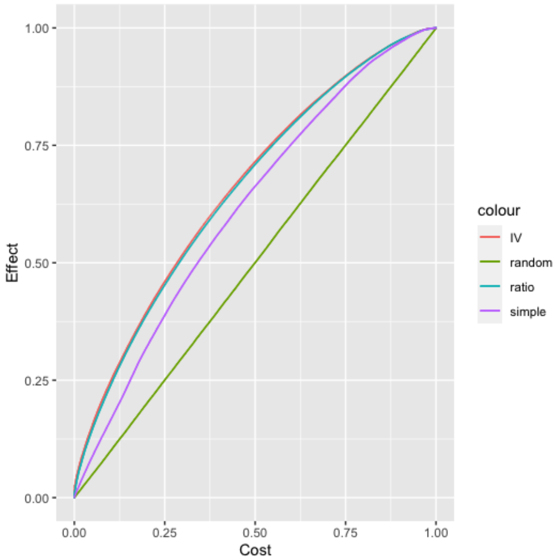
Marginal costs and policies based on them

Георгий Калашнов

Московский государственный университет, экономический факультет

29 октября 2020 г.

Final result



Plan

Table of Contents

Example 1: Development economics

Some literature in development economics focus on investigation of treatment regimes to optimize the cost effectiveness of intervention:

- Optimize the target sample (Crépon et al., 2015)
- Optimize the treatment itself (Cohen et al., 2015)

Most of the research provides average costs per treatment point estimate

Meta research (Give Well) focus on:

- external validity and proper account of spillover effects and side costs
- Budgets are planned on invest/don't invest basis

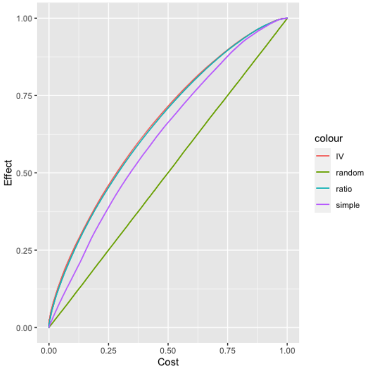
Example 1: Development economics

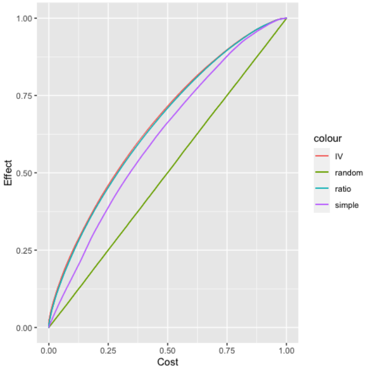
There is no or little discussion of

- Saturation and scalability
- Use of cunning methods to estimate marginal costs

These are important for optimization and budget planning

Example 1: Development economics





Example 2: Public finance

Governmental benefits (e.g. https://www.usa.gov/benefits)

- Are often targeted to specific groups of population
- Aim to introduce equitable income distribution
- Solve the poverty traps
- Can have other secondary goals

The problem of finding the best is essentially the same as in Development economics

Example 3: Price discrimination

- The firm can perfectly discriminate between workers
- Two possible price (or wage) levels
- Workers have heterogeneous and ex-ante unknown productivity and propensity to accept the wage

Table of Contents

Literature on Policy Learning

The literature proposes to

- Treat the costs as either constant or known in advance (Athey, Wager, 2017; Bhattacharya, Dupas, 2012)
- Optimize a linear combination of cost and effect (Zhao, Harinen, 2019)
- Solve the problem for a single budget constraint (Kallus, Mao, 2020; Kitagawa, Tetenov, 2018)

My research differs in the following sense

- Computes all possible budget-effect pairs
- Yield rules, which are monotonic in budget constraint

This may be useful to:

- Allocate budget between two initiatives
- Decrease time of computing
- Consider off-the-equilibrium-path outcomes
- Increase the budget at the implementation phase

The nearest work to mine is Goldenberg et al. (2020)

Among other ideas propose to use knapsack algorithm for the problem

$$\text{Sort the subjects according to } \frac{\hat{\tau}_o(X)}{\hat{\tau}_c(X)}$$

However, they

- Do not discuss limits of applicability
- Miss the analysis of the case, when treatment effect on the costs can be negative
- Do not evaluate the results with the effect-costs Pareto border

Literature on optimization methods

- Athey et al. (2019) solve local moment conditions with Random Forests and give confidence intervals

$$E((Y-\alpha-\tau T)T|X)=0$$

Other papers which I find useful, but do not use in my research now:

- Chernozhukov et al. (2018) argue that non-parametric regression methods require super-smoothness properties of data to be consistent and propose a generic solution and inference method for any regression method
- Kallus, Mao (2020) solves the stochastic optimization problem with random forests

Table of Contents

Method 1: Knapsack algorithm

- Estimate the two conditional treatment effects models $\hat{\tau}_c$, $\hat{\tau}_o$ (e.g. using generalized random forests, Athey et al., 2019)
- Swap the labels for the treatment and control groups if $\hat{\tau}_c(X) < 0$
- Estimate the inverse of the marginal costs as $\hat{\tau}_o(X)/\hat{\tau}_c(X)$
- Rank the items according to this inverse in ascending order and treat them while we meet budget constraint

Method 2: Empirical Welfare Maximization

$$W(\pi) = \frac{(O - \lambda C)(1 - T)}{1 - e(x)} + \pi(X)\tau(X) \to \max_{\pi \in \Pi}$$

$$\tau(O, C, T, \lambda) = \frac{(O - \lambda C)T}{e(x)} - \frac{(O - \lambda C)(1 - T)}{1 - e(x)}$$

- The sign of the τ determines optimal treatment
- I find λ, which solves the following using a generalized random forest (Athey et al., 2019)

$$\mathbb{E}(\tau(O, C, T, \lambda)|X) = 0$$

This is equivalent to a local instrumental estimate of the effect of O on C using T as an instrument.

Pareto Border

1. Split the sample to the main and evaluation sample, stratifying on the treatment
2. Fit the model on a main subsample
3. Build a cumulative effect and cost on evaluation subsample (and probably take a mean of several bootstrap estimates to smoothen it)

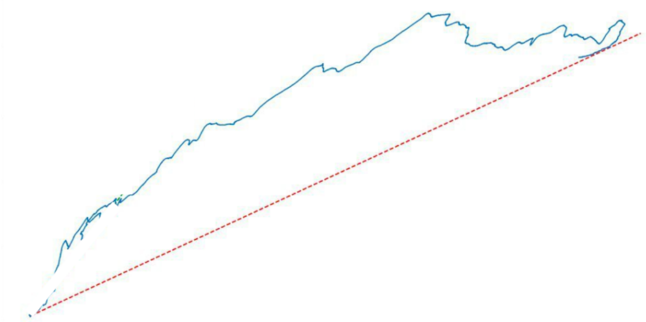


Table of Contents

Methods Comparisons

- Kitagawa, Tetenov (2018) show that the EWM method reaches asymptotic minimax regret efficiency bound though not all of their assumptions may hold here.
- EWM method requires $\tau_c(X)$ to be non-negative which is a ubiquitous case (monotonicity property)
- We have confidence intervals for $\hat{\lambda}(x)$ for EWM method
- The knapsack algorithm method requires to use e.g. delta method, but it can lack sample size locally.

I show these properties of the methods in simulations and compare them with naive approaches.

Table of Contents

Simulation Parameters

$$O(0) = C(0) = 0$$
$$C(1) = 2 + X_1 + \varepsilon$$
$$O(1) = \lambda C(1)$$
$$X_1 \sim N(0, 25)$$
$$X_2 \sim N(0, 1)$$
$$\varepsilon \sim N(0, 4)$$

I take absolute values of $O(1)$ and $C(1)$ to enforce τ_c to be non-negative to analyse the case when monotonicity property is satisfied

Simulation: Monotonicity assumption

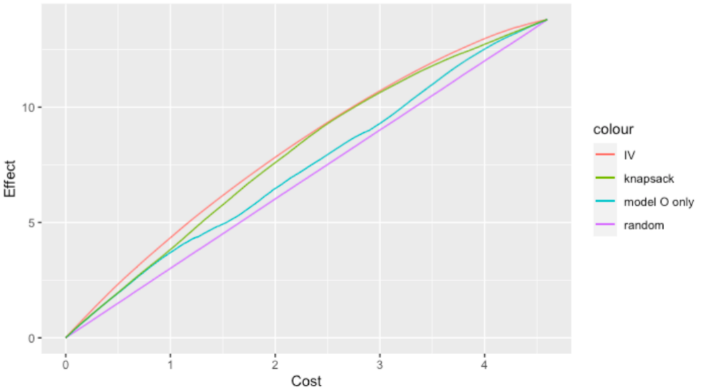


Рис.: Monotonicity in data

Simulation: No Monotonicity assumption

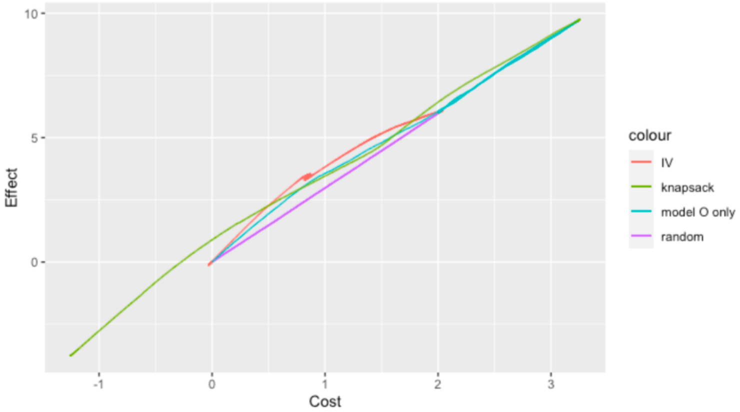
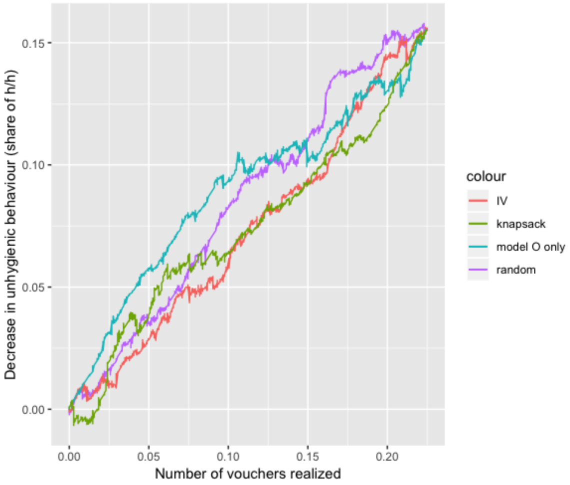


Рис.: No monotonicity in data

Table of Contents

Sanitation vouchers in Africa (Guiteras et al., 2015)



Malaria treatment subsidies in Bangladesh (Cohen et al., 2015)

Athey Susan, Tibshirani Julie, Wager Stefan. Generalized random forests // Annals of Statistics. 4 2019. 47, 2. 1179–1203.

Athey Susan, Wager Stefan. Efficient Policy Learning. 2017.

Bhattacharya Debopam, Dupas Pascaline. Inferring welfare maximizing treatment assignment under budget constraints // Journal of Econometrics. 3 2012. 167, 1. 168–196.

Chernozhukov Victor, Demirer Mert, Duflo Esther, Fernández-Val Iván. Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments. Cambridge, MA, 6 2018.

Cohen Jessica, Dupas Pascaline, Schaner Simone. Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial // American Economic Review. 2 2015. 105, 2. 609–645.

Crépon Bruno, Devoto Florencia, Duflo Esther, Parienté William. Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco // American Economic Journal: Applied Economics. 2015. 7, 1. 123–150.

Goldenberg Dmitri, Albert Javier, Bernardi Lucas, Estevez Pablo. Free Lunch! Retrospective Uplift Modeling for Dynamic Promotions Recommendation within ROI Constraints // RecSys 2020 - 14th ACM Conference on Recommender Systems. New York, NY, USA: Association for Computing Machinery, Inc, 9 2020. 486–491.

Guiteras Raymond, Levinsohn James, Mobarak Ahmed Mushfiq. Encouraging sanitation investment in the developing world: A cluster-randomized trial // Science. 5 2015. 348, 6237. 903–906.

Stochastic Optimization Forests. // . 8 2020.

Kitagawa Toru, Tetenov Aleksey. Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice // Econometrica. 3 2018. 86, 2. 591–616.

Zhao Zhenyu, Harinen Totte. Uplift Modeling for Multiple Treatments with Cost Optimization // Proceedings - 2019 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2019. 8 2019. 422–431.