Research on Design and Analysis of Experiments

- **Surrogates** (Athey, Chetty, Imbens, Kang, 2016, update coming shortly)
- **Heterogeneous treatment effects** (Wager & Athey (JASA, 2018); Athey, Tibshirani, and Wager (AOS, 2019); Friedberg, Athey, Tibshirani, and Wager (2018)
- **Offline policy estimation** (Athey and Wager, 2017; Zhou, Athey, and Wager 2018)
- **Improving estimation model used in contextual bandit algorithms** (Dimakopoulou, Zhou, Athey and Imbens, AAAI, 2018)
- **Survey** Athey & Imbens, The Econometrics of Randomized Experiments (Handbook of Experimental Economics)

TODAY:

- **Designing experiments with staggered rollouts** (Xiong, Athey, Bayati, Imbens 2019)
- **Testing hypotheses using adaptively collected data** (Hadad, Hirschberg, Zhan, Wager, Athey, 2019)

Recent work on causal inference with panel data

- **Matrix factorization in Consumer Choice settings**
  - Donnelly, Ruiz, Blei, and Athey (2019), Ruiz, Athey and Blei (AOAS, 2019), Athey, Blei, Donnelly, Ruiz, and Schmidt (AEA P&P, 2018)
  - Augment traditional mixed/nested logit models with matrix factorization for random coefficients
  - Tune for causal effects, test with held-out experiments

- **Matrix Completion for Panel Data Models**
  - Athey, Bayati, Doudchenko, Imbens, Khosravi, 2017
  - When evaluating staggered rollouts, use matrix completion to construct counterfactual outcomes

- **Synthetic Difference-in-Differences**
  - Arkhangelsky, Athey, Hirschberg, Imbens, Wager, 2018
  - Builds on ideas of double robustness in causal inference
  - Combine outcome models (e.g. matrix completion) with weighting that focuses on similar units and time periods

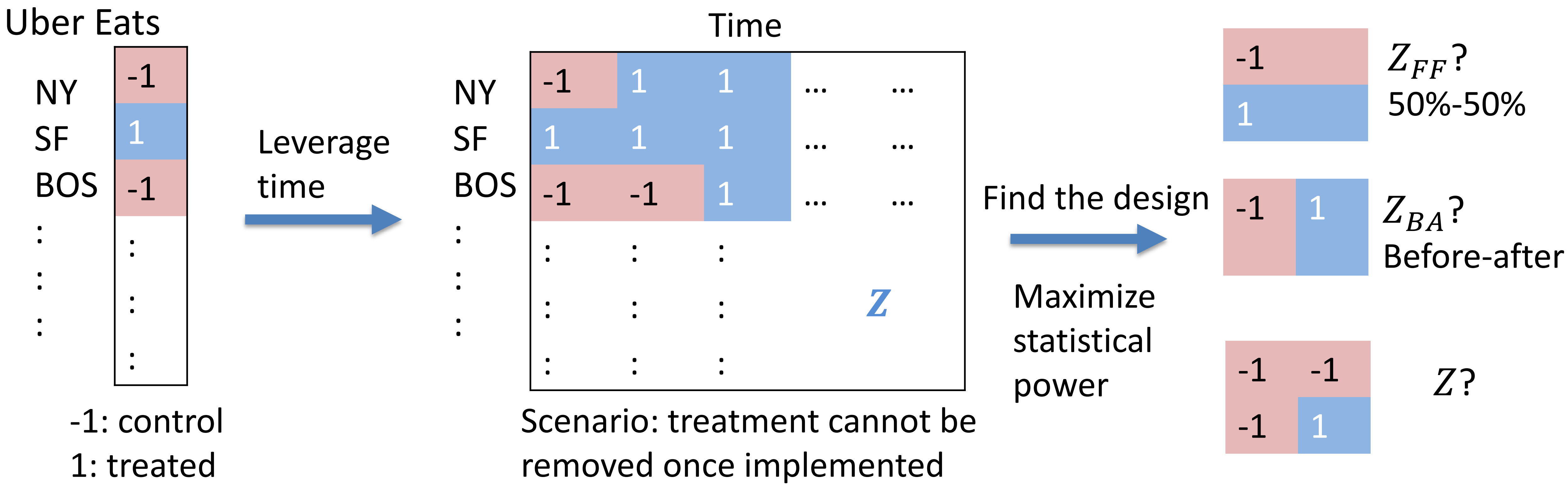
- **This paper:** *Design* experiments anticipating future hypothesis testing
  - Take the matrix completion approach to outcome modeling for tractability

Optimal Experimental Design for Staggered Rollouts

Ruoxuan Xiong, Susan Athey, Mohsen Bayati, Guido Imbens
Stanford University

Statistical Power: A Fundamental Challenge

- Randomized controlled **experiments** are the gold standard to draw inference about whether a product or an intervention is **effective**.
- **Fundamental challenge**: Low power due to small sample size:
  - Avoid contamination and network effects
  - Relevant units are market, product category, school, hospital…



Related Literature

- Related work: optimal allocation of clusters for a multi-period design
  - Hussey and Hughes 2007, Hemming et al. 2015, Li et al. 2018

- Contribution of this paper
  1. Study the optimal multi-period design under a general class of linear models
  2. Allow to use historical control data to find a better design

Problem Formulation: Optimization

- Model, for all $(i, t) \in \{1, \ldots, N\} \times \{1, \ldots, T\}$

$$\boldsymbol{Y_{it}} = \alpha_i + \beta_t + u_i^T v_t + \boldsymbol{\tau z_{it}} + \varepsilon_{it}$$

- Optimization problem: given $\hat{\tau}$, solve $\boldsymbol{Z} = [\boldsymbol{z_{it}}]$

$$\min_{z_{it}} \quad Var(\hat{\tau})$$ Maximize statistical power

$$s.t. \quad z_{it} \leq z_{it+1}$$ Treatment cannot be removed once implemented

$$z_{it} \in \{\text{-}1, 1\}$$ -1: control, 1: treated

- Integer programming

- Estimation: Best Linear Unbiased Estimator (BLUE) to estimate $\boldsymbol{\tau}$
  - Explicit formula for $Var(\hat{\tau})$ in terms of $\boldsymbol{Z}$

Observed : Y_{it}, z_{it}
Unknown : $\alpha_i, \beta_t, \boldsymbol{\tau}, \varepsilon_{it}, u_i, v_t$
α_i, β_t : unit, time fixed effects
u_i : latent covariates
z_{it}: treatment status
$\boldsymbol{\tau}$: treatment effect

Time β_t

α_i

NY	-1	1	1	…	…
SF	1	1	1	…	…
BOS	-1	-1	1	…	…
:	:	:	:		
:	:	:	:	$\boldsymbol{Z}$	
:	:	:	:		

Main Result

- N units, T periods, and $Y_{it} = \alpha_i + \beta_t + u_i^T v_t + \tau z_{it} + \varepsilon_{it}$ with $z_{it} \in \{-1,1\}$

Theorem. If u_i takes finitely many values, then any design that has **staggered rollouts** with $\frac{2t-1}{2T}$ 100% treated for each **strata** at time t is an **optimal** solution to $\min\limits_{\substack{[z_{it}]\in[-1,1]^{N\times T} \\ z_{it}\leq z_{it+1}}} Var(\hat{\tau})$

- Some intuition why?

-1	-1
-1	-1
1	1
1	1

Only time effect β_t
Z_{FF} : 50%-50%

-1	1
-1	1
-1	1
-1	1

Only unit effect α_i
Z_{BA} :Before-after

-1	-1
-1	1
-1	1
1	1

Both effects: $\frac{2t-1}{2T}$

Main Result

- N units, T periods, and $Y_{it} = \alpha_i + \beta_t + u_i^T v_t + \tau z_{it} + \varepsilon_{it}$ with $z_{it} \in \{-1,1\}$

Theorem. If u_i takes finitely many values, then any design that has **staggered rollouts** with $\frac{2t-1}{2T} \times 100\%$ treated for each **strata** at time t is an **optimal** solution to $\min_{\substack{[z_{it}] \in [-1,1]^{N\times T} \\ z_{it} \leq z_{it+1}}} Var(\hat{\tau})$

Strata: units with the same covariate values

$$\begin{array}{rl} u_i\text{=1} & \left\{\begin{array}{cc} -1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & 1 \end{array}\right. \\ u_i\text{=-1} & \left\{\begin{array}{cc} -1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & 1 \end{array}\right. \end{array}$$

Uber Eats

u_i: quality of restaurants

u_i=1: good

u_i=-1: bad

Main Result

- N units, T periods, and $Y_{it} = \alpha_i + \beta_t + u_i^T v_t + \tau z_{it} + \varepsilon_{it}$ with $z_{it} \in \{-1,1\}$

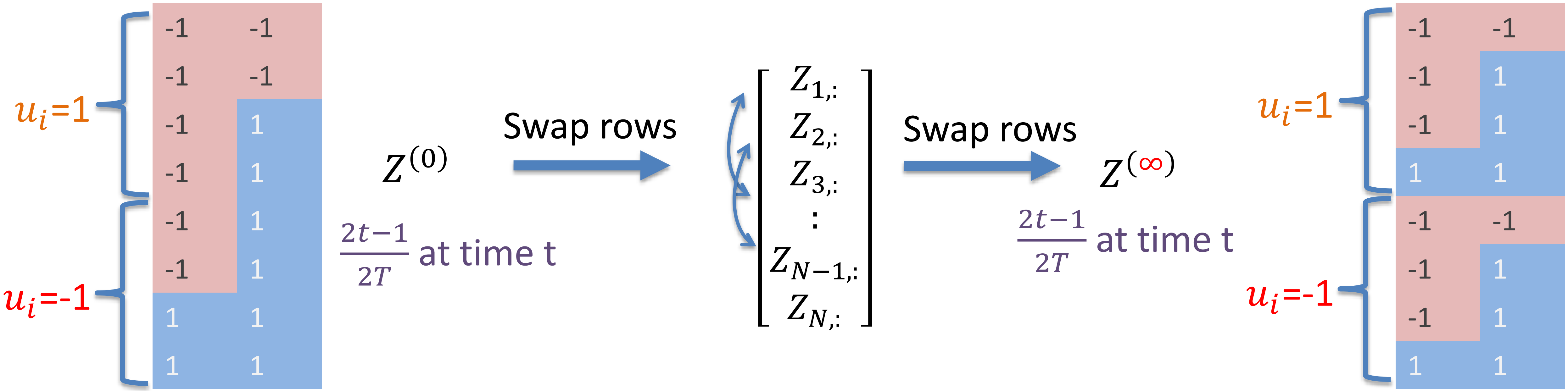
Theorem. If u_i takes finitely many values, then any design that has **staggered rollouts** with $\frac{2t-1}{2T} \times 100\%$ treated for each **strata** at time t is an **optimal** solution to $\min_{\substack{[z_{it}]\in\{-1,1\}^{N\times T} \\ z_{it} \leq z_{it+1}}} Var(\hat{\tau})$

Remark 1: This result holds regardless of whether u_i is known or latent

Remark 2: For general u_i use a clustering algorithm (e.g., K-means) to stratify

Next: A data-driven heuristic to stratify, when u_i's are latent

A Local Search Heuristic to Stratify



Step 1. Preprocess historical data: split historical data into several matrices

Step 2. Minimax decision rule: synthetic experiment τ on each matrix with estimator $\hat{\tau}^{(j)}$

- Find the Z to minimize $\max_j |\hat{\tau}^{(j)} - \tau|$. **(Robust measure)**

Step 3. Local Search Z (simulated annealing)

Empirical Application

- Purchases from a large grocery store
  - 17,880,248 transactions between May 2005 and May 2007
  - Aggregated weekly net expenditure by household
    - 63,375 households over 97 weeks
    - 85.4% entries to be 0
- Hypothetical experiments:
  - Promotions, coupons, ads, etc.
  - …

Empirical Application: Comparison

- Design matrices

$$\begin{matrix} -1 & -1 \\ -1 & -1 \\ 1 & 1 \\ 1 & 1 \end{matrix}$$

Z_{FF}

50% control 50% treated at every time period

$$\begin{matrix} -1 & 1 \\ -1 & 1 \\ -1 & 1 \\ -1 & 1 \end{matrix}$$

Z_{BA}

First half all control second half all treated

$$\begin{matrix} -1 & -1 \\ -1 & -1 \\ -1 & 1 \\ -1 & 1 \end{matrix}$$

Z_{FF+BA}

First half all control second half half treated

$$\begin{matrix} -1 & -1 \\ -1 & 1 \\ -1 & 1 \\ 1 & 1 \end{matrix}$$

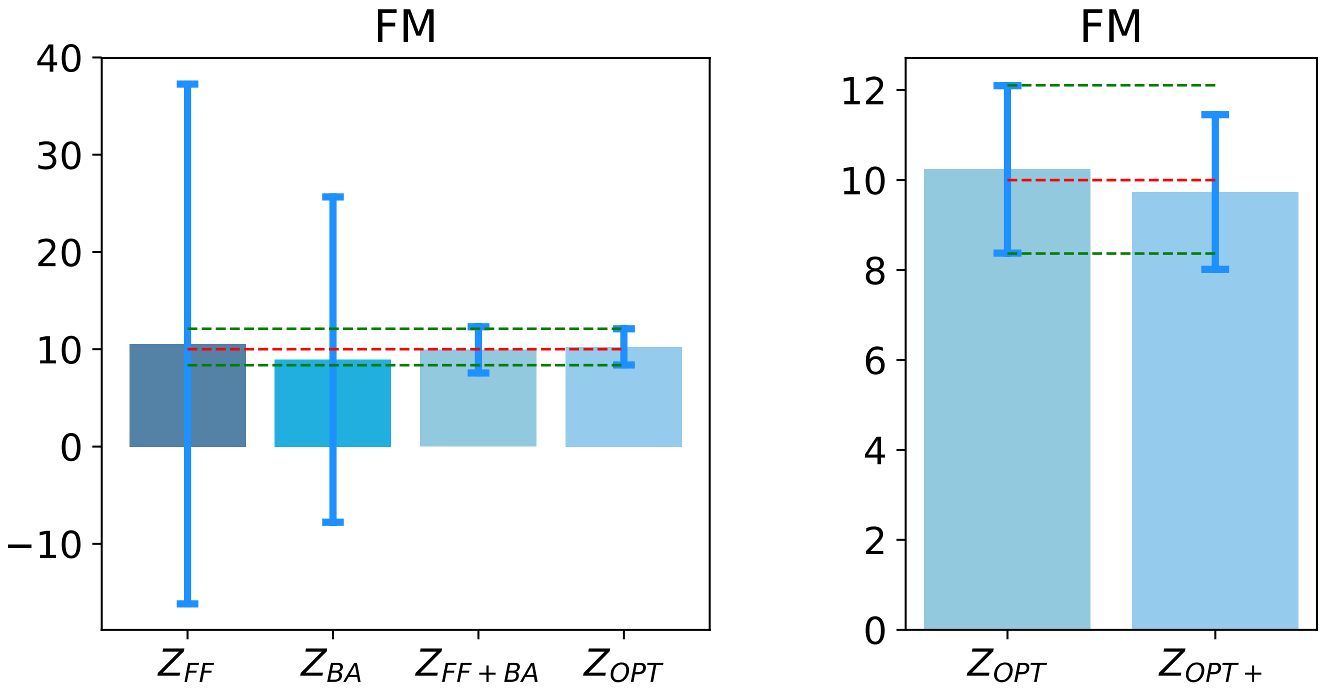
Z_{OPT}

Staggered rollouts

% treated: $\frac{2t-1}{2T}$

Results

- Z_{OPT} outperforms benchmark matrices $Z_{FF}, Z_{BA}, Z_{FF+BA}$ for all estimation methods
- Factor model has RMSE about half of linear fixed effect model
- Using historical data, local search heuristic with minimax decision rule provides a better matrix Z_{OPT+} than Z_{OPT} for all estimation methods



RMSE: 26.73 16.76 2.37 **1.88** 1.88 **1.74**

Red dash line: true value
Bar height: average estimated value
Error bar: variance

Conclusion

- **Optimal** multi-period design
  - Linear treated portion in time with **stratification** (randomize the treatment allocation within each strata, the units with the same covariates' value)

- A data driven data-driven heuristic to stratify
  - Based on local search with minimax decision rule using historical control data

- **Extension:** Treatment effect varies over time
  - Optimal design starts with lower % treated, increases faster, in order to learn time-dependent treatment effects

- Empirics: synthetic experiments on several real world data sets
  - Various domains, including retail and health care
  - Superior performance of the analytical solution and our algorithm compared with the benchmark designs

Confidence Intervals for Policy Evaluation in Adaptive Experiments

Vitor Hadad David A. Hirshberg Ruohan Zhan
Stefan Wager Susan Athey

Graduate School of Business
Stanford University
vitorh@stanford.edu

November, 2019

The promise

Multi-armed bandit experiments can be a substantially more efficient optimization method than traditional experiments. [1]

[1] Scott, S. L. (2015). Multi-armed bandit experiments in the online service economy. Applied Stochastic Models in Business and Industry, 31(1), 37-45.

Adaptive Experimentation

The promise

Multi-armed bandit experiments can be a substantially more efficient optimization method than traditional experiments. [1]

What do we want to *optimize*?

- Standard for bandits: *Low regret* during the experiment

But for innovation, understanding, scientific discovery, also want:

- Statistical *accuracy*
- Valid *confidence intervals*
- *Power* for the hypothesis test

[1] Scott, S. L. (2015). Multi-armed bandit experiments in the online service economy. Applied Stochastic Models in Business and Industry, 31(1), 37-45.

Adaptive Experimentation

The promise

Multi-armed bandit experiments can be a substantially more efficient optimization method than traditional experiments. [1]

What do we want to *optimize*?

- Standard for bandits: *Low regret* during the experiment

But for innovation, understanding, scientific discovery, also want:

- Statistical *accuracy*
- Valid *confidence intervals*
- *Power* for the hypothesis test

These goals can be *at odds* with each other.

[1] Scott, S. L. (2015). Multi-armed bandit experiments in the online service economy. Applied Stochastic Models in Business and Industry, 31(1), 37-45.

Inference challenges

You are given data $\{(W_t, Y_t)\}_{t=1}^{T}$ coming from an adaptive experiment. How would you estimate the value of arm w?

[1] [Xu et al., 2013, Bowden and Trippa, 2017, Nie et al., 2018, Shin et al., 2019]

Inference challenges

You are given data $\{(W_t, Y_t)\}_{t=1}^{T}$ coming from an adaptive experiment. How would you estimate the value of arm w?

Sample average

$$\widehat{Q}^{Sample}(w) = \frac{1}{n_w} \sum_{t:W_t=w} Y_t$$

[1] [Xu et al., 2013, Bowden and Trippa, 2017, Nie et al., 2018, Shin et al., 2019]

Inference challenges

You are given data $\{(W_t, Y_t)\}_{t=1}^{T}$ coming from an adaptive experiment. How would you estimate the value of arm w?

Sample average

$$\widehat{Q}^{Sample}(w) = \frac{1}{n_w} \sum_{t: W_t = w} Y_t$$

- In a randomized control trial: *unbiased*, *consistent*
- In a sequential experiment: possibly *biased* and *inconsistent* [1]

[1] [Xu et al., 2013, Bowden and Trippa, 2017, Nie et al., 2018, Shin et al., 2019]

Inference challenges

An *unbiased* estimator

Inverse Propensity Weighted Estimator

$$\widehat{Q}^{IPW}(w) = \frac{1}{T}\sum_{t=1}^{T}\frac{\mathbb{I}\{W_t = w\}}{e_t(w)}Y_t$$

- **Notation:** $e_t(w)$ is probability of selecting arm w at t

Inference challenges

An *unbiased* estimator

Inverse Propensity Weighted Estimator

$$\widehat{Q}^{IPW}(w) = \frac{1}{T}\sum_{t=1}^{T}\frac{\mathbb{I}\left\{W_t = w\right\}}{e_t(w)}Y_t$$

- **Notation:** $e_t(w)$ is probability of selecting arm w at t
- **Intuition:**
  - Adaptive data collection introduces sample bias
  - Up/down weight observations to counter bias

Inference challenges

An *unbiased* estimator

Inverse Propensity Weighted Estimator

$$\widehat{Q}^{IPW}(w) = \frac{1}{T}\sum_{t=1}^{T}\frac{\mathbb{I}\{W_t = w\}}{e_t(w)}Y_t$$

- **Notation:** $e_t(w)$ is probability of selecting arm w at t
- **Intuition:**
  - Adaptive data collection introduces sample bias
  - Up/down weight observations to counter bias
- **Problem:** High variance if $e_t(w)$ small

Inference challenges

A *lower variance* and *unbiased* estimator

Augmented Inverse Propensity Weighted Estimator

$$\widehat{Q}^{AIPW}(w) = \frac{1}{T}\sum_{t=1}^{T} \frac{\mathbb{I}\left\{W_t = w\right\}}{e_t(w)} Y_t + \left(1 - \frac{\mathbb{I}\left\{W_t = w\right\}}{e_t(w)}\right)\hat{\mu}_t$$

- **Notation:** $\hat{\mu}_t$ is any other estimator (including sample mean)
- *Additional term* helps *decrease variance*...
- ...but in bandit experiments, *variance is still too high*

(Remainder of the talk: IPW only)

Example

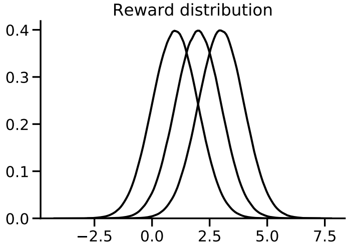
A simple example with *three treatments (arms)*.

True distribution
(unknown)

$Y_t(1) \sim \mathcal{N}(3,1)$ High reward

$Y_t(2) \sim \mathcal{N}(2,1)$

$Y_t(3) \sim \mathcal{N}(1,1)$ Low reward (Control)



Example

Let's compare

Adaptive vs *random* estimates of $Q(w)$.

Example

Let's compare

Adaptive vs *random* estimates of $Q(w)$.

Example

Let's compare

Adaptive vs *random* estimates of $Q(w)$.

Random

- Assign arms at random, with equal probability

Example

Let's compare

Adaptive vs *random* estimates of $Q(w)$.

Random

- Assign arms at random, with equal probability

Adaptive

- As we learn which arms are better, increase assignment probabilities to them (Thompson Sampling)

Example

Let's compare

Adaptive vs *random* estimates of $Q(w)$.

Random

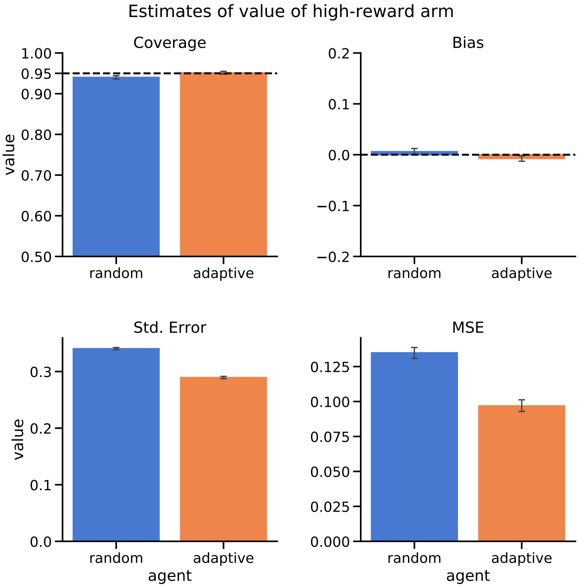
- Assign arms at random, with equal probability

Adaptive

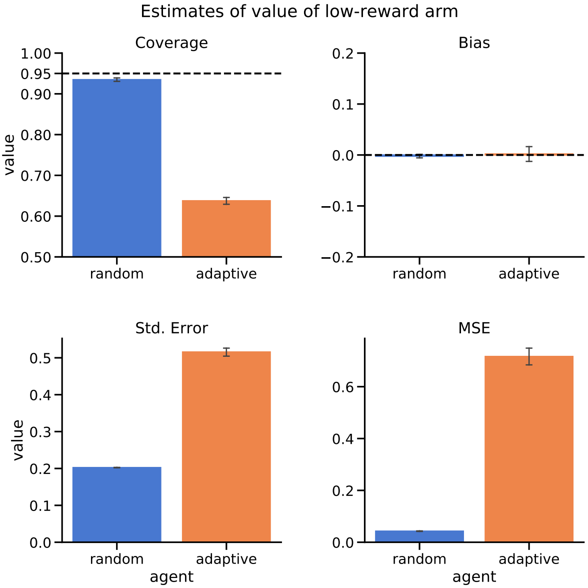
- As we learn which arms are better, increase assignment probabilities to them (Thompson Sampling)

Which data-collection algorithm
yields better *coverage*, *power*, and *mse*?

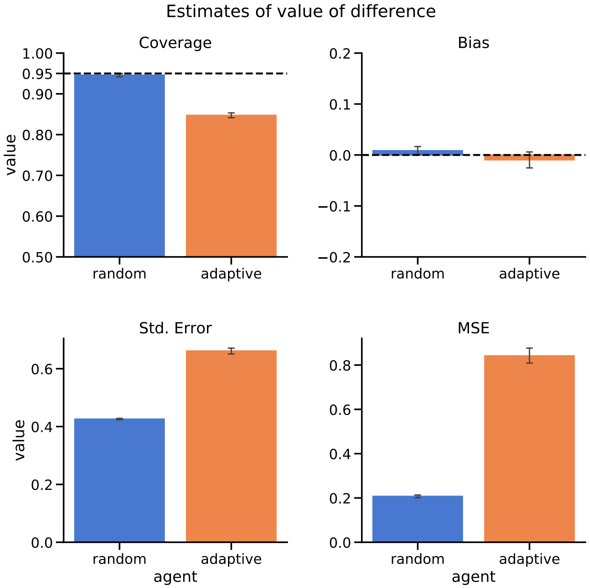
High-reward arm is estimated well



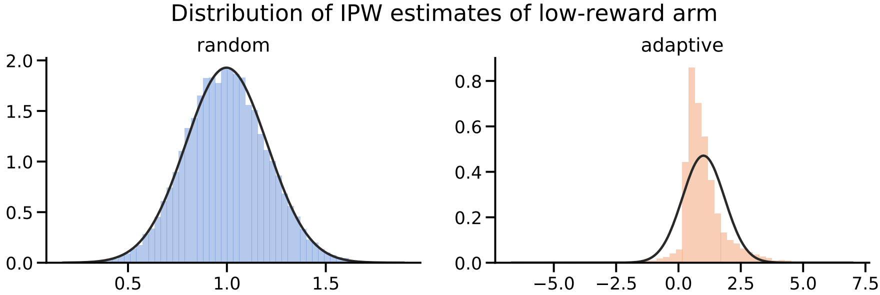
Low-reward arm is estimated poorly



Difference (High-Low) is estimated poorly too



Adaptivity hurts normality



central limit theorem may not hold

$\Longrightarrow$

Wrong confidence intervals!

Our contribution

Introduce *evaluation weights* $h_t(w)$.

Adaptively-Weighted Augmented Inverse Propensity Score Estimator

$$\widehat{Q}_T^{AW}(w) := \sum_{t=1}^{T} \frac{h_t(w)}{\sum_{s=1}^{T} h_s(w)} \left\{ \frac{\mathbb{I}\{W_t = w\}}{e_t(w)} Y_t + \left(1 - \frac{\mathbb{I}\{W_t = w\}}{e_t(w)}\right) \hat{\mu}_t \right\}$$

Our contribution

Introduce *evaluation weights* $h_t(w)$.

Adaptively-Weighted Augmented Inverse Propensity Score Estimator

$$\widehat{Q}_T^{AW}(w) := \sum_{t=1}^{T} \frac{h_t(w)}{\sum_{s=1}^{T} h_s(w)} \left\{ \frac{\mathbb{I}\left\{W_t = w\right\}}{e_t(w)} Y_t + \left(1 - \frac{\mathbb{I}\left\{W_t = w\right\}}{e_t(w)}\right) \hat{\mu}_t \right\}$$

- **Intuition:** $h_t(w)$ controls *flow of information*
- **Main result:** Conditions for CLT

Assumptions

A1: Bounded moments

Rewards have bounded fourth moments.

A2: Bounded $\hat{\mu}_t$

Estimate $\hat{\mu}_t$ is uniformly bounded, depends only on past data.

Assumptions

A3: Assignments

Assignment probabilities $e_t(w)$ depend on past data and converge in a.s. to a constant.

$$e_t(w) \xrightarrow[T\to\infty]{a.s.} e_\infty$$

A4: Evaluation weights

Weights $h_t(w)$ are non-negative, depend on past data.

Main result

A5: Growth and decay conditions (Main assumption)

Unnormalized weights $h_t(w)$ and assignment probabilities $e_t(w)$ jointly satisfy

$$\frac{\sum_{t=1}^{T} \frac{h_t(w)^p}{e_t(w)^{p-1}}}{\left(\sum_{t=1}^{T} \mathbb{E}\left[\frac{h_t(w)^2}{e_t(w)}\right]\right)^{p/2}} \xrightarrow[T\to\infty]{p} \begin{cases} \infty & \text{if } p = 1 \\ 1 & \text{if } p = 2 \\ 0 & \text{if } p \in \{3,4\} \end{cases}$$

In addition, for $p \in \{2,3,4\}$ this is bounded.

Joint constraint:

- *Data-collection algorithm*: $e_t(w)$ must decay slowly
- *Evaluation weights*: $h_t(w)$ chosen appropriately

Main result: Consistency and Asymptotic Normality

Theorem

If Assumption 1-5 are satisfied, then $\widehat{Q}_T^{AW}$ is a consistent estimator of the true arm value Q, and the following statistic is asymptotically normal.

$$\frac{\widehat{Q}_T(w)^{AW} - Q(w)}{\widehat{V}^{\frac{1}{2}}} \xrightarrow[T \to \infty]{d} \mathcal{N}(0, 1)$$

Main result: Consistency and Asymptotic Normality

Theorem

If Assumption 1-5 are satisfied, then $\widehat{Q}_T^{AW}$ is a consistent estimator of the true arm value Q, and the following statistic is asymptotically normal.

$$\frac{\widehat{Q}_T(w)^{AW} - Q(w)}{\widehat{V}^{\frac{1}{2}}} \xrightarrow[T\to\infty]{d} \mathcal{N}(0,1)$$

- **Notation:** $\widehat{V}_T$ is estimate of variance

Weighting schemes

Propensity-score based (`ps`)

Seeks minimum variance. When applied to IPW, gets us back approximately to simple averaging. Does not always satisfy A5.

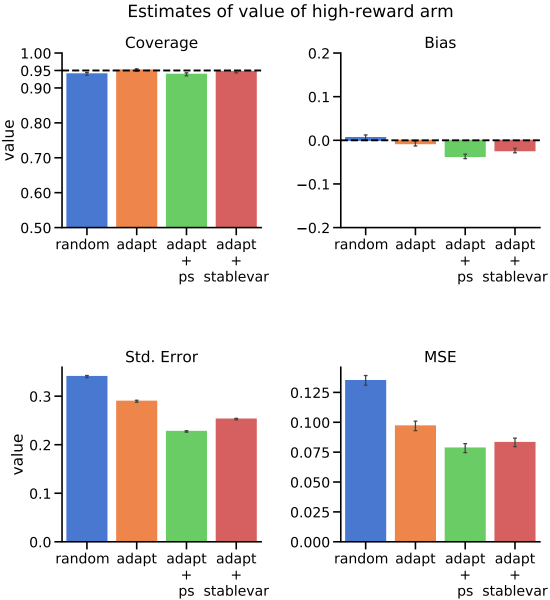
$$h_t = e_t(w)$$

Variance-stabilizing `stablevar`

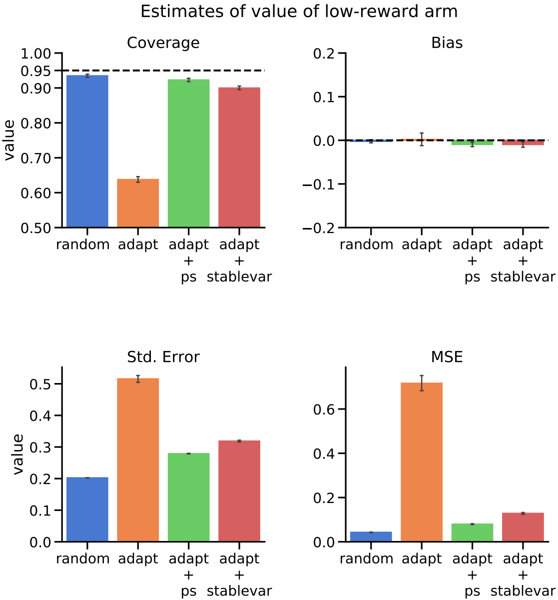
Ensures assumption A5(2) is satisfied.

$$h_t^2 = \left(1 - \sum_{s=1}^{t-1} h_s^2 e_s^{-1}\right) \mathbb{E}\left[\frac{e_t(w)}{\sum_{s=t}^{T} e_s} \,\Big|\, \mathcal{F}_{t-1}\right] e_t(w)$$

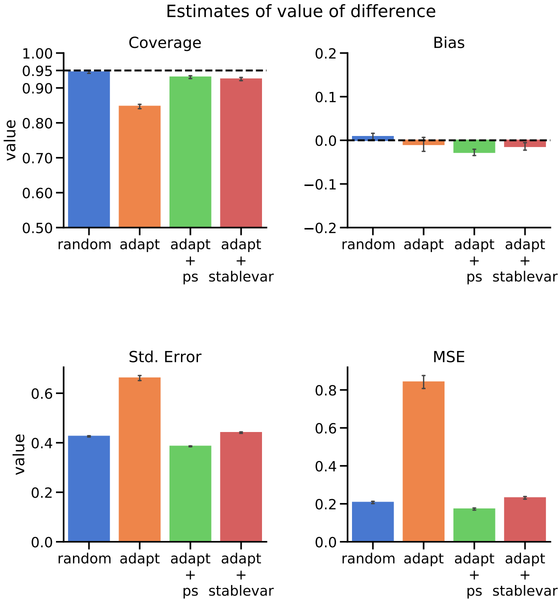
High-reward arms are (still) estimated well



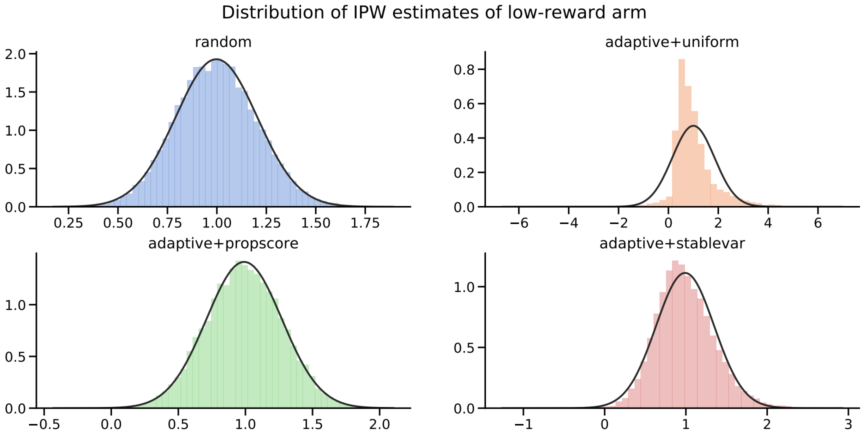
Low-reward arms are estimated well too



Difference is estimated well too



Normality is restored



Concluding thoughts

Appropriately chosen adaptive evaluation weights:

- Yield valid confidence intervals
- Produce estimates with smaller MSE
- Still allow for adaptive data collection

However: For hypothesis testing, beware of too-aggressive algorithms!