

Практическая эконометрика.

Лекция 8. Оценка гетерогенных эффектов: причинный лес и meta-learners

преподаватели 2023: Ольга Сучкова, Алексей Замниус, Анна Ставнийчук При поддержке Георгия Калашнова (автор курса и части слайдов). Используем материалы Дмитрия Архангельского и Мдины Курмангалиевой.

2 ноября 2023

Гетерогенность эффектов

Эффект может быть гетерогенным (Cunningham, п.04)

$$\begin{aligned} \mathbb{E}[Y_1|T=1] - \mathbb{E}[Y_0|T=0] = & \\ \underbrace{\mathbb{E}[Y_1] - \mathbb{E}[Y_0]}_{\text{ATE}} + & \underbrace{\mathbb{E}[Y_0|T=1] - \mathbb{E}[Y_0|T=0]}_{\text{Sample Bias}} + \\ \underbrace{(1-\pi)(ATT - ATnT)}_{\text{HTE bias}} & \end{aligned}$$

Пример гетерогенности - программа переселения из ветхого жилья в Чикаго: Chyn (2018)

- Различия в судьбе детей из переселённых семей и оставшихся жить в неблагополучном районе.



Эффекты для разных групп: Chyn (2018)

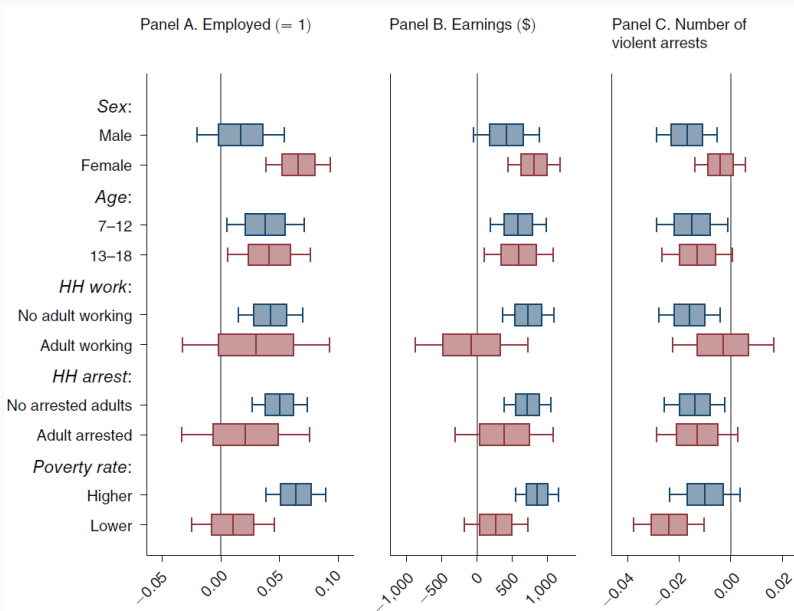


Table of Contents

Гетерогенность эффектов

Оценка гетерогенности как таковой

Что такое гетерогенность еще раз

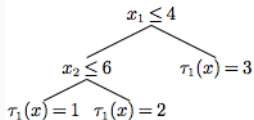
$$HTE(= CATE) = \tau(x) = \mathbb{E}(\tau|X = x)$$

Зачем мы можем хотеть оценить гетерогенность

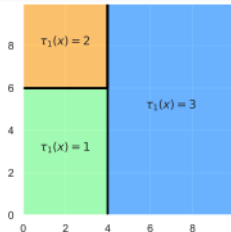
1. Чтобы установить ее наличие или отсутствие
2. Чтобы понять масштаб гетерогенности
3. Частичная идентификация
4. для рекомендаций (policy advice)
 - 4.1 Оценка кредитоспособности
 - 4.2 Таргетирование грантов
 - 4.3 Таргетирование маркетинга

На доске схемы!

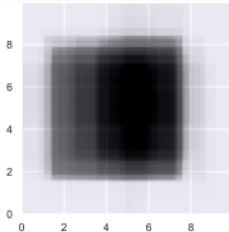
Random Forests (Breimer, 2001)



(a) A depth-3 tree. When the condition in a branching node holds, we take the left branch.



(b) Each tree gives a partition of \mathbb{R}^d , where each region corresponds to a leaf of the tree.



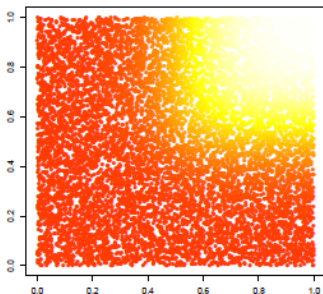
(c) Darker regions fall into the same region as $x = (0,0)$ for more trees in a forest.

Generalized (causal) random forests (Wager, Athey, 2017)

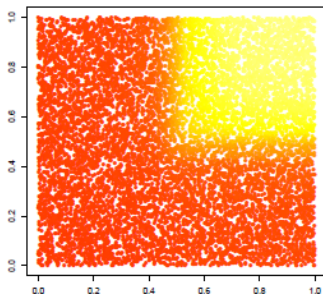
Схема на доске!

Что делать с полученными CATE? (Wager, Athey, 2017)

Heatmap: цветом - размер эффекта, по осям - иксы



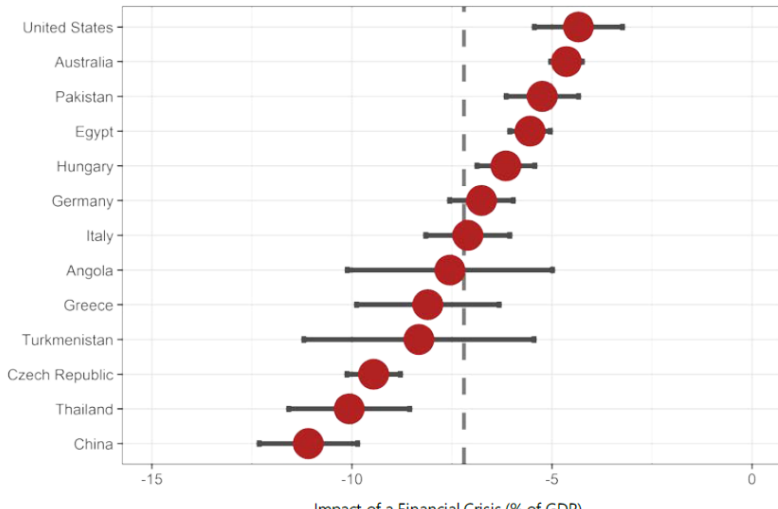
True effect $\tau(x)$



Causal forest

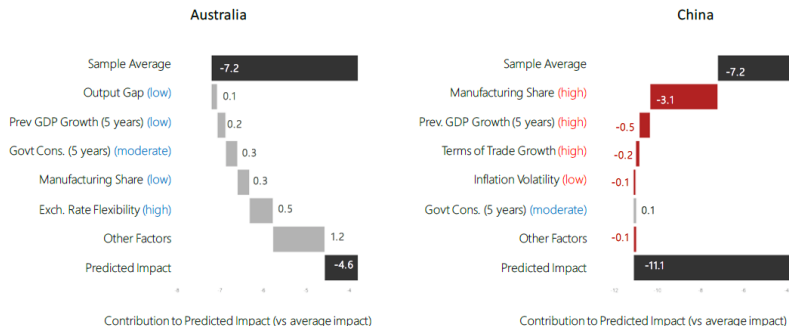
Пример: влияние банковских кризисов на рост ВВП (МВФ, 2019)

Figure 1. Causal Forest Predictions (Select Countries)



Вклад факторов в эффект (МВФ, 2019)

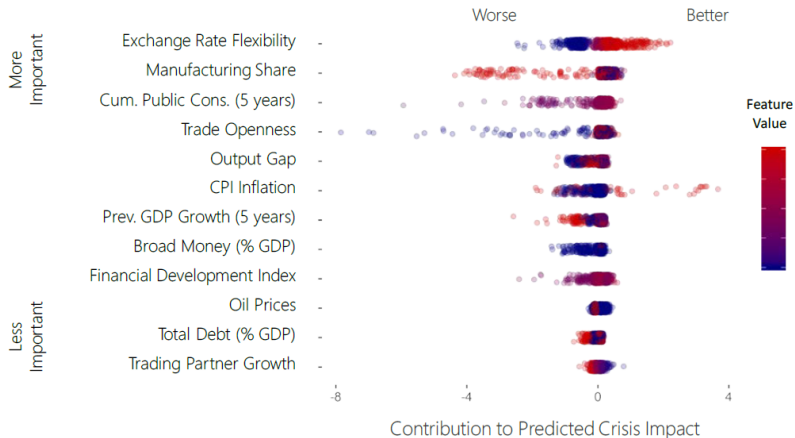
Figure 2. Shapley Values: Impact of a Banking Crisis on Output Growth



Source: IMF VE Database, author's calculations

Вклад факторов в эффект (МВФ, 2019)

Figure 3. Distribution of Shapley Values



Source: IMF VE Database, author's calculations

Meta-learners (Kunzel et al., 2017)

- ▶ T-learner (2 regressions)
- ▶ S-learner (single regression)
- ▶ X- learner
- ▶ PW-learner (propensity weights)
- ▶ DR-learner

Как соотносятся? какие преимущества? Проблемы? Как различаются их MSE ?

Meta-learners: cxema (Curth et al., 2021)

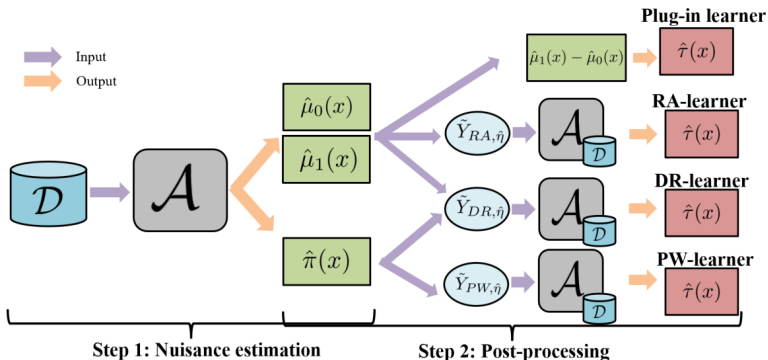
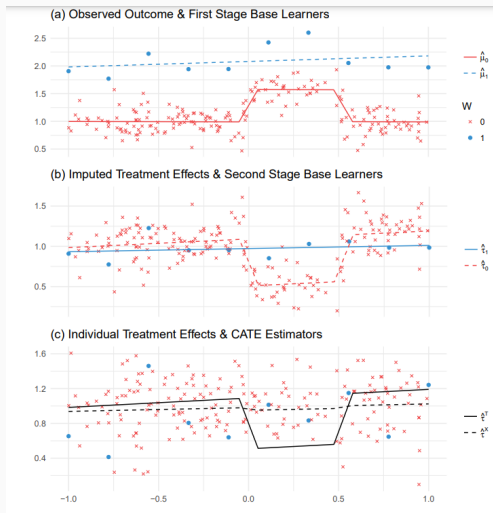


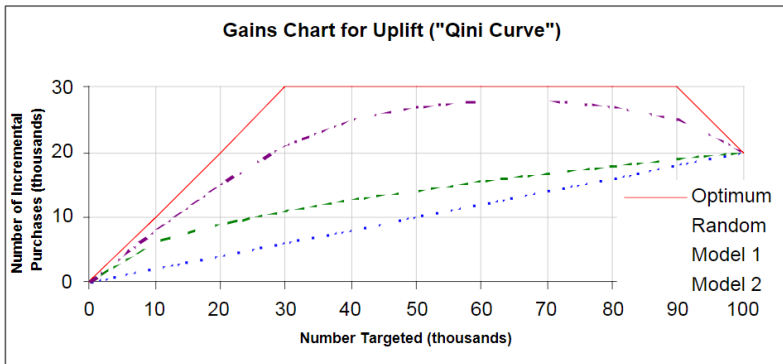
Figure 1: High-level overview of the four meta-learning

T- and X-learners with an unbalanced design (Kunzel et al., 2017)



Оценка качества моделей на исторических данных (No ML seminar, Окунева, 2022)

- ▶ uplift на $k\%$ клиентов
- ▶ График uplift по бинам
- ▶ uplift curve Radcliff, 2017



Оценка качества моделей на новых данных (No ML seminar, Окунева, 2022)

- ▶ response модели по сравнению с простой рандомизацией (эффект модели)
- ▶ response для 4 типов клиентов в модели и в простой рандомизации (эффект программы в целом)

Что почитать и посмотреть?

Причинный случайный лес:

- ▶ Статья (Wager, Athey, 2017)
- ▶ Лекция 4 Дмитрий Архангельский 2018
- ▶ Материалы лекций Мадины Курмангалиевой, Université Libre de Bruxelles 2021

Meta-learners

- ▶ Академическая базовая статья о meta-learners (Kunzel et al., 2017)
- ▶ Учебное пособие (Curth et al., 2021)
- ▶ Статья с примерами uplift curve (Radcliff, 2017)
- ▶ Видео отраслевика: (No ML seminar, П. Окунева, Glowbyte, 2022)