Machine Learning
for Economists

Optimal Policy Learning

Anthony Strittmatter

Literature

- Athey and Wager (2018): "Efficient Policy Learning", download.
- Kitagawa and Tetenov (2018): "Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice", Econometrica, 86(2), pp. 591-616, download.

What are Policy Rules?

- Determine the allocation of treatments to individuals based on observable covariates
- Policy rules are often labelled assignment rules, individualized treatment rules (ITR), personalized treatment rules, etc.

$\Rightarrow$ Optimal policy rules can potentially improve the allocation of limited resources

Scope of Applications

- Assignment of unemployed to training programs
- Targeting get-out-the-vote campaigns
- Allocation of preventive medical counselling (e.g. midwife or nutritionist)
- Telemedicine: Should the patient be send to the pharmacist/doctor or directly to the hospital?
- Targeting marketing campaigns:
  - Marketing for sustainable and renewable electricity rates by the energy provider
  - Marketing for supplementary insurances (e.g., dental insurance) by the health care provider
  - Marketing of a charitable organisation

Solicitation Letters

She's known nothing but abject poverty her entire life. Why on earth should Sebastiana have hope now? After forty-two years of toil in the unforgiving land of the high Andes, Sebastiana looks much older than her years. She has borne nine children and is alone to care for them after losing her husband six years ago. But a few months ago, Sebastiana joined a women's group sponsored by Freedom from Hunger. There she received a loan of $64 and training on how to grow her small, home-based business.

Treatment Definition

Treatment (efficiency story):
But does she really have a right to hope for something different? According to studies on our programs in Peru that used rigorous scientific methodologies, women who have received both loans and business education saw their profits grow, even when compared to women just received loans for their business. But the real difference comes when times are slow. The study showed that women in Freedom from Hunger's Credit with Education program kept their profits strong-ensuring that their families would not suffer, but thrive.

Control (emotional story):
But does she really have a right to hope for something different? Like Sophia and Carmen before her, the good news is, yes! Because of caring people like you, Freedom from Hunger was able to offer Sebastiana a self-help path toward achieving her dream of getting "a little land to farm" and pass down to her children. As Sebastiana's young son, Aurelio, runs up to hug her, she says, "I do whatever I can for my children."

Potential Effects

How to increase fundraising for Freedom for Hunger's Credit with Education program?

- **Altruistic donation motive:**
  Altruistic donors gain utility through the increased social welfare generated by the donation
- **Alternative donation motives:**
  Warm-glow, social norms, social pressure, own benefit, casually, participation (e.g., DellaVigna, List, and Malmendier, 2012)
- Mixtures between different motives possible. I will not distinguish sharply between different donation motives.

Fundamental Identification Problem

Potential Donations:

- Treatment dummy:

$$D_i = \begin{cases} 1 & \text{efficiency story} \\ -1 & \text{emotional story} \end{cases}$$

- $Y_i(1)$ potential donation under efficiency story
- $Y_i(-1)$ potential donation under emotional story

Observed Donations:

In the absence of spillover effects,

$$Y_i = Y_i(-1) + \frac{1+D_i}{2}\left(Y_i(1) - Y_i(-1)\right)$$

Individual Causal Effect:

$$\delta_i = Y_i(1) - Y_i(-1)$$

Infeasible Approach

- **Optimal Policy Rule:** $\pi_i^* = 1\{\delta_i > 0\} - 1\{\delta_i \leq 0\}$
  - Efficient story if $\delta_i > 0$
  - Emotional story if $\delta_i \leq 0$
- It is in most applications infeasible to identify and estimate individual causal effects

Approach Based on CATEs

- **Conditional Average Treatment Effect (CATE)**

$$\delta(x) = E[\delta_i | X_i = x] = E[Y_i(1) - Y_i(-1) | X_i = x]$$

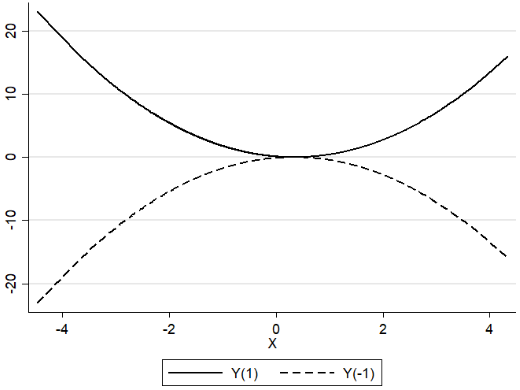
- X_i contains exogenous pre-treatment covariates/features/attributes that are potentially responsible for effect heterogeneity

- **CATE based policy rule:** $\pi(X_i) = 1\{\delta(X_i) > 0\} - 1\{\delta(X_i) \leq 0\}$
  - Efficient story if $\delta(X_i) > 0$
  - Emotional story if $\delta(X_i) \leq 0$

- **Caveat:**
  - The selection of a policy rule is a classification problem
  - CATEs are not targeted at this classification problem

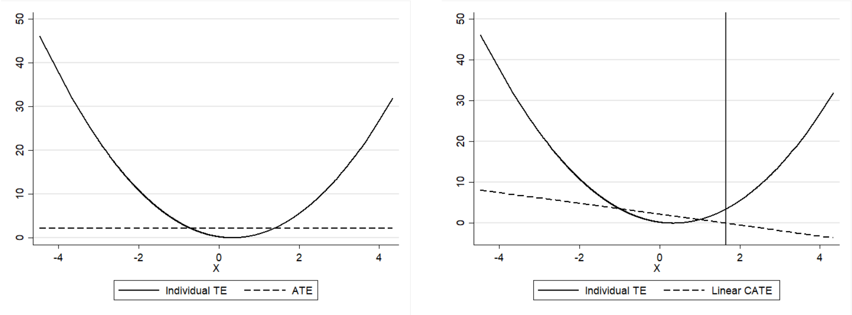
Simple Example

- $X \sim N(0,1)$
- $Y(1) = (X - 1/3)^2$
- $Y(-1) = -(X - 1/3)^2$



Reference: Qian and Murphy (2011)

CATEs Not Suited for Policy Rules



- Treating everybody is optimal
- ATEs find optimal policy rule ($MSE_{ATE} \approx 9.4$), even though linear prediction of CATEs approximate the individual treatment effects better ($MSE_{ATE} > MSE_{CATE} \approx 7.8$)

Regret Function

- For a given set of covariates X_i, the best achievable policy rule maximizes the utility of the policy

$$\pi^{best} = \max_{\pi} E[Y_i(\pi(X_i))]$$

- This is equivalent to selecting the π that minimizes the regret function

$$R(\pi) = E[Y_i(\pi_i^*)] - E[Y_i(\pi(X_i))]$$

$\rightarrow$ minimax regret criterion (Manski, 2004)

- The regret is the gap between the optimal and selected policy
  - **Type I regret:** mistakenly choosing an inferior treatment
  - **Type II regret:** mistakenly rejecting a superior treatment

Weighted Classification Representation

- Estimate the policy $\hat{\pi}(X_i)$ that maximizes

$$\hat{\pi} = \arg\max_{\pi} \frac{1}{2} E\left[\pi(X_i) \cdot sign(\delta_i) \cdot |\delta_i|\right]$$

- Classification of $sign(\delta_i)$ with weights $|\delta_i|$
- **Intuitively:**
  - Misclassifications hurts more when the (absolute) treatment effects are large
  - Misclassifications of individuals with (almost) zero effects is not very costly

Estimator

Apply sample analogy principle

$$\hat{\pi} = \arg\max_{\pi} \frac{1}{2 \cdot N} \sum_{i=1}^{N} \hat{\pi}(X_i) sign(\hat{\Gamma}_i) |\hat{\Gamma}_i|$$

with $\hat{\Gamma}_i$ being an approximation score of δ_i

- Augmented Inverse Probability Weighting:
  (Athey and Wager, 2017)

$$\Gamma_i = \mu_1(X_i) - \mu_{-1}(X_i) + \frac{D_i(Y_i - \mu_1(X_i))}{p(X_i)} - \frac{(1 - D_i)(Y_i - \mu_{-1}(X_i))}{1 - p(X_i)}$$

  with nuisance parameters $\mu_d(x) = E[Y_i(d)|X_i = x]$ and $p(x) = Pr(D_i = 1|X_i = x)$

- Weighted classification problem

Policy Learning Algorithm

1. Split the data in two samples A and B
2. Use ML to estimate $\hat{\mu}^A_{+1}(X_i)$, $\hat{\mu}^A_{-1}(X_i)$, and $\hat{p}^A_{+1}(X_i)$ in Sample A; as well as $\hat{\mu}^B_{+1}(X_i)$, $\hat{\mu}^B_{-1}(X_i)$, and $\hat{p}^B_{+1}(X_i)$ in Sample A
3. Estimate your preferred score function $\hat{\Gamma}_i$, for example,

$$\hat{\Gamma}^A_i = \hat{\mu}^B_{+1}(X_i) - \hat{\mu}^B_{-1}(X_i) + W_i \frac{Y_i - \hat{\mu}^B_{W_i}(X_i)}{\hat{p}^B_{W_i}(X_i)}$$

$$\hat{\Gamma}^B_i = \hat{\mu}^A_{+1}(X_i) - \hat{\mu}^A_{-1}(X_i) + W_i \frac{Y_i - \hat{\mu}^A_{W_i}(X_i)}{\hat{p}^A_{W_i}(X_i)}$$

4. Use ML to classify $sign(\hat{\Gamma}^A_i)$ with weight $|\hat{\Gamma}^A_i|$ in order to obtain the probability $\hat{q}^A_i(X_i) = Pr(\hat{\pi}^A(X_i) = 1)$. Proceed equivalently in sample B and obtain $\hat{q}^B_i(X_i)$.
5. Implement the policy rule $\pi(\hat{X}_i) = 2 \cdot 1\{\hat{q}^A_i(X_i) + \hat{q}^B_i(X_i) > 1\} - 1$

Classification Methods

- **Classification Trees**
  - In contrast to regression trees, classification trees use different performance measures
  - These measures are targeted to minimise the impurity (instead of the regression fit)
  - Entropy or Gini index
- **Logistic LASSO**
- **Support Vector Machines** (partition data in two samples)

Athey and Wager (2018)

(Main) Regularity Conditions:

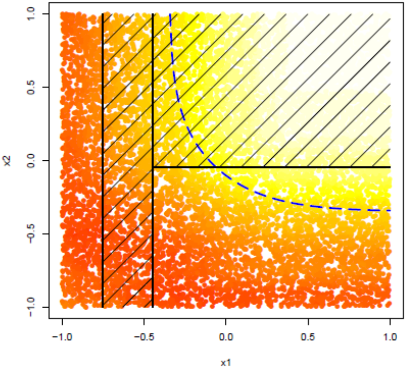
- Uniform consistency of $\hat{\mu}_w(X_i)$ and $\hat{p}_w(X_i)$
- $\sqrt[4]{N}$-convergence of $\hat{\mu}_w(X_i)$ and $\hat{p}_w(X_i)$
- $VC(\Pi)$ needs to have bounded complexity

Further Results:

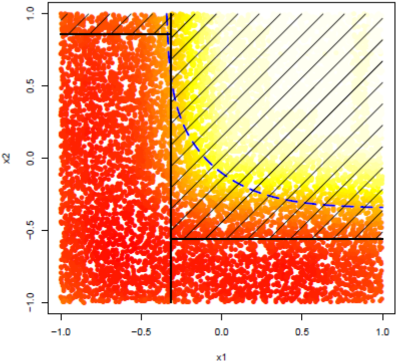
- $\hat{Q}_{DML}(\pi)$ is semi-parametrically efficient
- $\sqrt{N}(\hat{Q}_{DML}(\pi) - Q(\pi)) \xrightarrow{d} N(0, V(\pi))$
- $V(\pi) = Var(\pi(X_i)\delta(X_i)) + E\left[\frac{Var(Y(-1)|X_i)}{\hat{p}_{-1}(X_i)} + \frac{Var(Y(+1)|X_i)}{\hat{p}_{+1}(X_i)}\right]$

Simulation Exercise

Inverse Probability Weighting



Double Machine Learning



Data

- In June 2007, Freedom from Hunger sends 11,207 mails to previous donors who donated at least one gift since 2004
- Data obtained from Karlan and Wood (2017)
- I drop 214 individuals who donated more than \$200 in their last gift before the mailer (mean = \$43, max = \$10,000)

- **Randomized Control Trial:**
  - Stratified randomization by most recent donation year and previous donation amount (above/below \$100)
  - 5,524 previous donors received a mailer with the personal and efficiency story (treatment group)
  - 5,521 previous donors received a mailer with the personal and emotional story (control group)

Variable Definitions

Outcomes:

- Donation amount in response to the mailer (in US-dollars)
- Dummy indicating a donation in response to the mailer

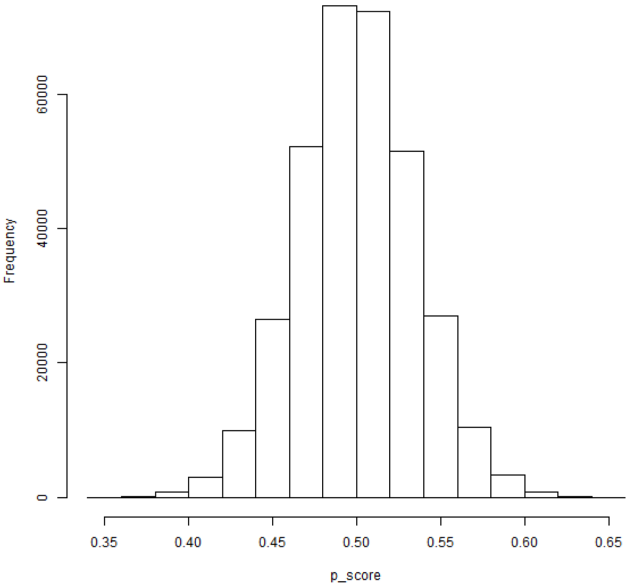
Covariates:

- Amount given before mailer, amount of last gift before mailer, largest gift before mailer (in US-dollars)
- Date first donation, date last donation before mailer (in days)
- # of gifts given in year prior to experiment, total # of gifts given, # of gifts per year
- Received mailer with big ask donor in past
- Median zip code income (in US-dollars)
- Average years of education in census tract

Descriptive Statistics

	Treated		Control		Std.
	mean	std. dev.	mean	std. dev.	Diff.
	(1)	(2)	(3)	(4)	(5)
char_giving	7.135	7.135	7.053	7.053	1.149
ext_marg	0.159	0.159	0.163	0.163	-2.418
amount_pre	236.5	236.5	229.0	229.0	3.205
amount_lastpre	33.70	33.70	34.21	34.21	-1.521
amount_maxpre	48.22	48.22	46.42	46.42	3.794
date_last	16914	16914	16915	16915	-0.007
date_first	15086	15086	15111	15111	-0.165
ngifts_yearbefore	0.736	0.736	0.773	0.773	-4.879
ngifts	6.753	6.753	6.719	6.719	0.508
ngifts_frequency	0.823	0.823	0.838	0.838	-1.836
bigask	0.096	0.096	0.096	0.096	-0.244
medinc	51849	51849	51522	51522	0.633
years_ed	13.98	13.98	13.97	13.97	0.083

Histogram Propensity Score

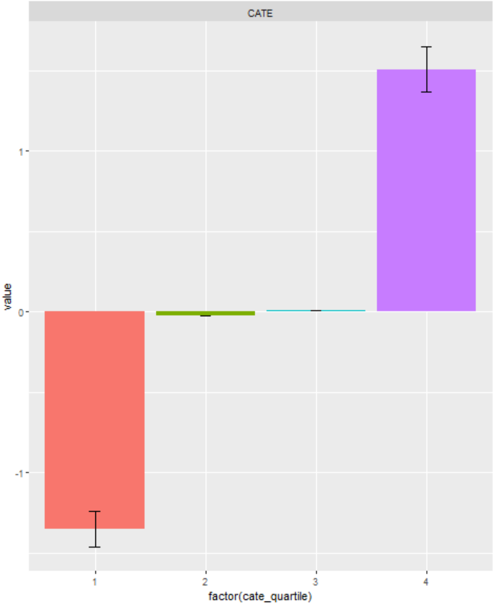


Average Treatment Effects

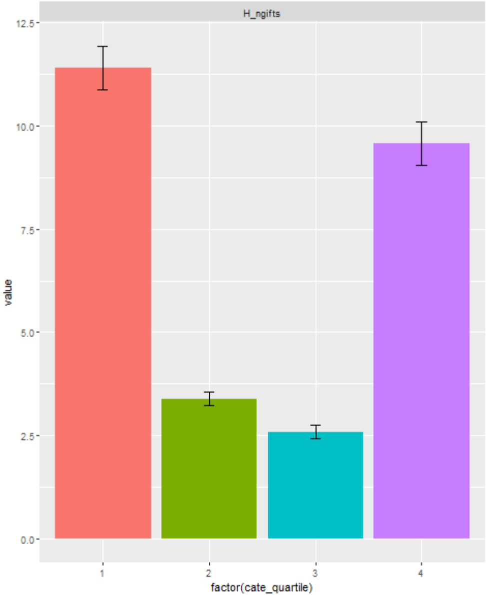
OLS	DML
(1)	(2)
Donation Amount (in US-dollars)	
0.0815	-0.012757
(0.5831)	(0.388609)
Donation Probability	
-0.003890	-0.0003409
(0.006993)	(0.00070150)

- Total donations \$78,353 in response to the mailer
- 1,777 donors (16%) give charity in response to the mailer

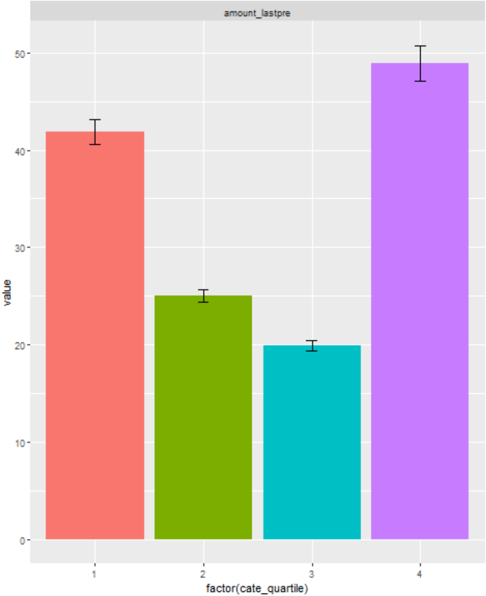
Quartiles of Predicted CATEs



Number of Previous Gifts by CATE Quartiles



Amount Last Gift by CATE Quartiles



In-Sample Results

CATEs	Policy Rule			
Rule	depth = 3	depth = 9	depth = 27	opt. depth
(1)	(2)	(3)	(4)	(5)
		Donation Amount (in US-dollars)		
2.339***	1.026***	1.796***	2.140***	2.668***
(0.184)	(0.171)	(0.181)	(0.183)	(0.190)
		Donation Probability		
0.013***	0.002***	0.005***	0.007***	0.010***
(0.0003)	(0.0001)	(0.0002)	(0.0003)	(0.0003)

avg. opt. depth $\approx$ 24

Opt. depth increases donations by $29,472 (38%) or donors by 106 persons (6%)

Share Treated

	CATEs	Policy Rule			
		depth = 3	depth = 9	depth = 27	opt. depth
	(1)	(2)	(3)	(4)	(5)
	Donation Amount (in US-dollars)				
Treated	48.9%	52.2%	52.0%	52.4%	52.2%
	5396	5764	5743	5789	5763
	Donation Probability				
Treated	49.5%	46.1%	48.4%	51.0%	50.9%
	5469	5088	5351	5631	5622

Out-of-Sample Results

CATEs Rule	Policy Rule depth = 3	depth = 9	depth = 27	opt. depth
(1)	(2)	(3)	(4)	(5)
Donation Amount (in US-dollars)				
-0.002	0.118	0.056	0.037	0.101*
(0.071)	(0.076)	(0.059)	(0.058)	(0.058)
Donation Probability				
-0.00009	-0.00010	0.00013	0.00024*	0.00025*
(0.00013)	(0.00009)	(0.00010)	(0.00014)	(0.00015)

Opt. depth increases donations by $1,114 (1.4%) or donors by 3 persons (0.2%)

Preliminary Conclusions

- Optimal policy rules can potentially increase charitable giving, but only to a limited extend in this application
- Out-of-sample all policy rules dominate the CATE rule
- What could improve the results?
  - More observations (11,000 observations)
  - More **relevant** covariates (11 covariates)
  - Stronger treatment

$\Rightarrow$ **Work in progress!**

Budget constraints

- Substract cost (e.g., Kitagawa and Tetenov, 2018):

$$\hat{\Gamma}_i^{Budget} = \hat{\Gamma}_i - c_i$$

- Fix number of participants (e.g., Bhattacharya and Dupas, 2012):

$$\hat{\pi}_i^{Budget} = 2 \cdot 1\{\hat{\pi}(X_i) \geq \bar{\pi}\} - 1$$

- Combination of both enables to fix the number of participants when the cost of participation vary

Batch vs. Bandit Algorithms

Batch

- Historical dataset
- Potentially optimal policy rules change over time
- Then findings cannot be extrapolated to the future

Bandit

- Data arrives sequentially (typically online data)
- Treatment decisions are made sequentially
- Presence of exploration vs. exploitation trade-off
- Example: targeted online advertisements
- Reference: Dimakopoulou, Zhou, Athey, and Imbens (2018)

Further Extensions

- **Multiple treatments**
  (e.g., Frölich, 2008, Kallus, 2017, Zhou, Athey, Wager, 2018)
- **Ordered treatments**
  (e.g., Chen, Fu, He, Kosorok, and Liu, 2018)
- **Dynamic treatments**
  (e.g., Zhang and Zhang, 2018, Zhao, Zheng, Laber, and Kosorok, 2015)
- **Continuous treatments**
  (e.g., Chen, Zheng, and Kosorok, 2016, Athey and Wager, 2018)

Ethical Concerns?

- Statistical discrimination even if we omit critical variable (e.g., gender, migration, etc.)
- Examples: hiring decisions, flight prices, program assignments
- More or less than discrimination than humans?
- Targeting rules also have the potential to reduce discrimination, but it has to be used appropriately
- Current scandals: Cambridge Analytica, Amazons' unethical hiring algorithm