Estimating average treatment effects from observational data using `teffects`

David M. Drukker

Director of Econometrics
Stata

2014 German Stata Users Group meeting
June 13, 2014
1. What do we want to estimate?

2. Estimators: Overview
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 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)

+-----------------+-----------------+-----------------+
<table>
<thead>
<tr>
<th>mother_id</th>
<th>bw_smoke</th>
<th>bw_nosmoke</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3183</td>
<td>3509</td>
</tr>
<tr>
<td>2</td>
<td>3060</td>
<td>3316</td>
</tr>
<tr>
<td>3</td>
<td>3165</td>
<td>3474</td>
</tr>
<tr>
<td>4</td>
<td>3176</td>
<td>3495</td>
</tr>
<tr>
<td>5</td>
<td>3241</td>
<td>3413</td>
</tr>
</tbody>
</table>
+-----------------+-----------------+-----------------+
```

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.
What do we want to estimate?

**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

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>bw_smoke</td>
<td>3171.72</td>
<td>.9088219</td>
<td>3169.938 3173.501</td>
</tr>
<tr>
<td>bw_nosmoke</td>
<td>3402.599</td>
<td>1.529189</td>
<td>3399.601 3405.597</td>
</tr>
</tbody>
</table>
```

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

|         | 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

\[
ATE = \mathbb{E}[bw_{\text{smoke}} - bw_{\text{nosmoke}}] = \mathbb{E}[bw_{\text{smoke}}] - \mathbb{E}[bw_{\text{nosmoke}}]
\]
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 analysis to estimate treatment effects.

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...
What do we want to estimate?

Difference in means

```
. regress bweight ibn.mbsmoke, noconstant
Source | SS    df   MS
-------+--------+------
Model  | 5.2512e+10   2  2.6256e+10
Residual | 1.5016e+09  4640  323622.478
       | 5.4014e+10  4642  11635851.6
Number of obs =  4642
F(  2,  4640) =81131.59
Prob > F =  0.0000
R-squared = 0.9722
Adj R-squared = 0.9722
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 (nonsmoker vs smoker) | -275.2519  | -317.3096  -233.1942 |
```

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.
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 \( x_i \) in the population and each treatment level \( t \), \( 0 < P(t_i = t | x) < 1 \).
What do we want to estimate?

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.
What do we want to estimate?

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).
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.

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>outcome</td>
<td>Regression adjustment (RA)</td>
</tr>
<tr>
<td>treatment</td>
<td>Inverse-probability weighted (IPW)</td>
</tr>
<tr>
<td>outcome and treatment</td>
<td>Augmented IPW (AIPW)</td>
</tr>
<tr>
<td>outcome and treatment</td>
<td>IPW RA (IPWRA)</td>
</tr>
<tr>
<td>outcome (nonparametrically)</td>
<td>Nearest-neighbor matching (NNMATCH)</td>
</tr>
<tr>
<td>treatment</td>
<td>Propensity-score matching (PSMATCH)</td>
</tr>
</tbody>
</table>
Regression adjustment (RA) estimators:

- RA estimators run separate regressions for each treatment level, then
  - means of predicted outcomes using all the data and the estimated coefficients for treatment level \( i \) all the data estimate \( \text{POM}_i \)
  - use differences of \( \text{POM}s \), or conditional on the treated \( \text{POM}s \), to estimate ATEs or ATETs

- Formally, the CMI assumption implies that we can estimate \( \mathbb{E}(y_t|x_i) \) directly from the observations for which person \( i \) gets treatment \( t \)
  - \( y_t \) is the potential outcome for treatment level \( t \)
  - Averages of predicted \( \mathbb{E}(y_t|x_i) \) yield estimates of the POM \( \mathbb{E}[y_t] \)

See (Cameron and Trivedi, 2005, chapter 25), (Wooldridge, 2010, chapter 21), and (Vittinghoff et al., 2012, chapter 9)
. use cattaneo2
. 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

<table>
<thead>
<tr>
<th>bweight</th>
<th>Robust</th>
<th></th>
<th></th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
</tr>
<tr>
<td>ATE</td>
<td>-230.9541</td>
<td>24.34012</td>
<td>-9.49</td>
<td>0.000</td>
</tr>
<tr>
<td>mbsmoke (smoker vs nonsmoker)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POmean</td>
<td>3402.548</td>
<td>9.546721</td>
<td>356.41</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**RA with linear regression to model outcome**
. teffects ra (bweight mmarried prenatal1 fbaby medu, poisson) (mbsmoke)
Iteration 0:   EE criterion =  3.926e-17
Iteration 1:   EE criterion =  1.612e-23
Treatment-effects estimation
Estimator : regression adjustment
Outcome model : Poisson
Treatment model: none

| bweight            | Robust 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.

. teffects ipw (bweight) (mbsmoke mmarried prenatal1 fbaby medu)
Iteration 0:  EE criterion =  1.701e-23
Iteration 1:  EE criterion =  6.339e-27
Treatment-effects estimation  Number of obs   =   4642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: logit

|                | Coef.  | Std. Err. |      z  |     P>|z| | [95% Conf. Interval] |
|----------------|--------|-----------|---------|--------|----------------------|
| bweight        |        |           |         |        |                      |
| ATE mbsmoke    | -231.1516 | 24.03183  | -9.62   | 0.000  | [-278.2531, -184.0501]  
| (smoker vs nonsmoker) |   |           |         |        |                      |
| P0mean mbsmoke | 3402.219  | 9.589812  | 354.77  | 0.000  | [3383.423, 3421.015]   |
| nonsmoker      |        |           |         |        |                      |

- IPW with logit to model treatment
Estimators: Overview

IPW example II

. teffects ipw (bweight) (mbsmoke mmarried prenatal1 fbaby medu, hetprobit(medu > ))
Iteration 0:  EE criterion =  7.158e-16
Iteration 1:  EE criterion =  1.681e-26
Treatment-effects estimation Number of obs  =  4642
Estimator : inverse-probability weights
Outcome model : weighted mean
Treatment model: heteroskedastic probit

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.  Std. Err.</td>
</tr>
<tr>
<td>bweight</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
</tr>
<tr>
<td>(smoker vs</td>
<td>-217.7521 28.5796</td>
</tr>
<tr>
<td>nonsmoker)</td>
<td></td>
</tr>
<tr>
<td>POmean</td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td>3401.788 9.570692</td>
</tr>
<tr>
<td>nonsmoker</td>
<td></td>
</tr>
</tbody>
</table>

- 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

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 =  2.179e-26
Treatment-effects estimation                     Number of obs  =  4642
Estimator : augmented IPW
Outcome model : linear by ML
Treatment model: logit

|                  | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|------------------|-------|-----------|------|------|----------------------|
|                  | bweight |            |      |      |                      |
| 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 =  8.662e-24
Treatment-effects estimation                       Number of obs =  4642
Estimator:      augmented IPW
Outcome model:  Poisson by ML
Treatment model: heteroskedastic probit

| bweight       | Coef. | Std. Err.  | z     | P>|z|   | 95% Conf. Interval |
|---------------|-------|------------|-------|-------|-------------------|
| ATE           |       |            |       |       |                   |
| mbsmoke       | -220.496 | 28.30292      | -7.79 | 0.000 | -275.9687 -165.0233 |
| (smoker vs    |       |            |       |       |                   |
| nonsmoker)    |       |            |       |       |                   |
| P0mean        | 3402.429 | 9.557345 | 356.00 | 0.000 | 3383.697 3421.161 |
| mbsmoke       |       |            |       |       |                   |
| nonsmoker     |       |            |       |       |                   |
```

- **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).
. teffects ipwra (bweight mmarried prenatal1 fbaby medu) ///
> (mbsmoke mmarried prenatal1 fbaby medu)
Iteration 0:  EE criterion =  3.016e-22
Iteration 1:  EE criterion =  1.361e-25
Treatment-effects estimation
Number of obs  =  4642
Estimator     : IPW regression adjustment
Outcome model : linear
Treatment model: logit

|              | Robust Coef. | Std. Err. | z     | P>|z| |  [95% Conf. Interval] |
|--------------|--------------|-----------|-------|------|----------------------|
|              | bweight      |           |       |      |                      |
| ATE mbsmoke  |              |           |       |      |                      |
|   (smoker    |              |           |       |      |                      |
|     vs       |              |           |       |      |                      |
|     nonsmoker) | -227.4408   | 25.62591  | -8.88 | 0.000 | -277.6667 -177.215   |
|              |              |           |       |      |                      |
| P0mean mbsmoke | 3403.027    | 9.56025   | 355.96| 0.000 |  3384.289 3421.765  |
|  nonsmoker |              |           |       |      |                      |

- IPWRA with linear model for outcome and logit for treatment
. 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 =  6.131e-24
Treatment-effects estimation Number of obs =  4642
Estimator : IPW regression adjustment
Outcome model : Poisson
Treatment model: heteroskedastic probit

| 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 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 (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.
See Abadie and Imbens (2006) and Abadie and Imbens (2011) for formal results, rates of convergence, and the details of the bias-correction methods.


tefffect 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

<table>
<thead>
<tr>
<th></th>
<th>Number of obs</th>
<th>Matches: requested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimator: nearest-neighbor matching</td>
<td>4642</td>
<td>1</td>
</tr>
<tr>
<td>Outcome model: matching</td>
<td>min = 1</td>
<td>max = 645</td>
</tr>
<tr>
<td>Distance metric: Mahalanobis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                  | 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 (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 mmmarried 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

| 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**
. teffects psmatch (bweight) (mbsmoke mmarrried 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

| 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 and why QTE are useful

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 $\tau$th quantile of the distribution of the treated potential outcome $y_1$
- $q_0(\tau) = F_{y_0}^{-1}(\tau)$ is the $\tau$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
- $F_{y_j}(y) = \mathbb{E}_x[F_{y_j|x}(y|x)]$
- Keep in mind that $q_j(\tau) = F_{y_j}^{-1}(\tau) \neq \mathbb{E}[q_j(\tau|x)]$, where $q_j(\tau|x)$ is condition-on-$x$ quantile of the potential-outcome distribution
Imagine a study that followed middle-aged men for two years after suffering a heart attack

- Does exercise affect the time to a second heart attack?
- Some observations on the time to second heart attack are censored
- Observational data implies that treatment allocation depends on covariates
- We use a model for the outcome to adjust for this dependence
QTEs for survival data

- Exercise could help individuals with relatively strong hearts but not help those with weak hearts.
- For each treatment level, a strong-heart individual is in the .75 quantile of the marginal, over the covariates, distribution of time to second heart attack.
  - QTE(.75) is difference in .75 marginal quantiles.
- Weak-heart individual would be in the .25 quantile of the marginal distribution for each treatment level.
  - QTE(.25) is difference in .25 marginal quantiles.
- Our story indicates that the QTE(.75) should be significantly larger than the QTE(.25).
A regression-adjustment estimator for QTEs

- Estimate the $\theta_1$ parameters of $F(y|x, t = 1, \theta_1)$ the CDF conditional on covariates and conditional on treatment level
  - Conditional independence implies that this conditional on treatment level CDF estimates the CDF of the treated potential outcome
- Similarly, estimate the $\theta_0$ parameters of $F(y|x, t = 0, \theta_1)$
- At the point $y$,
  \[
  1/N \sum_{i=1}^{N} F(y|x_i, \hat{\theta}_1)
  \]
  estimates the marginal distribution of the treated potential outcome
- The $\hat{q}_{1,.75}$ that solves
  \[
  1/N \sum_{i=1}^{N} F(\hat{q}_{1,.75}|x_i, \hat{\theta}_1) = .75
  \]
  estimates the .75 marginal quantile for the treated potential outcome
Using mqgamma and poparms for estimation and inference

**A regression-adjustment estimator for QTEs**

- The \( \hat{q}_{0,.75} \) that solves

\[
\frac{1}{N} \sum_{i=1}^{N} F(\hat{q}_{0,.75} | x_i, \hat{\theta}_0) = .75
\]

estimates the .75 marginal quantile for the control potential outcome.

- \( \hat{q}_1(.75) - \hat{q}_0(.75) \) consistently estimates QTE(.75)

- See Drukker (2014) for details
mqgamma is a user-written command documented in Drukker (2014)

```
. use exercise
. mqgamma t active, treat(exercise) fail(fail) lns(health) quantile(.25 .75)
Iteration 0:  EE criterion =  .7032254
Iteration 1:  EE criterion =  .05262105
Iteration 2:  EE criterion =  .00028553
Iteration 3:  EE criterion =  6.892e-07
Iteration 4:  EE criterion =  4.706e-12
Iteration 5:  EE criterion =  1.604e-22
Gamma quantile-treatment-effect estimation       Number of obs  =   2000

                   | Coef.     Robust Std. Err.       z     P>|z|    [95% Conf. Interval]
-------------|-------------|-----------------|-------|---------|-------------------|-------------------|
         t     |       z     |             |       |         |                   |                   |
q25_0      |              |               |       |         |                   |                   |
   _cons   |   .2151604  |   .0159611    |  13.48|  0.000  |   .1838771        |   .2464436        |
q25_1      |              |               |       |         |                   |                   |
   _cons   |   .2612655  |   .0249856    |  10.46|  0.000  |   .2122946        |   .3102364        |
q75_0      |              |               |       |         |                   |                   |
   _cons   |   1.591147  |   .0725607    |  21.93|  0.000  |   1.44893         |   1.733363        |
q75_1      |              |               |       |         |                   |                   |
   _cons   |   2.510068  |   .1349917    |  18.59|  0.000  |   2.245489        |   2.774647        |
```
Using `mqgamma` and `poparms` for estimation and inference

**mqgamma example**

```stata
.nlcom (_b[q25_1:_cons] - _b[q25_0:_cons]) ///
> (_b[q75_1:_cons] - _b[q75_0:_cons])

_nl_1: _b[q25_1:_cons] - _b[q25_0:_cons]
_nl_2: _b[q75_1:_cons] - _b[q75_0:_cons]
```

|     | Coef.    | Std. Err. |     | P>|z| | [95% Conf. Interval] |
|-----|----------|-----------|-----|-----|----------------------|
| _nl_1 | 0.0461051 | 0.0295846 | 1.56 | 0.119 | -0.0118796 - 0.1040899 |
| _nl_2 | 0.9189214 | 0.1529012 | 6.01 | 0.000 | 0.6192405 - 1.218602  |
poparms is a user-written command documented in Cattaneo, Drukker, and Holland (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, Drukker, and Holland (2013) call the AIPW estimator an efficient-influence function (EIF) estimator because EIF theory is what produces the augmentation term
QTE can differ over $\tau$

- 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.
- \( \text{ATE} \) and \( \text{ATET} \) cannot investigate this type of hypothesis.
- QTE can investigate this type of hypothesis.
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 estimates

```
. clear all
. use cattaneo2  
. 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)

|     | bweight | bootstrap       | z   | P>|z<|      | [95% Conf. Interval] |
|-----|---------|-----------------|-----|---------|---------------------|
|     |         | Coef. Std. Err. |     |         |                     |
| 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 estimates

```
. poparms, coeflegend
Treatment Mean and Quantiles Estimation
(efficient influence function)
Number of obs = 4642

<table>
<thead>
<tr>
<th>bweight</th>
<th>Coef. Legend</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
</tr>
<tr>
<td>nonsmoker</td>
<td>3403.35 _b[mean:0bn.mbsmoke]</td>
</tr>
<tr>
<td>smoker</td>
<td>3183.081 _b[mean:1.mbsmoke]</td>
</tr>
<tr>
<td>q20</td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
</tr>
<tr>
<td>nonsmoker</td>
<td>3000 _b[q20:0bn.mbsmoke]</td>
</tr>
<tr>
<td>smoker</td>
<td>2778 _b[q20:1.mbsmoke]</td>
</tr>
<tr>
<td>q80</td>
<td></td>
</tr>
<tr>
<td>mbsmoke</td>
<td></td>
</tr>
<tr>
<td>nonsmoker</td>
<td>3840 _b[q80:0bn.mbsmoke]</td>
</tr>
<tr>
<td>smoker</td>
<td>3625 _b[q80:1.mbsmoke]</td>
</tr>
</tbody>
</table>
```
Using {mqgamma} and {poparms} for estimation and inference

**{poparms} example**

- **poparms estimates**

```stata
. 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 and Guido Imbens. 2012. “Matching on the estimated propensity score,”.


