Treatment Effects Using Stata

Enrique Pinzón

StataCorp LP

October, 2013
Madrid

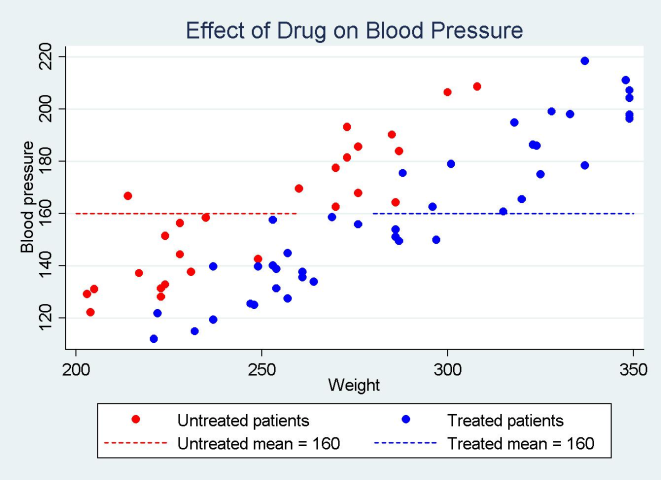
Motivation

We are interested in the outcomes of receiving a treatment in scenarios were researchers have observational data.

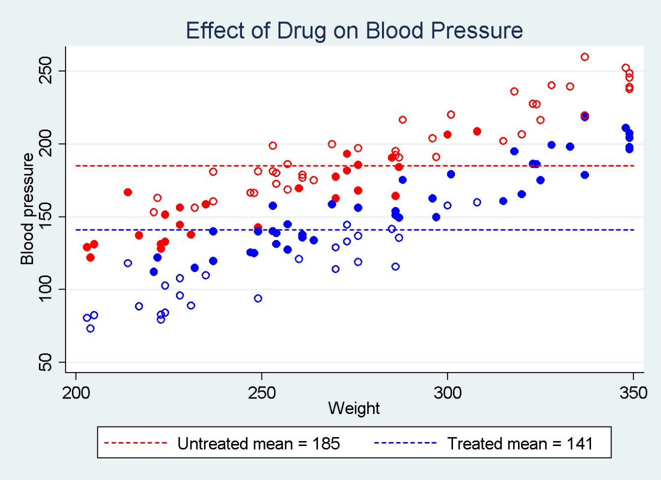
For instance:

- The impact on public education outcomes for schools that received a transfer and those that did not.
- Employment outcomes for individuals that participated in a job training program and those that did not.
- The effect on birth weight for babies of mothers that smoked relative to those of mothers that did not.

Observed Effect of Statin on Blood Presure



Potential Outcomes of Statin on Blood Presure



How We Approach Treatment Effects

- We cannot observe individuals in both states simultaneously
  - Design a random experiment
  - We cannot do this because of technical or ethical concerns
- We need to account for covariates that are correlated with the treatment
- We will think of the problem in terms of models that govern the treatment result and the outcome

How We Approach Treatment Effects

- We cannot observe individuals in both states simultaneously
  - Design a random experiment
  - We cannot do this because of technical or ethical concerns
- We need to account for covariates that are correlated with the treatment
- We will think of the problem in terms of models that govern the treatment result and the outcome

How We Approach Treatment Effects

- We cannot observe individuals in both states simultaneously
  - Design a random experiment
  - We cannot do this because of technical or ethical concerns
- We need to account for covariates that are correlated with the treatment
- We will think of the problem in terms of models that govern the treatment result and the outcome

How We Approach Treatment Effects

- We cannot observe individuals in both states simultaneously
  - Design a random experiment
  - We cannot do this because of technical or ethical concerns
- We need to account for covariates that are correlated with the treatment
- We will think of the problem in terms of models that govern the treatment result and the outcome

Notation and Definitions

- The potential outcome is denoted by the random variable y_τ with $\tau \in \{0, 1, \ldots, K\}$. The potential realizations will be denoted by:
  - y_{0i} is the outcome individual i if they do not receive the treatment, where $i = 1 \ldots n$
  - y_{ki} is the potential outcome for individual i if they receive different discrete levels of the treatment, where $k = 1 \ldots K$
  - Usually people think about the binary case where there are only two levels y_{0i} and y_{1i}
- Potential outcome mean

$$POM = E\left(y_\tau\right)$$

- Average treatment effect

$$ATE = E\left(y_{ki} - y_{0i}\right)$$

- Average treatment effect on the treated

$$ATET = E\left(y_{ki} - y_{0i} | \tau = k\right)$$

- From now on we will focus on binary treatments. All results are valid for multivariate treatments unless explicitly noted.

Assumptions

- We will be dealing with a cross-sectional random sample of n individuals
- **Overlap:**

$$0 < P(\tau_i = 1 | X_i = x) < 1$$

- **Conditional Independence**: Conditional on the covariates, X, the potential outcomes, y_0, y_1, and the treatment, τ, are independent

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$\begin{aligned} y_0 &= x\beta_0 + \varepsilon_0 \\ y_1 &= x\beta_1 + \varepsilon_1 \\ y &= \tau y_1 + (1-\tau)\, y_0 \end{aligned}$$

TREATMENT MODEL:

$$\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$
\begin{aligned}
y_0 &= x\beta_0 + \varepsilon_0 \\
y_1 &= x\beta_1 + \varepsilon_1 \\
y &= \tau y_1 + (1-\tau)\, y_0
\end{aligned}
$$

TREATMENT MODEL:

$$
\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}
$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$\begin{aligned} y_0 &= x\beta_0 + \varepsilon_0 \\ y_1 &= x\beta_1 + \varepsilon_1 \\ y &= \tau y_1 + (1-\tau)\, y_0 \end{aligned}$$

TREATMENT MODEL:

$$\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$
\begin{aligned}
y_0 &= x\beta_0 + \varepsilon_0 \\
y_1 &= x\beta_1 + \varepsilon_1 \\
y &= \tau y_1 + (1-\tau)\, y_0
\end{aligned}
$$

TREATMENT MODEL:

$$
\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}
$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$\begin{aligned} y_0 &= x\beta_0 + \varepsilon_0 \\ y_1 &= x\beta_1 + \varepsilon_1 \\ y &= \tau y_1 + (1-\tau)\, y_0 \end{aligned}$$

TREATMENT MODEL:

$$\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$\begin{aligned} y_0 &= x\beta_0 + \varepsilon_0 \\ y_1 &= x\beta_1 + \varepsilon_1 \\ y &= \tau y_1 + (1-\tau)\, y_0 \end{aligned}$$

TREATMENT MODEL:

$$\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

General Framework Illustrated with a Linear Example

OUTCOME MODEL:

$$\begin{aligned} y_0 &= x\beta_0 + \varepsilon_0 \\ y_1 &= x\beta_1 + \varepsilon_1 \\ y &= \tau y_1 + (1-\tau)\, y_0 \end{aligned}$$

TREATMENT MODEL:

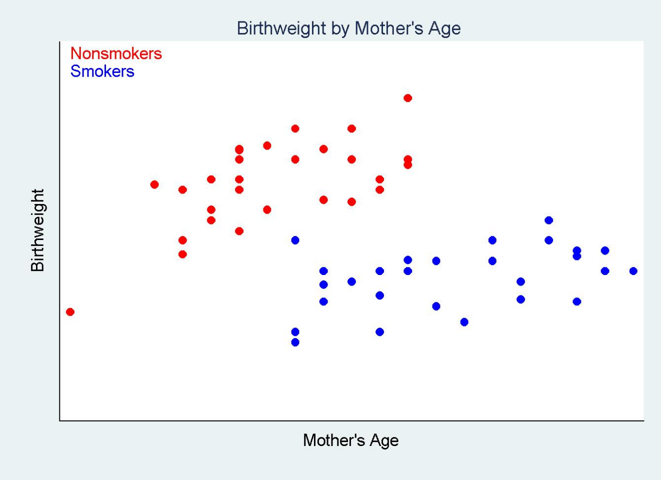
$$\tau = \begin{cases} 1 & \text{if} \quad w\gamma + \eta > 0 \\ 0 & \text{otherwise} \end{cases}$$

- w refers to the covariates that determine the treatment
- y_0 and y_1 are not observed. Only y, x, w, and τ are observed
- The random disturbances η, ε_0, and ε_1 are independent
- The functional forms for the outcome model do not need to be linear
- All the estimators we will see arise from combinations of the outcome model and the treatment model

Estimators Discussed Today

- Regression Adjustment (RA)
- Inverse Probability Weighting (IPW)
- Augmented Inverse Probability Weighting (AIPW)
- Inverse Probability Weighted Regression Adjustment (IPWRA)
- Nearest Neighbor Matching
- Propensity Score Matching

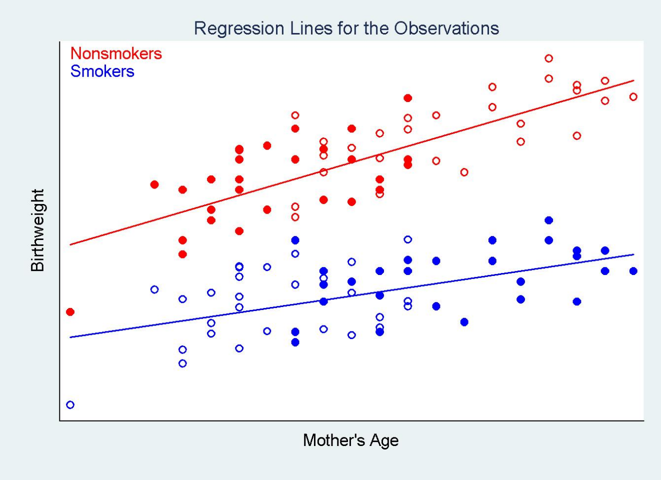
Effect of Smoking Mothers on Birthweight



Regression Adjustment (RA)

- We model the potential outcome and do not say anything about the treatment mechanism
- A conditional expectation is estimated for the treatment and control groups.
- The results from the estimations are used to compute POMs and thereafter ATEs, and ATETs.

Graphical Representation of RA Estimation



Models for the Potential Outcome

Outcome Model	$E(y\|x,z,\tau)$
`linear`	$x\beta_\tau$
`logit`	$exp(x\beta_\tau)/\{1+exp(x\beta_\tau)\}$
`probit`	$\Phi(x\beta_\tau)$
`poisson`	$exp(x\beta_\tau)$
`hetprobit`	$\Phi(x\beta_\tau/z\alpha_\tau)$

Data from Cattaneo (2010) Journal of Econometrics

`bweight`:	infant birth weight (grams)
`lbweight`:	1 if low birthweight baby
`mbsmoke`:	1 if mother smoked
`prenatal`:	trimester of first prenatal care visit
`fbaby`:	1if first baby
`mmarried`:	1 if mother married
`mage`:	mother´s age
`fage`:	father´s age
`alcohol`:	1 if alcohol consumed during pregnancy

- Sample of newborns from the United States from 1997

Data from Cattaneo (2010) Journal of Econometrics

`bweight`:	infant birth weight (grams)
`lbweight`:	1 if low birthweight baby
`mbsmoke`:	1 if mother smoked
`prenatal`:	trimester of first prenatal care visit
`fbaby`:	1if first baby
`mmarried`:	1 if mother married
`mage`:	mother´s age
`fage`:	father´s age
`alcohol`:	1 if alcohol consumed during pregnancy

- Sample of newborns from the United States from 1997

RA Linear Outcome Average Treatment Effect (ATE)

```
. teffects ra (bweight prenatal1 mmarried mage fbaby) (mbsmoke)
Iteration 0:   EE criterion =  7.734e-24
Iteration 1:   EE criterion =  1.196e-25
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)	-239.6392	23.82402	-10.06	0.000	-286.3334	-192.945
POmean						
mbsmoke nonsmoker	3403.242	9.525207	357.29	0.000	3384.573	3421.911

RA Average Treatment Effect on the Treated (ATET)

```
. teffects ra (bweight prenatal1 mmarried mage fbaby) (mbsmoke), atet
Iteration 0:   EE criterion =  7.629e-24
Iteration 1:   EE criterion =  2.697e-26
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : linear
Treatment model: none
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATET         |
     mbsmoke |
    (smoker  |
          vs |
  nonsmoker) |  -223.3017    22.7422    -9.82   0.000    -267.8755   -178.7278
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3360.961   12.75749   263.45   0.000     3335.957    3385.966
------------------------------------------------------------------------------
```

RA Probit Outcome ATE

```
. teffects ra (lbweight prenatal1 mmarried mage fbaby, probit) (mbsmoke)
Iteration 0:   EE criterion =  1.018e-18
Iteration 1:   EE criterion =  6.251e-34
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : probit
Treatment model: none
```

lbweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	.0500546	.0118733	4.22	0.000	.0267833	.0733259
POmean						
mbsmoke nonsmoker	.0517931	.003734	13.87	0.000	.0444745	.0591116

RA Probit ATET

```
. teffects ra (lbweight prenatal1 mmarried mage fbaby, probit) (mbsmoke), atet
Iteration 0:   EE criterion =  1.018e-18
Iteration 1:   EE criterion =  2.165e-34
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : probit
Treatment model: none
```

lbweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
mbsmoke (smoker vs nonsmoker)	.0458142	.0119394	3.84	0.000	.0224134	.0692149
POmean						
mbsmoke nonsmoker	.0641478	.0054295	11.81	0.000	.0535063	.0747894

Inverse Probability Weighting (IPW)

- In contrast to RA estimators, IPW estimate models for the treatment
- We fit a model for the treatment and compute the probabilities of treatment
- We then compute a weighted average, using the inverse of the probability of being in each group.

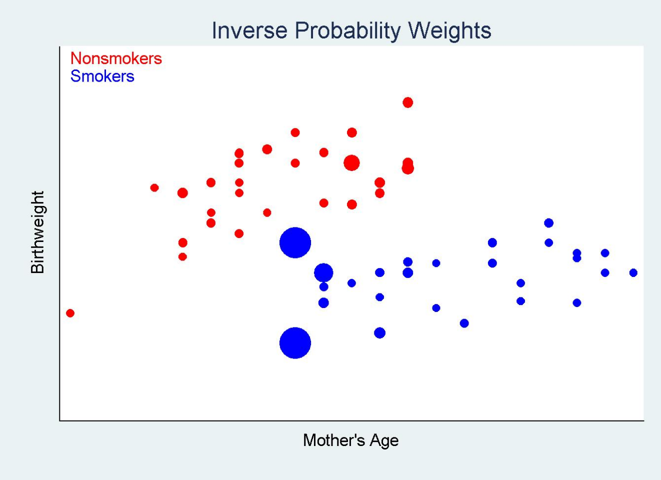
Inverse Probability Weight Calculation

```
. logistic mbsmoke mmarried alcohol mage fedu
Logistic regression                               Number of obs     =         60
                                                  LR chi2(4)        =      46.50
                                                  Prob > chi2       =     0.0000
Log likelihood = -18.339432                       Pseudo R2         =     0.5590
```

mbsmoke	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mmarried	.0785086	.0909212	-2.20	0.028	.0081122	.7597976
alcohol	18.81727	27.98003	1.97	0.048	1.020649	346.9259
mage	2.147569	.459327	3.57	0.000	1.41218	3.265909
fedu	.8189843	.1157528	-1.41	0.158	.6208252	1.080393
_cons	4.46e-07	2.12e-06	-3.07	0.002	3.96e-11	.0050329

```
. predict ps
(option pr assumed; Pr(mbsmoke))
. replace ps = 1/ps if mbsmoke==1
(30 real changes made)
. replace ps = 1/(1-ps) if mbsmoke==0
(30 real changes made)
```

Inverse Propability Weighting Graphically



Treatment Models

Treatment Model	$P(\tau \vert w, z)$
`logit`	$exp(w\gamma_\tau) / \{1 + exp(w\gamma_\tau)\}$
`probit`	$\Phi(w\gamma_\tau)$
`hetprobit`	$\Phi(w\gamma_\tau / z\theta_\tau)$

- Only the logit model is available for multivalued treatments

$$P(\tau|w) = \frac{exp(w\gamma_\tau)}{1 + \sum_{k=1}^{K} exp(w\gamma_k)}$$

Treatment Models

Treatment Model	$P(\tau \vert w, z)$
`logit`	$exp(w\gamma_\tau) / \{1 + exp(w\gamma_\tau)\}$
`probit`	$\Phi(w\gamma_\tau)$
`hetprobit`	$\Phi(w\gamma_\tau / z\theta_\tau)$

- Only the logit model is available for multivalued treatments

$$P(\tau|w) = \frac{exp(w\gamma_\tau)}{1 + \sum_{k=1}^{K} exp(w\gamma_k)}$$

IPW ATE

```
. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu)
Iteration 0:   EE criterion =  1.713e-21
Iteration 1:   EE criterion =  4.794e-27
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : inverse probability weighted
Outcome model  : weighted mean
Treatment model: logit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -231.7203   25.17975    -9.20   0.000    -281.0717   -182.3689
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3403.527   9.576358   355.41   0.000     3384.757    3422.296
------------------------------------------------------------------------------
```

IPW ATET

```
. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu), atet
Iteration 0:   EE criterion =  1.714e-21
Iteration 1:   EE criterion =  3.735e-27
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : inverse probability weighted
Outcome model  : weighted mean
Treatment model: logit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATET         |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -225.6992    23.7133    -9.52   0.000    -272.1764    -179.222
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3363.359   14.28989   235.37   0.000     3335.351    3391.367
------------------------------------------------------------------------------
```

IPW ATE

```
. teffects ipw (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu, probit)
Iteration 0:   EE criterion =  4.622e-21
Iteration 1:   EE criterion =  8.622e-26
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : inverse probability weighted
Outcome model  : weighted mean
Treatment model: probit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -230.6886   25.81524    -8.94   0.000    -281.2856   -180.0917
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3403.463   9.571369   355.59   0.000     3384.703    3422.222
------------------------------------------------------------------------------
```

IPW ATET

```
. teffects ipw (bweight) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu, probit), atet
Iteration 0:   EE criterion =  4.621e-21
Iteration 1:   EE criterion =  7.103e-27
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : inverse probability weighted
Outcome model  : weighted mean
Treatment model: probit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATET         |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -225.1773   23.66458    -9.52   0.000     -271.559   -178.7955
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3362.837   14.20149   236.79   0.000     3335.003    3390.671
------------------------------------------------------------------------------
```

Doubly Robust Estimators

- Doubly robust estimators model both the treatment and the outcome model
- These models are interesting because they are consistent even if one of the models is misspecified
- Augmented Inverse Probability Weighting (AIPW) and Inverse Probability Weighted Regression Adjustment(IPWRA) have this property

Double Robust Estimators AIPW

- Estimate a treatment model and compute inverse-probability weights
- Estimate separate regression model of the outcome for each treatment level
  - We allow the outcome model to be estimated by nonlinear least squares or weighted nonlinear least squares
- Compute the weighted means of the treatment-specific predicted outcomes, where the weights are the inverse-probability weights computed in step.

ATE for AIPW

```
. teffects aipw (bweight prenatal1 mmarried mage fbaby) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu)
Iteration 0:   EE criterion =  1.721e-21
Iteration 1:   EE criterion =  2.247e-26
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : augmented IPW
Outcome model  : linear by ML
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-232.0409	25.66973	-9.04	0.000	-282.3527	-181.7292
POmean						
mbsmoke nonsmoker	3403.457	9.570043	355.64	0.000	3384.7	3422.214

ATE for AIPW with Nonlinear Least Squares

```
. teffects aipw (bweight prenatal1 mmarried mage fbaby, poisson) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu), nls
Iteration 0:   EE criterion =  .00018418
Iteration 1:   EE criterion =  1.991e-17
Treatment-effects estimation                Number of obs      =      4642
Estimator      : augmented IPW
Outcome model  : Poisson by NLS
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-232.1593	25.69692	-9.03	0.000	-282.5244	-181.7943
POmean						
mbsmoke nonsmoker	3403.444	9.57036	355.62	0.000	3384.687	3422.202

Displaying Treatment and Outcome Equations

```
. teffects aipw (bweight prenatal1 mmarried mage fbaby, poisson) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu), aequations nolog
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : augmented IPW
Outcome model  : Poisson by ML
Treatment model: logit
------------------------------------------------------------------------------
             |               Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
  nonsmoker) |  -232.1369   25.68896    -9.04   0.000    -282.4864   -181.7875
-------------+----------------------------------------------------------------
POmean       |
     mbsmoke |
   nonsmoker |   3403.444   9.570363   355.62   0.000     3384.686    3422.202
-------------+----------------------------------------------------------------
OME0         |
   prenatal1 |   .0191803   .0082502     2.32   0.020     .0030102    .0353503
    mmarried |   .0480049   .0080048     6.00   0.000     .0323158    .0636939
        mage |   .0007522   .0006106     1.23   0.218    -.0004447     .001949
       fbaby |  -.0209166   .0057619    -3.63   0.000    -.0322097   -.0096235
       _cons |   8.072261   .0159896   504.84   0.000     8.040922      8.1036
-------------+----------------------------------------------------------------
OME1         |
   prenatal1 |   .0080848    .012943     0.62   0.532    -.0172831    .0334526
    mmarried |   .0426096   .0130351     3.27   0.001     .0170612    .0681579
        mage |  -.0023601   .0013552    -1.74   0.082    -.0050163    .0002961
       fbaby |   .0131662   .0126163     1.04   0.297    -.0115613    .0378937
       _cons |    8.07972   .0334184   241.77   0.000     8.014221    8.145219
-------------+----------------------------------------------------------------
TME1         |
    mmarried |  -1.145706   .0975846   -11.74   0.000    -1.336969   -.9544439
        mage |    .321518   .0657363     4.89   0.000     .1926773    .4503588
             |
c.mage#c.mage|  -.0060368   .0012234    -4.93   0.000    -.0084346   -.0036389
             |
       fbaby |  -.3864258   .0894428    -4.32   0.000    -.5617305   -.2111211
        medu |  -.1420833   .0179132    -7.93   0.000    -.1771926    -.106974
       _cons |  -2.950915   .8302955    -3.55   0.000    -4.578264   -1.323565
------------------------------------------------------------------------------
```

Double Robust Estimators Inverse Probability Weighted Regression Adjustment (IPWRA)

- Estimate a treatment model and compute inverse-probability weights
- Use the estimated inverse-probability weights and fit weighted regression models of the outcome for each treatment level
- Compute the means of the treatment-specific predicted outcomes

ATET for Inverse Probability Weighted Regression Adjustment

```
. teffects ipwra (bweight prenatal1 mmarried mage fbaby) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu), atet
Iteration 0:   EE criterion =  4.620e-21
Iteration 1:   EE criterion =  1.345e-26
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
mbsmoke (smoker vs nonsmoker)	-224.0108	23.846	-9.39	0.000	-270.7481	-177.2735
POmean						
mbsmoke nonsmoker	3361.671	14.54939	231.05	0.000	3333.154	3390.187

Displaying Treatment and Outcome Equations

```
. teffects ipwra (bweight prenatal1 mmarried mage fbaby) ///
> (mbsmoke mmarried c.mage##c.mage fbaby medu), atet aequations
Iteration 0:   EE criterion =  4.620e-21
Iteration 1:   EE criterion =  1.345e-26
Treatment-effects estimation                    Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
mbsmoke (smoker vs nonsmoker)	-224.0108	23.846	-9.39	0.000	-270.7481	-177.2735
POmean						
mbsmoke nonsmoker	3361.671	14.54939	231.05	0.000	3333.154	3390.187
OME0						
prenatal1	77.07926	40.4633	1.90	0.057	-2.227341	156.3859
mmarried	138.9961	29.48776	4.71	0.000	81.20114	196.791
mage	4.482273	3.033008	1.48	0.139	-1.462313	10.42686
fbaby	-73.85266	32.55461	-2.27	0.023	-137.6585	-10.0468
_cons	3157.337	72.75786	43.40	0.000	3014.734	3299.939
OME1						
prenatal1	25.11133	40.37541	0.62	0.534	-54.02302	104.2457
mmarried	133.6617	40.86443	3.27	0.001	53.5689	213.7545
mage	-7.370881	4.21817	-1.75	0.081	-15.63834	.8965804
fbaby	41.43991	39.70712	1.04	0.297	-36.38461	119.2644
_cons	3227.169	104.4059	30.91	0.000	3022.537	3431.801
TME1						
mmarried	-1.145706	.0975846	-11.74	0.000	-1.336969	-.9544439
mage	.321518	.0657363	4.89	0.000	.1926773	.4503588
c.mage#c.mage	-.0060368	.0012234	-4.93	0.000	-.0084346	-.0036389
fbaby	-.3864258	.0894428	-4.32	0.000	-.5617305	-.2111211
medu	-.1420833	.0179132	-7.93	0.000	-.1771926	-.106974
_cons	-2.950915	.8302955	-3.55	0.000	-4.578264	-1.323565

Nearest Neighbor Matching

- Can be understood as an outcome model within our framework
- Matches the closest individuals in terms of covariates
- Is a nonparametric estimate with an asymptotic bias.
- These estimators are nondifferentiable therefore the bootstrap is not allowed
- These estimators do not allow for multivalued treatments

Nearest Neighbor Matching

- Can be understood as an outcome model within our framework
- Matches the closest individuals in terms of covariates
- Is a nonparametric estimate with an asymptotic bias.
- These estimators are nondifferentiable therefore the bootstrap is not allowed
- These estimators do not allow for multivalued treatments

ATE with Nearest Neighbor Matching

```
. teffects nnmatch (bweight mage prenatal1 mmarried fbaby) (mbsmoke)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : nearest-neighbor matching     Matches: requested =         1
Outcome model  : matching                                     min =         1
Distance metric: Mahalanobis                                  max =       139
```

bweight	Coef.	AI Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-240.3306	28.43006	-8.45	0.000	-296.0525	-184.6087

Exact Matching and Different Distance

```
. teffects nnmatch (bweight mage) (mbsmoke), ///
> ematch(prenatal1 mmarried fbaby) metric(euclidean)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : nearest-neighbor matching     Matches: requested =         1
Outcome model  : matching                                     min =         1
Distance metric: Euclidean                                    max =       139
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -240.3306   28.43006    -8.45   0.000    -296.0525   -184.6087
------------------------------------------------------------------------------
```

Bias Adjustment

```
. teffects nnmatch (bweight mage fage) (mbsmoke), ///
> ematch(prenatal1 mmarried fbaby) biasadj(mage fage)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : nearest-neighbor matching     Matches: requested =         1
Outcome model  : matching                                     min =         1
Distance metric: Mahalanobis                                  max =        25
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -223.8389   26.19973    -8.54   0.000    -275.1894   -172.4883
------------------------------------------------------------------------------
```

Propensity Score Matching

- Can be classified within the class of treatment models
- Estimate the treatment probabilities (propensity scores)
- Assign values to unobserved outcomes based on observed ones with similar propensity scores
- Estimate ATE
- These estimators are nondifferentiable therefore the bootstrap is not allowed
- These estimators do not allow for multivalued treatments

Propensity Score Matching

- Can be classified within the class of treatment models
- Estimate the treatment probabilities (propensity scores)
- Assign values to unobserved outcomes based on observed ones with similar propensity scores
- Estimate ATE
- These estimators are nondifferentiable therefore the bootstrap is not allowed
- These estimators do not allow for multivalued treatments

Propensity Score Matching controlling matches

```
. teffects psmatch (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu), ///
> nneighbor(2)
Treatment-effects estimation                   Number of obs      =      4642
Estimator      : propensity-score matching     Matches: requested =         2
Outcome model  : matching                                     min =         2
Treatment model: logit                                        max =        74
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
         vs  |
 nonsmoker)  |  -214.2469   27.47783    -7.80   0.000    -268.1025   -160.3914
------------------------------------------------------------------------------
```

Conclusion

- We have presented a host of treatment effects estimators within a unified framework
- The estimators are parametric and nonparametric and in the parametric cases can be consistent under misspecification of the potential outcome or treatment models
- The estimators provide estimates and inference for quantities of interest for researchers, POM, ATE, ATET.

Double Robustness I

- Let $P(\tau|x,z,\hat{\gamma}) =: M_P(\hat{\gamma})$ be our estimated conditional treatment probabilities
- Let $E(y|x,z,\tau,\hat{\beta}) =: M_E(\hat{\beta}_\tau)$ define our estimated conditional means
- We define the following estimators for the POMs

$$\hat{E}(y_1) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\tau_i y_i}{M_P(\hat{\gamma})} - \frac{\{\tau_i - M_P(\hat{\gamma})\}}{M_P(\hat{\gamma})} M_E(\hat{\beta}_1)\right]$$

$$\hat{E}(y_0) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{(1-\tau_i)y_i}{1-M_P(\hat{\gamma})} - \frac{\{\tau_i - M_P(\hat{\gamma})\}}{1-M_P(\hat{\gamma})} M_E(\hat{\beta}_0)\right]$$

Double Robustness I

- Let $P\left(\tau|x,z,\hat{\gamma}\right) =: M_P\left(\hat{\gamma}\right)$ be our estimated conditional treatment probabilities
- Let $E(y|x,z,\tau,\hat{\beta}) =: M_E(\hat{\beta}_\tau)$ define our estimated conditional means
- We define the following estimators for the POMs

$$
\begin{aligned}
\hat{E}\left(y_1\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\tau_i y_i}{M_P\left(\hat{\gamma}\right)} - \frac{\left\{\tau_i - M_P\left(\hat{\gamma}\right)\right\}}{M_P\left(\hat{\gamma}\right)} M_E(\hat{\beta}_1)\right] \\
\hat{E}\left(y_0\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\left(1-\tau_i\right) y_i}{1-M_P\left(\hat{\gamma}\right)} - \frac{\left\{\tau_i - M_P\left(\hat{\gamma}\right)\right\}}{1-M_P\left(\hat{\gamma}\right)} M_E(\hat{\beta}_0)\right]
\end{aligned}
$$

Intuition Behind Double Robustness II

- We will focus on $\hat{E}(y_1)$ (a similar argument follows for $\hat{E}(y_0)$)
- By the law of large numbers it follows that $\hat{E}(y_1)$ has the following probability limit:

$$\hat{E}(y_1) \xrightarrow{p} E\left[\frac{\tau y}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1)\right]$$

$$\begin{aligned}
&= E\left[\frac{\tau y_1}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1 - y_1\right] \\
&= E\left[\frac{\tau y_1}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1 - y_1 \frac{M_P(\gamma)}{M_P(\gamma)}\right] \\
&= E\left[\frac{y_1(\tau - M_P(\gamma))}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1\right] \\
&= E\left[\frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} (y_1 - M_E(\beta_1)) + y_1\right] \\
&= E(y_1) + E\left[\frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} (y_1 - M_E(\beta_1))\right]
\end{aligned}$$

Intuition Behind Double Robustness II

- We will focus on $\hat{E}(y_1)$ (a similar argument follows for $\hat{E}(y_0)$)
- By the law of large numbers it follows that $\hat{E}(y_1)$ has the following probability limit:

$$\hat{E}(y_1) \xrightarrow{p} E\left[\frac{\tau y}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1)\right]$$

$$\begin{aligned}
&= E\left[\frac{\tau y_1}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1 - y_1\right] \\
&= E\left[\frac{\tau y_1}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1 - y_1 \frac{M_P(\gamma)}{M_P(\gamma)}\right] \\
&= E\left[\frac{y_1(\tau - M_P(\gamma))}{M_P(\gamma)} - \frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} M_E(\beta_1) + y_1\right] \\
&= E\left[\frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} (y_1 - M_E(\beta_1)) + y_1\right] \\
&= E(y_1) + E\left[\frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)} (y_1 - M_E(\beta_1))\right]
\end{aligned}$$

Intuition Behind Double Robustness III

$$\hat{E}\left(y_1\right) \xrightarrow{p} E\left(y_1\right) + E\left[\frac{\{\tau - M_P\left(\gamma\right)\}}{M_P\left(\gamma\right)}\left(y_1 - M_E(\beta_1)\right)\right]$$

- Given conditional independence of treatment and outcome conditional on the regressors by the law of iterated expectations:
  - If the outcome model is correctly specified $E\left[y_1 - M_E(\beta_1)\right] = 0$. This implies that even if the treatment model is incorrectly specified, $\hat{E}\left(y_1\right) \xrightarrow{p} E\left(y_1\right)$
  - Similarly if the treatment model is correctly specified $E\left[\tau - M_P\left(\gamma\right)\right] = 0$. Thus, even if $E\left[y_1 - M_E(\beta_1)\right] \neq 0$ we have that $\hat{E}\left(y_1\right) \xrightarrow{p} E\left(y_1\right)$

Intuition Behind Double Robustness III

$$\hat{E}(y_1) \xrightarrow{p} E(y_1) + E\left[\frac{\{\tau - M_P(\gamma)\}}{M_P(\gamma)}(y_1 - M_E(\beta_1))\right]$$

- Given conditional independence of treatment and outcome conditional on the regressors by the law of iterated expectations:
  - If the outcome model is correctly specified $E[y_1 - M_E(\beta_1)] = 0$. This implies that even if the treatment model is incorrectly specified, $\hat{E}(y_1) \xrightarrow{p} E(y_1)$
  - Similarly if the treatment model is correctly specified $E[\tau - M_P(\gamma)] = 0$. Thus, even if $E[y_1 - M_E(\beta_1)] \neq 0$ we have that $\hat{E}(y_1) \xrightarrow{p} E(y_1)$

Estimation I: An Example of Moment Based Estimation

We define the projection model by:

$$\begin{aligned} y &= X\beta + \varepsilon \\ E\left(X'\varepsilon\right) &= 0 \end{aligned}$$

β is then given by:

$$\begin{aligned} 0 &= E\left(X'\varepsilon\right) \\ 0 &= E\left(X'\left\{y - X\beta\right\}\right) \\ \beta &= E\left(X'X\right)^{-1} E\left(X'y\right) \end{aligned}$$

A consistent estimator of β is:

$$\hat{\beta} = \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'y}{n}\right)$$

Estimation I: An Example of Moment Based Estimation

We define the projection model by:

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
E\left(X'\varepsilon\right) &= 0
\end{aligned}
$$

β is then given by:

$$
\begin{aligned}
0 &= E\left(X'\varepsilon\right) \\
0 &= E\left(X'\left\{y - X\beta\right\}\right) \\
\beta &= E\left(X'X\right)^{-1} E\left(X'y\right)
\end{aligned}
$$

A consistent estimator of β is:

$$
\hat{\beta} = \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'y}{n}\right)
$$

Estimation I: An Example of Moment Based Estimation

We define the projection model by:

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
E\left(X'\varepsilon\right) &= 0
\end{aligned}
$$

β is then given by:

$$
\begin{aligned}
0 &= E\left(X'\varepsilon\right) \\
0 &= E\left(X'\left\{y - X\beta\right\}\right) \\
\beta &= E\left(X'X\right)^{-1} E\left(X'y\right)
\end{aligned}
$$

A consistent estimator of β is:

$$
\hat{\beta} = \left(\frac{X'X}{n}\right)^{-1} \left(\frac{X'y}{n}\right)
$$

Estimation I: An Example of Moment Based Estimation

We define the projection model by:

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
E\left(X'\varepsilon\right) &= 0
\end{aligned}
$$

β is then given by:

$$
\begin{aligned}
0 &= E\left(X'\varepsilon\right) \\
0 &= E\left(X'\left\{y - X\beta\right\}\right) \\
\beta &= E\left(X'X\right)^{-1} E\left(X'y\right)
\end{aligned}
$$

A consistent estimator of β is:

$$
\hat{\beta} = \left(\frac{X'X}{n}\right)^{-1} \left(\frac{X'y}{n}\right)
$$

Estimation I: An Example of Moment Based Estimation

We define the projection model by:

$$\begin{aligned} y &= X\beta + \varepsilon \\ E\left(X'\varepsilon\right) &= 0 \end{aligned}$$

β is then given by:

$$\begin{aligned} 0 &= E\left(X'\varepsilon\right) \\ 0 &= E\left(X'\left\{y - X\beta\right\}\right) \\ \beta &= E\left(X'X\right)^{-1} E\left(X'y\right) \end{aligned}$$

A consistent estimator of β is:

$$\hat{\beta} = \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'y}{n}\right)$$

Estimation II: Methodology We Employ

- The different specifications for the outcome generate moment conditions
- We can then use GMM to estimate the parameters of interest
- For the linear model:

$$0 = E\left[\tau(y - x\beta_1)'x + (1-\tau)(y - x\beta_0)'x\right]$$

- For the probit and logit models

$$0 = E\left(\tau\left[\frac{g(x\beta_1)\{y - G(x\beta_1)\}}{G(x\beta_1)\{1 - G(x\beta_1)\}}\right] + (1-\tau)\left[\frac{g(x\beta_1)\{y - G(x\beta_0)\}}{G(x\beta_0)\{1 - G(x\beta_0)\}}\right]\right)$$

  - $G(.)$ is either the standard normal CDF or the logistic function
  - $g(.)$ is the derivative of $G(.)$

Estimation II: Methodology We Employ

- The different specifications for the outcome generate moment conditions
- We can then use GMM to estimate the parameters of interest
- For the linear model:

$$0 = E\left[\tau(y - x\beta_1)'x + (1-\tau)\,(y - x\beta_0)'x\right]$$

- For the probit and logit models

$$0 = E\left(\tau\left[\frac{g\,(x\beta_1)\,\{y - G\,(x\beta_1)\}}{G\,(x\beta_1)\,\{1 - G\,(x\beta_1)\}}\right] + (1-\tau)\left[\frac{g\,(x\beta_1)\,\{y - G\,(x\beta_0)\}}{G\,(x\beta_0)\,\{1 - G\,(x\beta_0)\}}\right]\right)$$

  - $G(.)$ is either the standard normal CDF or the logistic function
  - $g(.)$ is the derivative of $G(.)$

Distance

The distance function is given by:

$$\left\|x_i - x_j\right\|_S = \left\{\left(x_i - x_j\right)' S^{-1} \left(x_i - x_j\right)\right\}^{1/2}$$

where S can be:

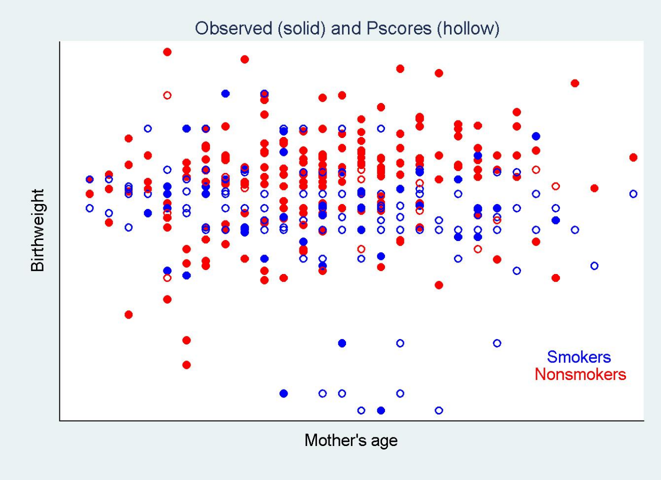
$$S = \begin{cases} \frac{(X - x1_n)' W (X - x1_n)}{\sum_{i=1}^{n} w_i - 1} & \text{if} \quad \text{metric is mahalanobis} \\ \text{diagonal}\left\{\frac{(X - x1_n)' W (X - x1_n)}{\sum_{i=1}^{n} w_i - 1}\right\} & \text{if} \quad \text{metric is ivariance} \\ I_k & \text{if} \quad \text{metric is euclidean} \end{cases}$$

Above 1_n is an n vector of ones, W is a matrix of frequency weights

ATE for Propensity Score Matching

```
. teffects psmatch (bweight) (mbsmoke mmarried c.mage##c.mage fbaby medu), ///
> generate(ps)
Treatment-effects estimation                   Number of obs      =       4642
Estimator      : propensity-score matching     Matches: requested =          1
Outcome model  : matching                                     min =          1
Treatment model: logit                                        max =         74
------------------------------------------------------------------------------
             |              AI Robust
     bweight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ATE          |
     mbsmoke |
    (smoker  |
          vs |
  nonsmoker) |  -210.9683     32.021    -6.59   0.000    -273.7284   -148.2083
------------------------------------------------------------------------------
```

Matches Generated by the Estimator



A Nonsequitur

- Imbens and Wooldridge (2009) JEL for a recent survey
- Regression Discontinuity. Lee and Lemiux (2010) JEL
- Nonparametric Multivariate Treatment Effects. See Cattaneo 2010 in the New Palgrave Dictionary and Cattaneo 2010 JOE.
- Stata also offers estimation in the presence of endogeneity