. "Matching on the estimated propensity score," .

Abadie, Alberto and Guido W. Imbens. 2006. "Large sample properties of matching estimators for average treatment effects," *Econometrica*, 235–267.

———. 2011. "Bias-corrected matching estimators for average treatment effects," *Journal of Business and Economic Statistics*, 29(1), 1–11.

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

Bang, Heejung and James M. Robins. 2005. "Doubly robust estimation in missing data and causal inference models," *Biometrics*, 61(4), 962–973.

Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and applications*, Cambridge: Cambridge University Press.

Cattaneo, Matias D., David M. Drukker, and Ashley D. Holland. 2013. "Estimation of multivalued treatment effects under conditional independence," *Stata Journal*, 13(3), ??

Cattaneo, M.D. 2010. "Efficient semiparametric estimation of multi-valued treatment effects under ignorability," *Journal of Econometrics*, 155(2), 138–154.

Heckman, James and Salvador Navarro-Lozano. 2004. "Using matching, instrumental variables, and control functions to estimate economic choice models," *Review of Economics and statistics*, 86(1), 30–57.

Heckman, James J. 1997. "Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations," *Journal of Human Resources*, 32(3), 441–462.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient estimation of average treatment effects using the estimated propensity score," *Econometrica*, 71(4), 1161–1189.

Holland, Paul W. 1986. "Statistics and causal inference," *Journal of the American Statistical Association*, 945–960.

Horvitz, D. G. and D. J. Thompson. 1952. "A Generalization of Sampling Without Replacement From a Finite Universe," *Journal of the American Statistical Association*, 47(260), 663–685.

Imbens, Guido W. 2000. "The role of the propensity score in estimating dose-response functions," *Biometrika*, 87(3), 706–710.

———. 2004. "Nonparametric estimation of average treatment effects under exogeneity: A review," *Review of Economics and statistics*, 86(1), 4–29.

Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47, 5–86.

Lunceford, Jared K and Marie Davidian. 2004. "Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study," *Statistics in medicine*, 23(19), 2937–2960.

Quade, Dana. 1982. "Nonparametric Analysis of Covariance by Matching," *Biometrics*, 38(3), 597–611.

Robins, James M. and Andrea Rotnitzky. 1995. "Semiparametric Efficiency in Multivariate Regression Models with Missing Data," *Journal of the American Statistical Association*, 90(429), 122–129.

Robins, James M., Andrea Rotnitzky, and Lue Ping Zhao. 1994. "Estimation of Regression Coefficients When Some Regressors Are Not Always Observed," *Journal of the American Statistical Association*, 89(427), 846–866.

———. 1995. "Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data," *Journal of the American Statistical Association*, 90(429), 106–121.

Rosenbaum, P. and D. Rubin. 1983. "Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55.

Rubin, Donald. 1973. "Matching to Remove Bias in Observational Studies," *Biometrics*, 29, 159–183.

Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology*, 66(5), 688.

———. 1977. "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational and Behavioral statistics*, 2(1), 1–26.

Tsiatis, Anastasios A. 2006. *Semiparametric theory and missing data*, New York: Springer Verlag.

Vittinghoff, E., D. V. Glidden, S. C. Shiboski, and C. E. McCulloch. 2012. *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*, New York: Springer, 2 ed.

Wooldridge, Jeffrey M. 2002. "Inverse probability weighted M-estimators for sample selection, attrition, and stratification," *Portuguese Economic Journal*, 1, 117–139.

———. 2007. "Inverse probability weighted estimation for general missing data problems," *Journal of Econometrics*, 141(2), 1281–1301.

———. 2010. *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Massachusetts: MIT Press, second ed.