

Conditional Average Treatment-Effects Estimation using Stata

Di Liu

Principal Econometrician

Stata

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary
- 7 Overview of the `cate` suite

ATE versus CATE

- ATE is a popular way to measure the treatment effects.

$$ATE = \mathbf{E}[y(1) - y(0)] \quad (1)$$

When each individual or group has **different (heterogeneous)** treatment effects, ATE may **oversimplify** the treatment effects.

- Conditional Average Treatment Effects (CATE) measure the treatment effects conditional on a set of variables.

$$CATE = \mathbf{E}[y(1) - y(0)|\mathbf{x}] \quad (2)$$

CATE measures the treatment effects as a **function** of **x**.

Advantages of studying CATE

- ① It improves the **understanding of the treatment-effect heterogeneity**.
 - ▶ Are the treatment effects heterogeneous?
 - ▶ How do the treatment effects vary with some variables?
 - ▶ Do the treatment effects vary between prespecified groups?
 - ▶ Do the data discover groups where treatment effects are different?
- ② It helps to **evaluate the treatment-assignment policy**.
 - ▶ If we implement a treatment-assignment policy, how would the average outcome in the population change?
 - ▶ Which policy is better among a candidate set of policies?

Different versions of CATE

$$CATE = \mathbf{E}[y(1) - y(0)|\mathbf{x}]$$

Depending on the definition of $\mathbf{x}$, CATE helps us to understand the heterogeneous treatment effects at different levels.

- **IATE**: Individualized average treatment effects when $\mathbf{x}$ is individual characteristics (finest level of treatment effects).
- **GATE**: Group average treatment effects when $\mathbf{x}$ is a group (prespecified group analysis).
- **GATES**: Sorted group average treatment effects when $\mathbf{x}$ ranks IATEs (data-driven group hypothesis testing).

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity**
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary
- 7 Overview of the `cate` suite

Partial linear outcome model

- We want to estimate the effect of 401(k) eligibility on net financial assets.

$$\mathbf{E}[\text{asset}(1) - \text{asset}(0)|\mathbf{x}]$$

where $\mathbf{x}$ are individual characteristics such as income, age, education, pension, marital status, etc.

- The outcome model is

$$\text{asset} = e401k * \tau(\mathbf{x}) + g(\mathbf{x}, \mathbf{w}) + \epsilon$$

where $\mathbf{w}$ is high-dimensional controls.

- So the potential outcomes are

$$\text{asset}(1) = \tau(\mathbf{x}) + g(\mathbf{x}, \mathbf{w}) + \epsilon$$

$$\text{asset}(0) = g(\mathbf{x}, \mathbf{w}) + \epsilon$$

- Thus,

$$\mathbf{E}[\text{asset}(1) - \text{asset}(0)|\mathbf{x}] = \tau(\mathbf{x})$$

Partialling-out estimator

$$\text{asset} = \text{e401k} * \tau(\mathbf{x}) + g(\mathbf{x}, \mathbf{w}) + \epsilon \quad (3)$$

$$\text{e401k} = f(\mathbf{x}, \mathbf{w}) + u \quad (4)$$

Taking conditional expectation in eq. (3) on both sides

$$\mathbf{E}[\text{asset} | \mathbf{x}, \mathbf{w}] = f(\mathbf{x}, \mathbf{w}) * \tau(\mathbf{x}) + g(\mathbf{x}, \mathbf{w}) \quad (5)$$

Eq. (3) minus (5) partialled-out $g(\mathbf{x}, \mathbf{w})$

$$\overbrace{\text{asset} - \mathbf{E}[\text{asset} | \mathbf{x}, \mathbf{w}]}^{\widetilde{\text{asset}}} = \overbrace{(\text{e401k} - f(\mathbf{x}, \mathbf{w}))}^{\widetilde{\text{e401k}}} * \tau(\mathbf{x}) + \epsilon \quad (6)$$

Estimate $\tau(\mathbf{x})$ by solving a local moment condition via generalized random forest.

$$\mathbf{E} \left[\alpha(\mathbf{x}) * \widetilde{\text{e401k}} * \left(\widetilde{\text{asset}} - \widetilde{\text{e401k}} * \tau(\mathbf{x}) \right) \right] = 0$$

Load data

```
. webuse assets3
(Excerpt from Chernozhukov and Hansen (2004))
. global catecovars age educ i.(incomecat pension married twoearn ira ownhome)
.
. global fvars incomecat pension married twoearn ira ownhome
. global controls c.(educ age)#i.($fvars)
```

- catecovars refers to **x**
- controls refers to **w**

Using cate to estimate IATE

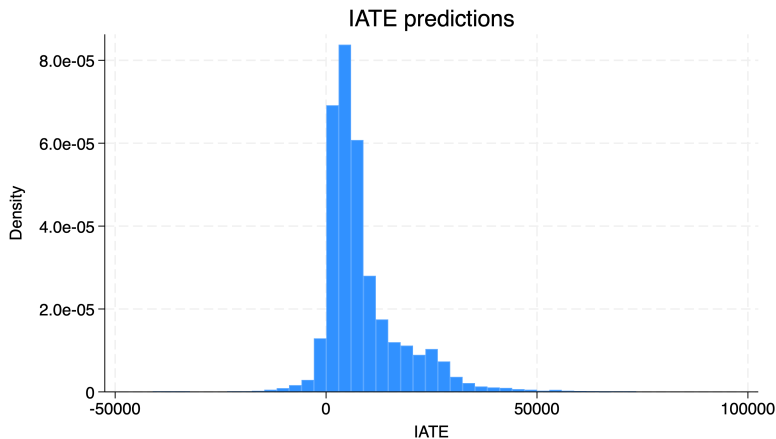
```
. cate po (assets $catecovars) (e401k), rseed(12345671) controls($controls) nol  
> og
```

Conditional average treatment effects	Number of observations	= 9,913
Estimator: Partialing out	Number of folds in cross-fit	= 10
Outcome model: Linear lasso	Number of outcome controls	= 47
Treatment model: Logit lasso	Number of treatment controls	= 47
CATE model: Random forest	Number of CATE variables	= 17

assets	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
ATE e401k (Eligible vs Not eligi..)	8107.563	1144.817	7.08	0.000	5863.763	10351.36
POmean e401k Not eligi..	13902.88	838.5924	16.58	0.000	12259.27	15546.49

The output shows ATE. Under the hood, `cate` also estimates a nonparametric function for IATE via generalized random forest.

categraph histogram: IATE predictions histogram



IATE distribution has a fat right tail, so the ATE possibly overestimates the treatment effects.

estat heterogeneity: test of treatment-effects heterogeneity

```
. estat heterogeneity  
Treatment-effects heterogeneity test  
H0: Treatment effects are homogeneous  
      chi2(1) =    4.19  
Prob > chi2 = 0.0406
```

We reject the null hypothesis of homogeneous treatment effects.

estat projection: linear projection of IATE

```
. estat projection
```

```
Treatment-effects linear projection
```

```
Number of obs =      9,913
F(11, 9901)    =        5.12
Prob > F       =       0.0000
R-squared      =       0.0047
Adj R-squared  =       0.0036
Root MSE      =      1.138e+05
```

	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
age	164.3654	116.2698	1.41	0.157	-63.54715	392.2779
educ	-440.1495	472.5372	-0.93	0.352	-1366.419	486.1197
incomecat						
1	-3093.247	1981.377	-1.56	0.119	-6977.15	790.6558
2	2216.006	2195.87	1.01	0.313	-2088.346	6520.359
3	6116.068	3244.506	1.89	0.059	-243.8253	12475.96
4	18355.28	5321.146	3.45	0.001	7924.749	28785.81
pension						
Receives ..	4320.983	2439.087	1.77	0.076	-460.1247	9102.09
married						
Married	-2103.475	3370.329	-0.62	0.533	-8710.007	4503.056
twoearn						
Yes	-1957.787	4326.422	-0.45	0.651	-10438.45	6522.88
ira						
Yes	-1284.949	3578.426	-0.36	0.720	-8299.392	5729.495
ownhome						
Yes	2963.537	1630.756	1.82	0.069	-233.0765	6160.15
_cons	1728.235	7880.15	0.22	0.826	-13718.46	17174.93

categraph iateplot (I): IATE function plot

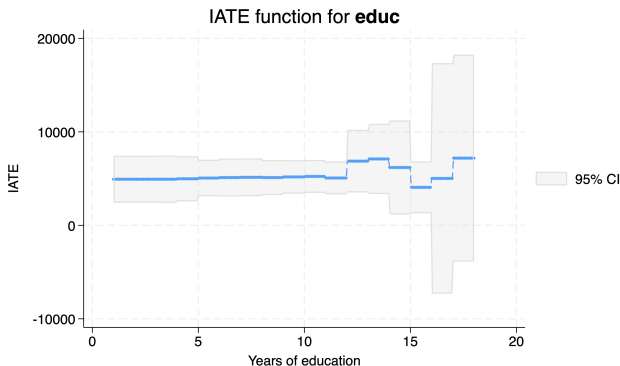
```
. categraph iateplot educ
```

Note: IATE estimated at fixed values of covariates other than **educ**.

Variable	Statistic	Value	Type
age	mean	41.05891	continuous
incomecat	base	0	factor
ira	base	0	factor
married	base	0	factor
ownhome	base	0	factor
pension	base	0	factor
twoearn	base	0	factor

Notice that $\tau(\mathbf{x})$ is a function of several parameters when $\dim(\mathbf{x}) > 1$. To plot a multiple dimension function, we **fix all** the variables to specific values **except** **educ**.

categraph iateplot (II)



Think about this graph as a slice in a bread in a specific direction. Each point is $\tau(\mathbf{x})$ when $\mathbf{x}$ takes a specific value. For example, $\mathbf{E}[y(1) - y(0) | \text{educ} = 10, \text{others} = \text{fixed}]$.

estat series: ATE over a continuous variable

```
. estat series educ, graph
warning: you have entered variable educ as continuous but it only has 18
        distinct values. The estimates may differ substantially if you
        inadvertently include a discrete variable as continuous

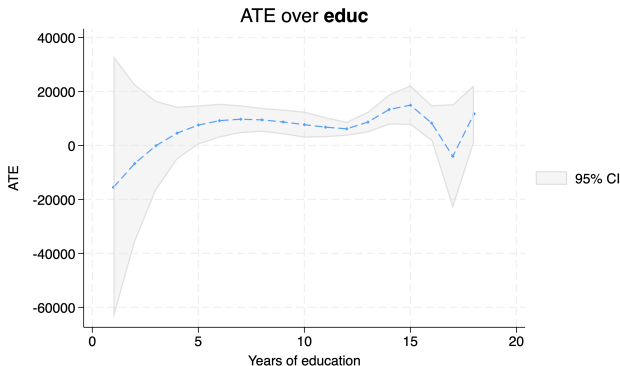
Computing approximating function
Minimizing cross-validation criterion
Iteration 0: Cross-validation criterion = 1.30e+10
Iteration 1: Cross-validation criterion = 1.30e+10
Computing average derivatives
Nonparametric series regression for IATE
Cubic B-spline estimation      Number of obs      =      9,913
Criterion: cross-validation    Number of knots    =      3
```

	Effect	Robust std. err.	z	P> z	[95% conf. interval]	
educ	2532.489	1377.915	1.84	0.066	-168.1735	5233.152

Note: Effect estimates are averages of derivatives.

The output shows the marginal effects of education on the treatment effects.

estat series: ATE over a continuous variable



Notice that each point shows the ATE if the education is fixed at a specific value. For example, $E[y(1) - y(0)|educ = 10]$.

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary
- 7 Overview of the `cate` suite

Using `cate ... , group(varname)` for GATE

We want to know the effects of e401k on asset for each income category.

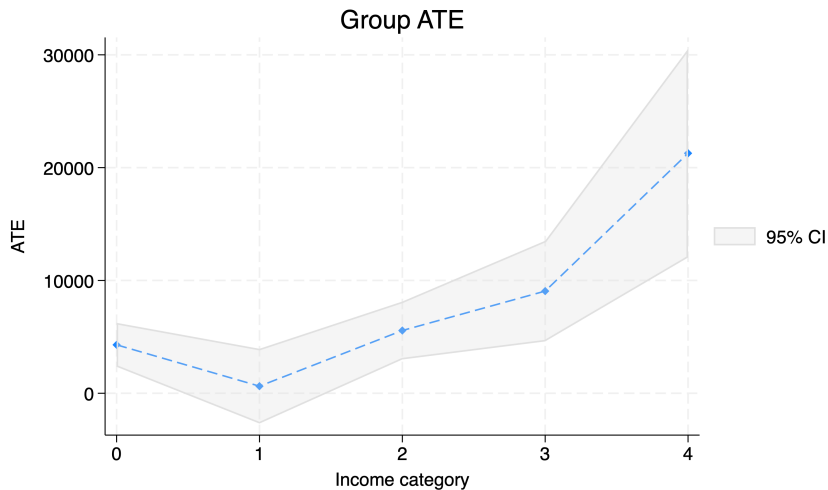
```
. cate aipw (assets $catecovars) (e401k), rseed(12345671) ///  
> controls($controls) group(incomecat) nolog
```

```
Conditional average treatment effects      Number of observations      = 9,913  
Estimator:      Augmented IPW              Number of folds in cross-fit = 10  
Outcome model:  Linear lasso                Number of outcome controls  = 47  
Treatment model: Logit lasso                Number of treatment controls = 47  
CATE model:     Random forest                Number of CATE variables    = 17
```

assets	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
GATE						
incomecat						
0	4295.829	992.7063	4.33	0.000	2350.16	6241.497
1	628.2236	1690.636	0.37	0.710	-2685.362	3941.809
2	5562.85	1310.006	4.25	0.000	2995.284	8130.415
3	9058.087	2276.042	3.98	0.000	4597.125	13519.05
4	21275.42	4716.757	4.51	0.000	12030.74	30520.09
ATE						
e401k (Eligible vs Not elig..)	8164.364	1151.125	7.09	0.000	5908.2	10420.53
POmean						
e401k Not eligi..	13910.87	842.0945	16.52	0.000	12260.39	15561.34

categraph gateplot: Visualize GATE

```
. categraph gateplot
```



estat gatetest: Test GATE homogeneity

```
. estat gatetest
```

Group treatment-effects heterogeneity test

H0: Group average treatment effects are homogeneous

(1) [GATE]0bn.incomecat - [GATE]1.incomecat = 0

(2) [GATE]0bn.incomecat - [GATE]2.incomecat = 0

(3) [GATE]0bn.incomecat - [GATE]3.incomecat = 0

(4) [GATE]0bn.incomecat - [GATE]4.incomecat = 0

chi2(4) = 22.39

Prob > chi2 = 0.0002

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)**
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary
- 7 Overview of the `cate` suite

Using cate ..., group(#) for GATES

```
. cate aipw (assets $catecovars) (e401k), rseed(12345671) ///  
> controls($controls) group(4) nolog
```

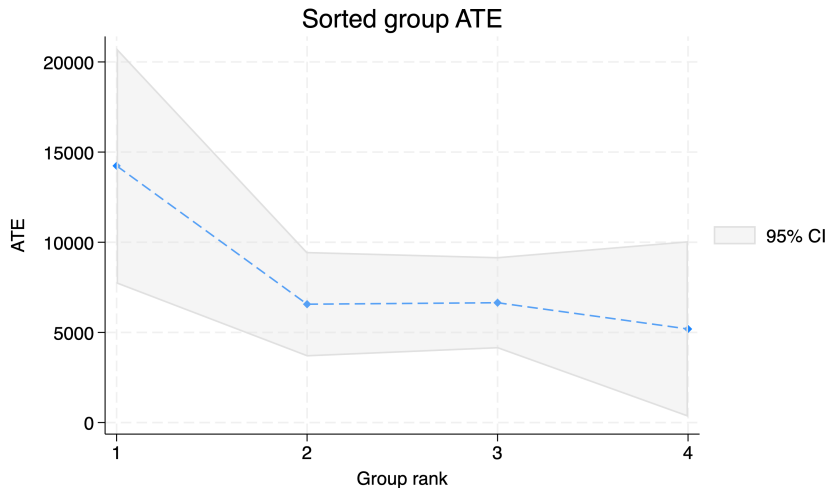
Conditional average treatment effects	Number of observations	= 9,913
Estimator: Augmented IPW	Number of folds in cross-fit	= 10
Outcome model: Linear lasso	Number of outcome controls	= 47
Treatment model: Logit lasso	Number of treatment controls	= 47
CATE model: Random forest	Number of CATE variables	= 17

assets	Robust					
	Coefficient	std. err.	z	P> z	[95% conf. interval]	
GATES						
rank						
1	14238.01	3335.108	4.27	0.000	7701.317	20774.7
2	6565.533	1482.069	4.43	0.000	3660.732	9470.334
3	6646.957	1294.802	5.13	0.000	4109.191	9184.723
4	5190.023	2487.992	2.09	0.037	313.6494	10066.4
ATE						
e401k (Eligible vs Not elig..)	8164.364	1151.125	7.09	0.000	5908.2	10420.53
POMean						
e401k Not eligi..	13910.87	842.0945	16.52	0.000	12260.39	15561.34

The group is defined by the IATE quantiles in a **cross-fitting** manner.
So higher rank **does not** necessarily imply higher ATE.

categraph gateplot: Visualize GATES

```
. categraph gateplot
```



In this example, group 1 has higher ATE than group 4. We can test it!

estat gatetest: Test GATES homogeneity

```
. estat gatetest 1 4
Sorted group treatment-effects heterogeneity test
H0: Sorted group average treatment effects are homogeneous
( 1)  [GATES]1bn.rank - [GATES]4.rank = 0
      chi2(1) =    4.73
Prob > chi2 = 0.0297
```

- The test rejects the null hypothesis of homogeneous GATE between groups 1 and 4. So people in group 1 have higher ATE than those in group 4.
- Question: Do the people in groups 1 and 4 have different characteristics?

estat classification: Classification analysis

Question: Is a variable's mean different between groups 1 and 4?

```
. estat classification ownhome
```

Classification t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf. interval]	
1	2,482	.8585818	.0069957	.3485227	.8448638	.8722998
4	2,469	.4641555	.0100387	.4988145	.4444703	.4838407
Combined	4,951	.6618865	.0067239	.4731152	.6487047	.6750683
diff		.3944263	.0122248		.3704603	.4183923

```
diff = mean(1) - mean(4)                                t = 32.2645
H0: diff = 0                                             Degrees of freedom = 4949
Ha: diff < 0                                           Ha: diff != 0
Pr(T < t) = 1.0000      Pr(|T| > |t|) = 0.0000      Ha: diff > 0
Pr(T > t) = 0.0000
```

```
. estat classification age
```

Classification t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf. interval]	
1	2,482	45.52337	.1785502	8.895315	45.17325	45.87349
4	2,469	37.2017	.2236204	11.11148	36.7632	37.6402
Combined	4,951	41.37346	.1547295	10.88729	41.07012	41.6768
diff		8.321667	.2859933		7.760993	8.882341

```
diff = mean(1) - mean(4)                                t = 29.0974
H0: diff = 0                                             Degrees of freedom = 4949
Ha: diff < 0                                           Ha: diff != 0
Pr(T < t) = 1.0000      Pr(|T| > |t|) = 0.0000      Ha: diff > 0
Pr(T > t) = 0.0000
```

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation**
- 6 Summary
- 7 Overview of the `cate` suite

Treatment-assignment policy

- Policy value:

$$\Pi(\pi) = \mathbf{E}[\pi_i y_i(1) + (1 - \pi_i) y_i(0)] \quad (7)$$

where $\pi_i \in [0, 1]$ is a prespecified treatment-assignment probability for the i th observation. π_i is also referred to as the policy.

- Notice that, from IATE estimates, we can already estimate $y(1)$ and $y(0)$. Thus, policy evaluation is closely related to CATE.
- Compare two policies:

$$\Pi(\pi_A) - \Pi(\pi_B) \quad (8)$$

ATE is a special case of policy comparison

Let $\pi_A = 1$ and $\pi_B = 0$. Then

$$\begin{aligned}ATE &= \mathbf{E}[y(1)] - \mathbf{E}[y(0)] \\&= \mathbf{E}[\textcolor{blue}{1} * y(1) + 0 * y(0)] - \mathbf{E}[\textcolor{red}{0} * y(1) + 1 * y(0)] \\&= \Pi(\pi_A) - \Pi(\pi_B)\end{aligned}$$

Thus, ATE is the contrast of the two special policy values. π_A means treat all the units, while π_B means treat none.

Lung transplant treatment-assignment policy evaluation

```
. webuse lung, clear
(Fictional data on lung transplant)

.
. global cvars bmip heightp o2amt lungals centervol walkdist      ///
>      bmid heightd distd lungpo2 hratio ischemict
. global fvars diabetesp karn racep sexp lifesvent assisvent      ///
>      o2rest raced smoked cmv deathcause diabetesd            ///
>      expandd sexd lungalloc genderm racem
.
. global controls c.($cvars)#i.($fvars)
.
. global catecovars c.($cvars) i.($fvars)
```

Treatment: Bilateral lung transplant vs. single lung transplant

Outcome: Forced expiratory volume in one second relative to a healthy person

Using cate to estimate IATE

```
. cate aipw (fevlp $catecovars) (transtype), rseed(12345671)    ///  
> controls($controls) nolog  
Conditional average treatment effects      Number of observations      = 937  
Estimator:      Augmented IPW              Number of folds in cross-fit = 10  
Outcome model:   Linear lasso               Number of outcome controls  = 454  
Treatment model: Logit lasso                Number of treatment controls = 454  
CATE model:      Random forest              Number of CATE variables    = 46
```

fevlp	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
ATE transtype (BLT vs SLT)	37.5243	.1646795	227.86	0.000	37.20153	37.84707
POmean transtype SLT	46.49502	.2025403	229.56	0.000	46.09805	46.892

Replicate ATE

```
. generate treatall = 1  
. generate treatnone = 0
```

```
. estat policyeval treatall treatnone
```

Treatment-assignment policy evaluation

Number of obs = 937

	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
Value						
policy						
treatall	84.01932	.3085432	272.31	0.000	83.41459	84.62406
treatnone	46.49502	.2025403	229.56	0.000	46.09805	46.892
Contrast						
policy						
(treatall						
vs						
treatnone)	37.5243	.1646795	227.86	0.000	37.20153	37.84707

Compare hypothetical policy with the observed policy

Hypothetical policy: Assigns patient to BLT if the patient's walking distance is greater than 500 meters in 6 mins and if the patient does not have diabetes.

```
. generate policy1 = walkdist > 500 & !diabetesp & !missing(walkdist)
.
. estat policyeval policy1 transtype
```

Treatment-assignment policy evaluation Number of obs = 937

	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
Value						
policy						
policy1	72.66426	.714435	101.71	0.000	71.26399	74.06452
transtype	66.53891	.5149955	129.20	0.000	65.52954	67.54828
Contrast						
policy						
(policy1						
vs						
transtype)	6.125348	.9130896	6.71	0.000	4.335725	7.91497

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary**
- 7 Overview of the `cate` suite

Methodological building blocks

- **Generalized random forest**: estimates the IATE function $\tau(\mathbf{x}) = \mathbf{E}[y(1) - y(0)|\mathbf{x}]$ (Athey et al., 2019)
- **Debiased/double machine learning**: partialling-out and AIPW estimators + cross-fitting (Athey et al. 2019, Semenova and Chernozhukov 2021, Nie and Wager 2021, Kennedy 2020, and Knaus 2022)

Benefits of modern methods:

- 1 Flexible IATE estimation **without assuming parametric assumptions**
- 2 **High-dimensional controls** in both the outcome and the treatment models
- 3 **Guard** against **machine learning bias**

Discussion

What can `cate` do ?

- Study treatment-effects heterogeneity at different levels (IATE, GATE, and GATES) in cross-sectional data
- Policy evaluation
- Nonparametric (GRF) or parametric (add linear interaction term) estimation of IATE
- Cross-fitting to guard against machine learning mistakes

Table of Contents

- 1 Motivation: Go beyond the ATE
- 2 Example 1: Exploit treatment-effects heterogeneity
- 3 Example 2: Group average treatment effect (GATE)
- 4 Example 3: Sorted group average treatment effect (GATES)
- 5 Example 4: Treatment-assignment policy evaluation
- 6 Summary
- 7 Overview of the `cate` suite

The cate suite (I)

- **Estimation:**

- ▶ `cate po` estimates **IATE** function (**partialling-out** model)
- ▶ `cate aipw` estimates **IATE** function (**AIPW** model)
- ▶ `cate ..., group(varname)` estimates **GATE**
- ▶ `cate ..., group(#)` estimates **GATES**

- **Prediction:** `predict` observational level IATEs, its standard error and CI

- **Visualization**

- ▶ `categraph histogram`: histogram of predictions of IATEs
- ▶ `categraph gateplot`: plot of GATE or GATES
- ▶ `categraph iateplot`: plot of the IATE function

The cate suite (II)

● Inference:

- ▶ `estat heterogeneity`: Heterogeneous treatment-effects test
- ▶ `estat gatetest`: GATE or GATES heterogeneity test
- ▶ `estat classification`: Classification analysis of the data-driven groups
- ▶ `estat ate`: ATE for a subpopulation
- ▶ `estat projection`: IATE function **linear approximation**
- ▶ `estat series`: IATE function **series approximation**
- ▶ `estat policyeval`: Treatment-assignment policy evaluation

References

- Athey, S., J. Tibshirani, and S. Wager. 2019. Generalized random forests. *Annals of Statistics* 47: 1179–1203.
- Kennedy, E. H. 2020. Towards optimal doubly robust estimation of heterogeneous causal effects . URL <http://arxiv.org/abs/2004.14497>.
- Knaus, M. C. 2022. Double machine learning-based programme evaluation under unconfoundedness. *Econometrics Journal* 25: 602–627.
- Nie, X., and S. Wager. 2021. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika* 108: 299–319.
- Semenova, V., and V. Chernozhukov. 2021. Debiased machine learning of conditional average treatment effects and other causal functions. *Econometrics Journal* 24: 264–289.