Effective Adaptive Exploration of Prices and Promotions in Choice-Based Demand Models

Lalit Jain, Zhaoqi Li, Erfan Loghmani, Blake Mason, Hema Yoganarasimhan

FOSTER SCHOOL OF BUSINESS
UNIVERSITY of WASHINGTON

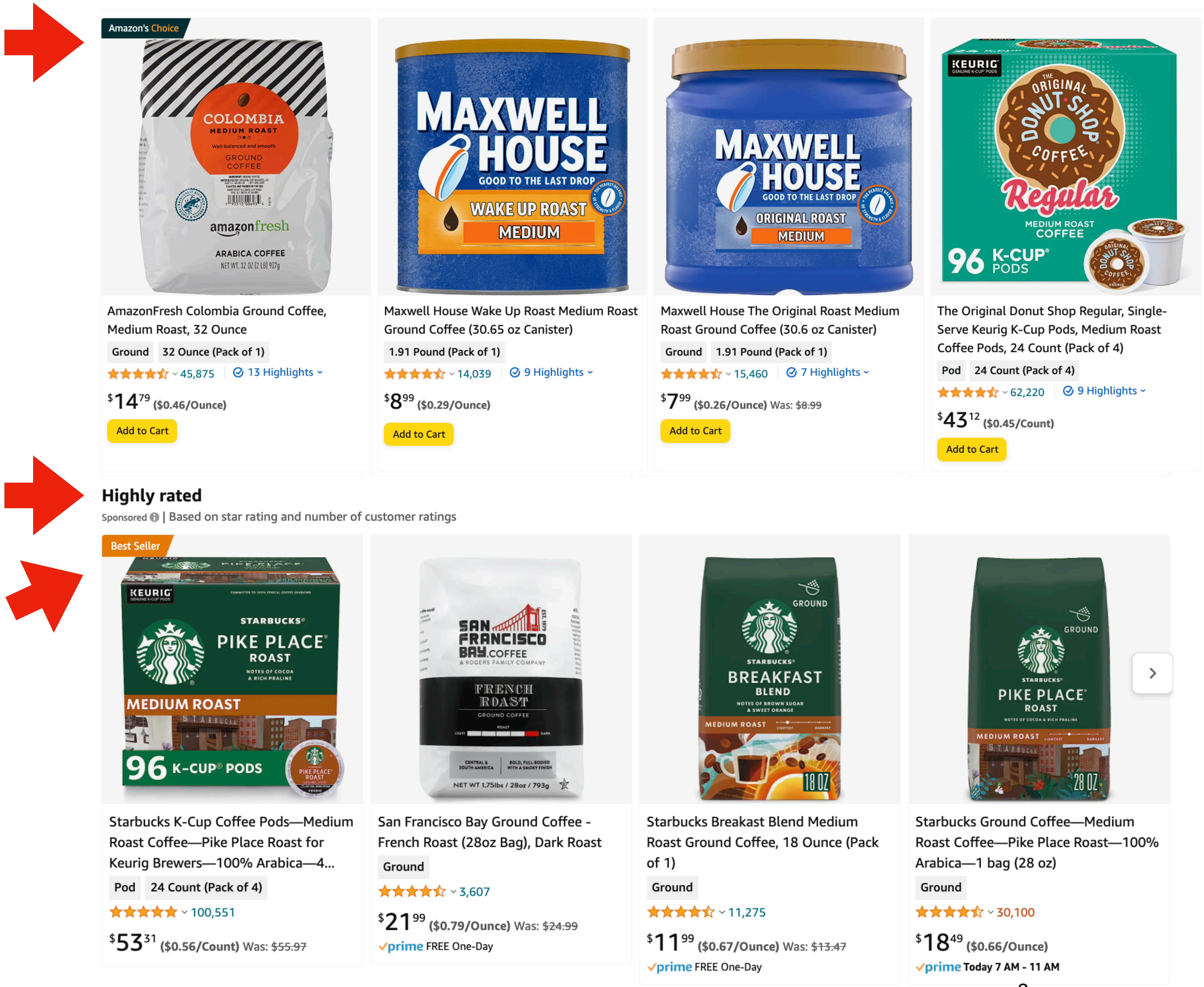
Outline

1. Introduction

2. Adaptively Setting Prices and Promotions

3. Incorporating Context

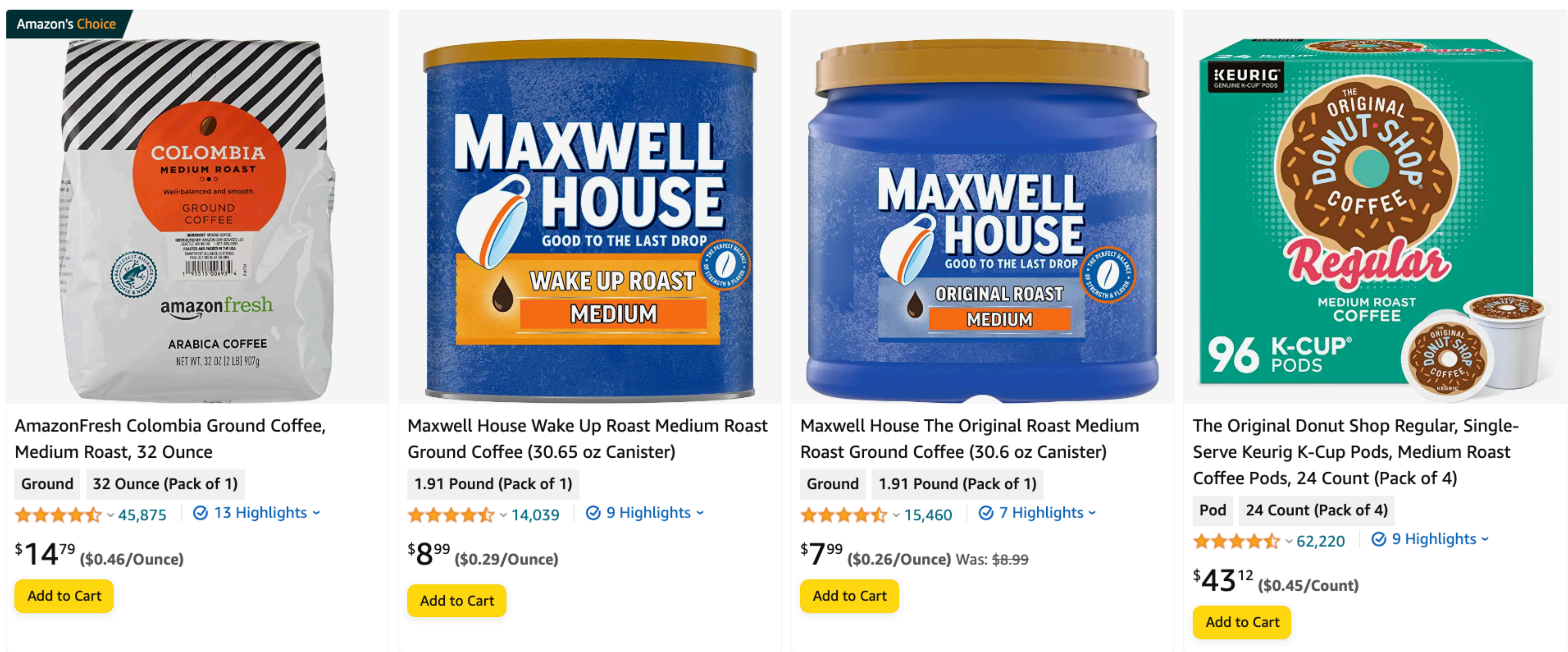
Problem: Setting Prices and Promotion



Firm needs to

- **set prices**
- decide **what items to promote**

Problem: Setting Prices and Promotion



Firm needs to

- **set prices**
- decide **what items to promote**
- by **how much** given a budget.

Why study this problem?

Amazon changes it's prices hourly, and can change a product's price as frequently as once every 10 minutes!

How can an online retailer **jointly** decide how to set prices and promotions in order to maximize profit?

If the demand curve for each consumer/product was known - we could jointly optimize.

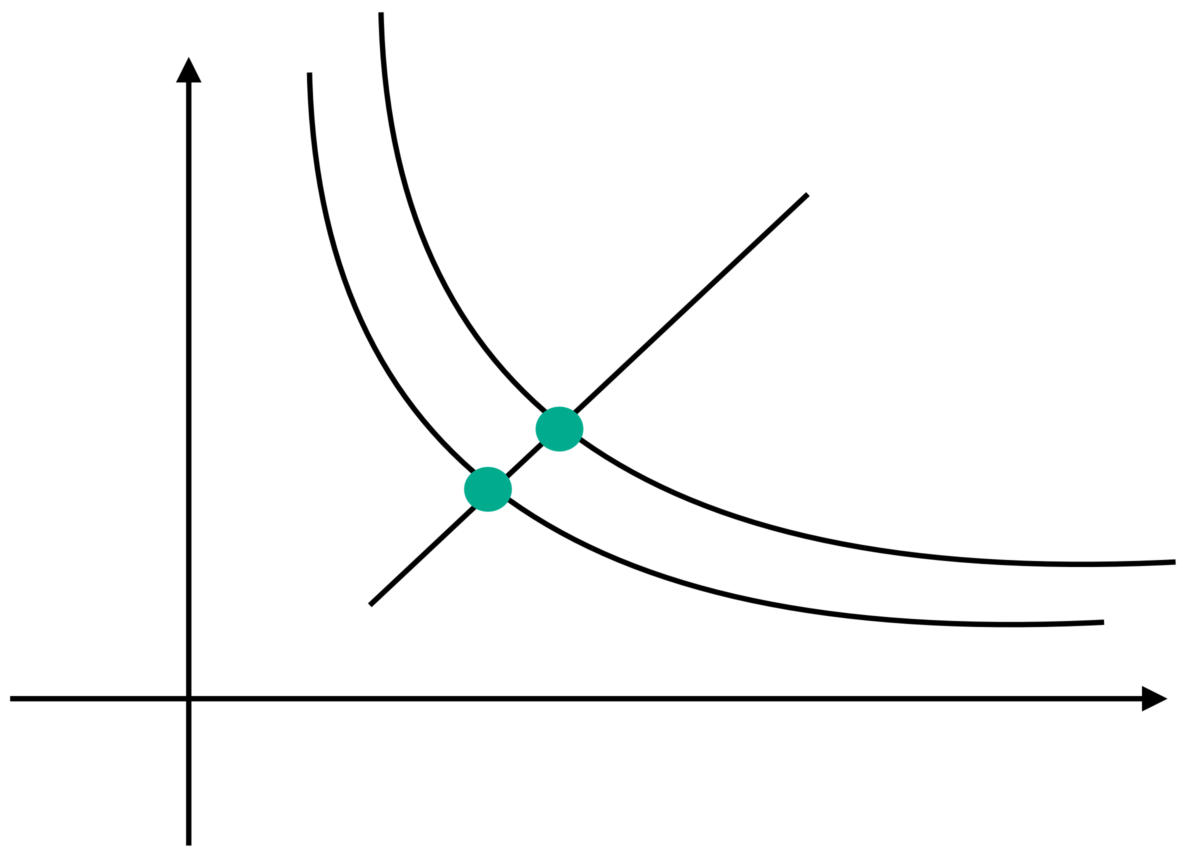
But in practice, the demand is rarely known and can depend on the customer segment.

Traditional Approaches for Pricing

Existing Approaches Suffer from a Lack of Strategic Exploration

Exploiting Historical Data

- Lack of Exogenous Variation
- "Greedy method" - no exploration



Structural Demand Estimation:

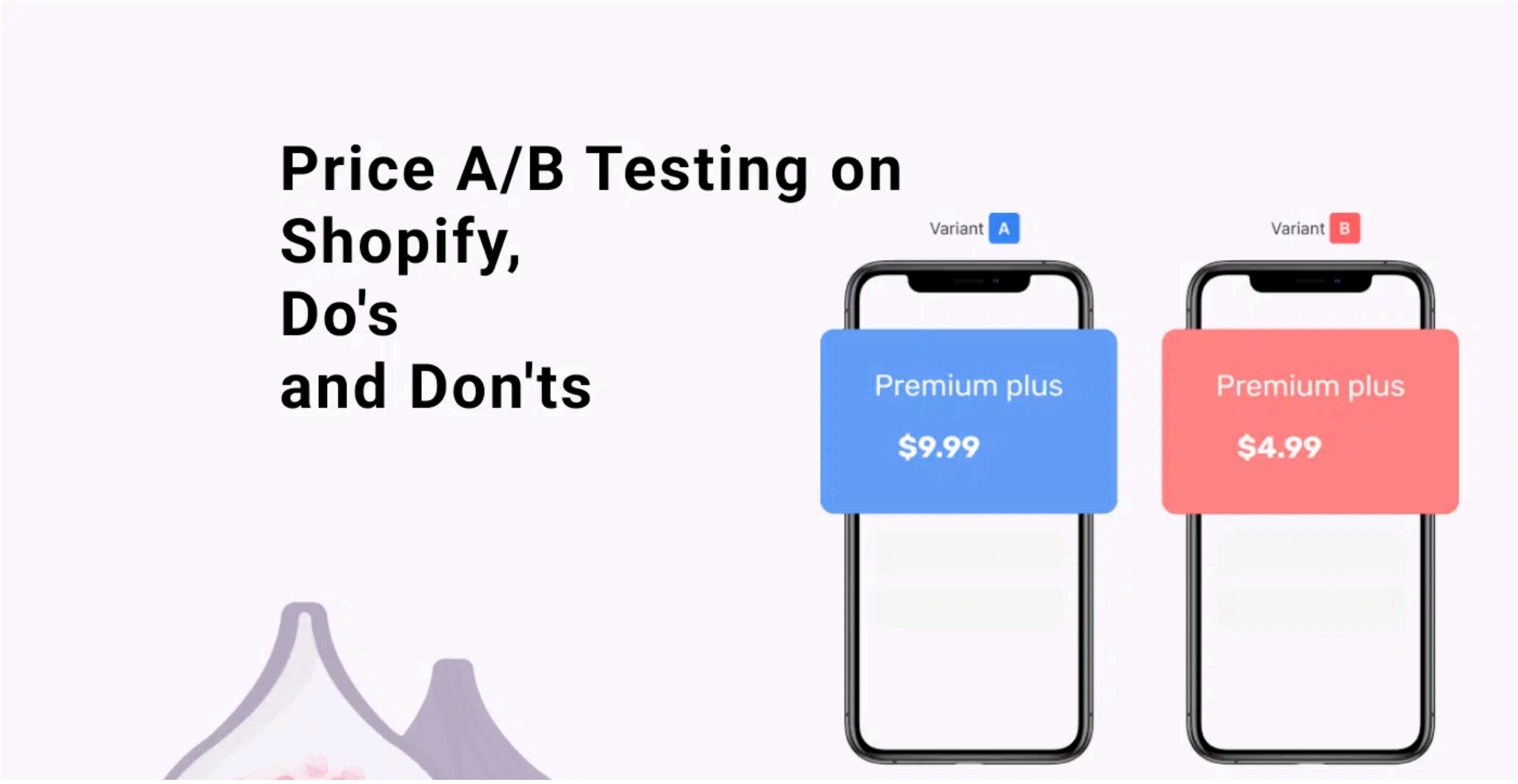
Berry+Levinson+Pakes '95, Guadagni+Little '83, Hitsch '06, and more…

Traditional Approaches for Pricing

Existing Approaches Suffer from a Lack of Strategic Exploration

Pricing Experiments (A/B tests)

- Potentially High Opportunity cost
- Extremely large action space with many products



Literature:

Aghnion et al '91, Dube and Misra '17, …

Adaptive Pricing: Protocol

Assume K products and one outside option

For $t = 1,2,3,\cdots, T$

Customer(s) arrives at platform

Firm chooses price $\mathbf{p}_t = (p_{1t}, \cdots, p_{Kt}) \in [\ell, u]^K$

and promotion $\mathbf{x}_t = (x_{1t}, \cdots, x_{Kt}) \in \mathsf{X} \subset [0,1]^K$,

Observe purchase decision(s) $I_t \in \{0,1,\cdots, K\}$, and collect revenue p_{I_t}

*Easily extended to the batched setting

Adaptive Pricing: Minimize Regret

Expected Profit

$$R(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} \mathbb{P}(I = k \,|\, \mathbf{p}, \mathbf{x})(p_k - m_k)$$

Adaptive Pricing: Minimize Regret

Expected Profit

$$R(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} \mathbb{P}(I = k \,|\, \mathbf{p}, \mathbf{x})(p_k - m_k)$$

Optimal Price/Promotion

$$\mathbf{p}_\star, \mathbf{x}_\star = \arg \max_{\mathbf{p} \in [\ell, u]^K, x \in \mathsf{X}} R(\mathbf{p}, \mathbf{x})$$

Adaptive Pricing: Minimize Regret

Expected Profit

$$R(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} \mathbb{P}(I = k \mid \mathbf{p}, \mathbf{x})(p_k - m_k)$$

Optimal Price/Promotion

$$\mathbf{p}_\star, \mathbf{x}_\star = \arg \max_{\mathbf{p} \in [\ell, u]^K, x \in \mathsf{X}} R(\mathbf{p}, \mathbf{x})$$

Goal: Minimize Regret

$$Reg_T = \sum_{t=1}^{T} R(\mathbf{p}_\star, \mathbf{x}_\star) - R(\mathbf{p}_t, \mathbf{x}_t)$$

Adaptive Pricing: Minimize Regret

Expected Profit

$$R(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} \mathbb{P}(I = k \mid \mathbf{p}, \mathbf{x})(p_k - m_k)$$

Optimal Price/Promotion

$$\mathbf{p}_\star, \mathbf{x}_\star = \arg \max_{\mathbf{p} \in [\ell, u]^K, x \in \mathsf{X}} R(\mathbf{p}, \mathbf{x})$$

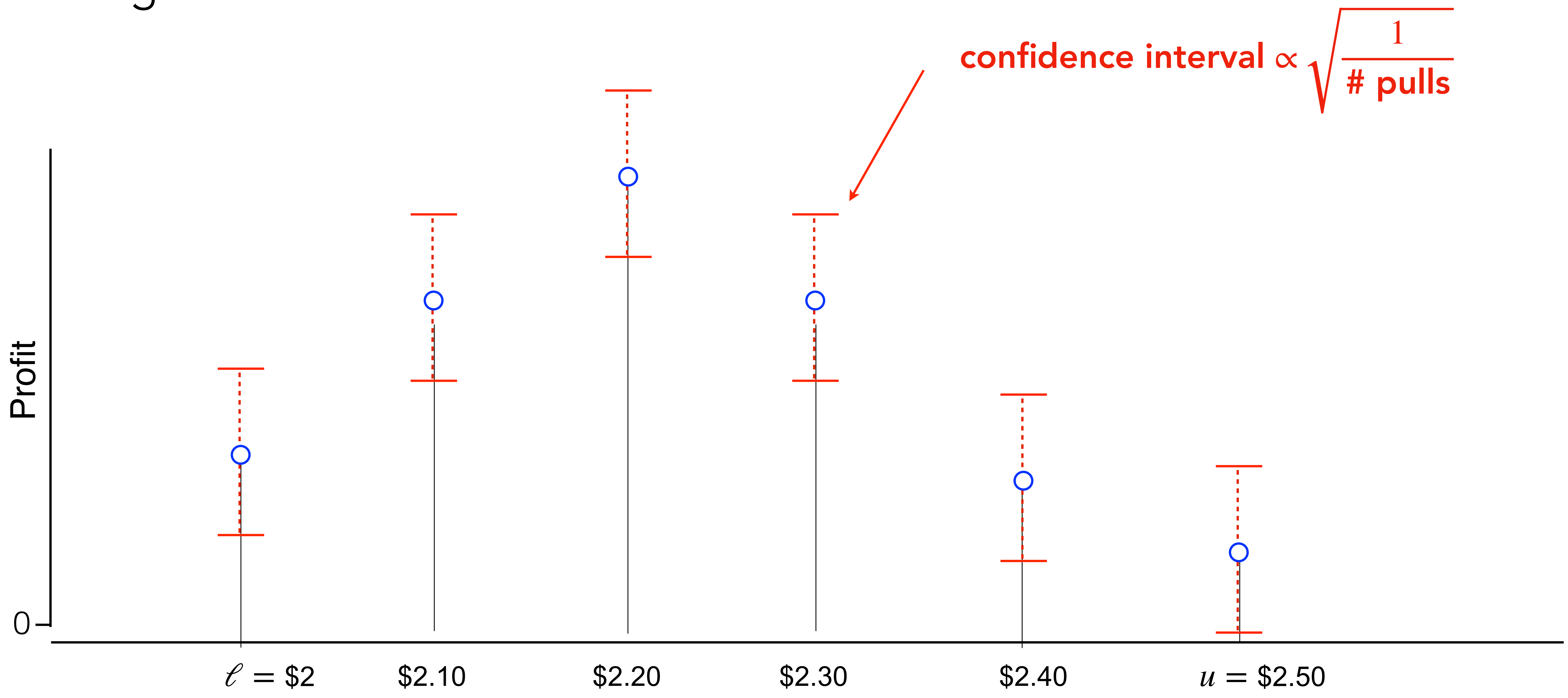
Goal: Minimize Regret

$$Reg_T = \sum_{t=1}^{T} R(\mathbf{p}_\star, \mathbf{x}_\star) - R(\mathbf{p}_t, \mathbf{x}_t)$$

We want to minimize our opportunity cost of learning the optimal price and promotion. Ideally $Reg_T / T \to 0$

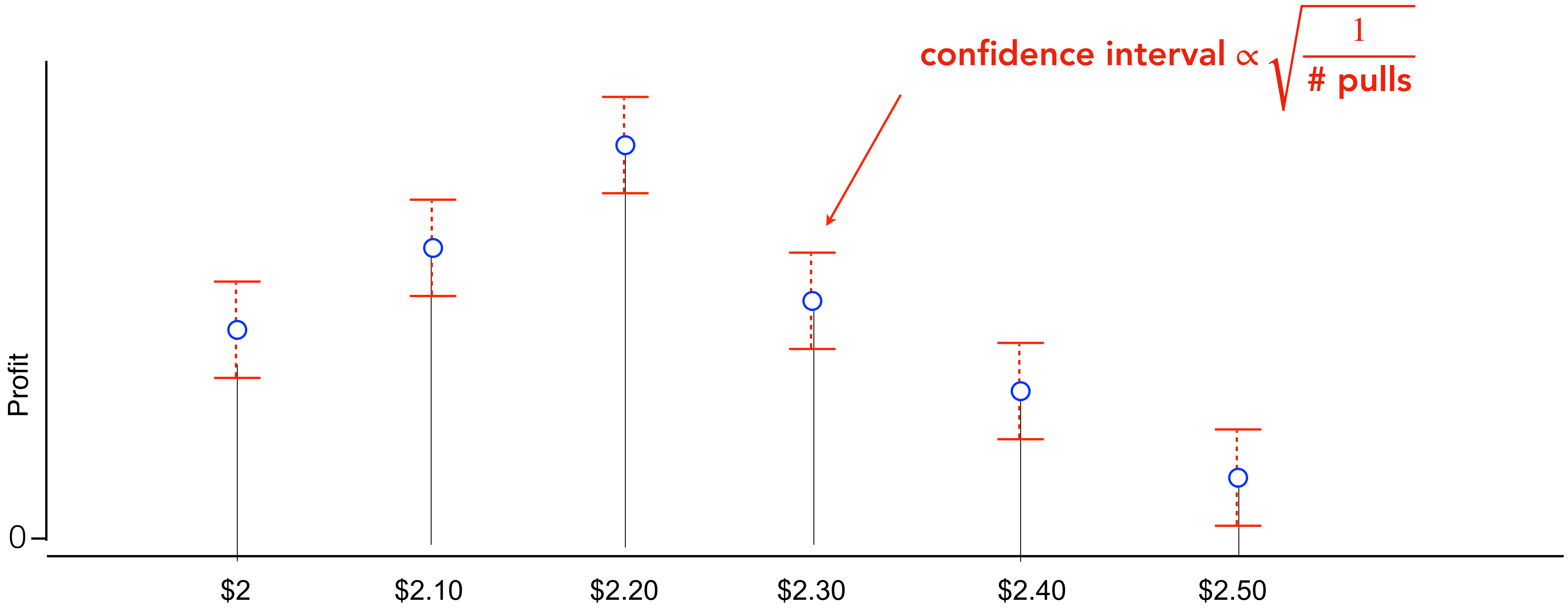
Review of Single Product Pricing: A/B/N testing

- A/B/n Testing:



Review of Single Product Pricing: A/B/N testing

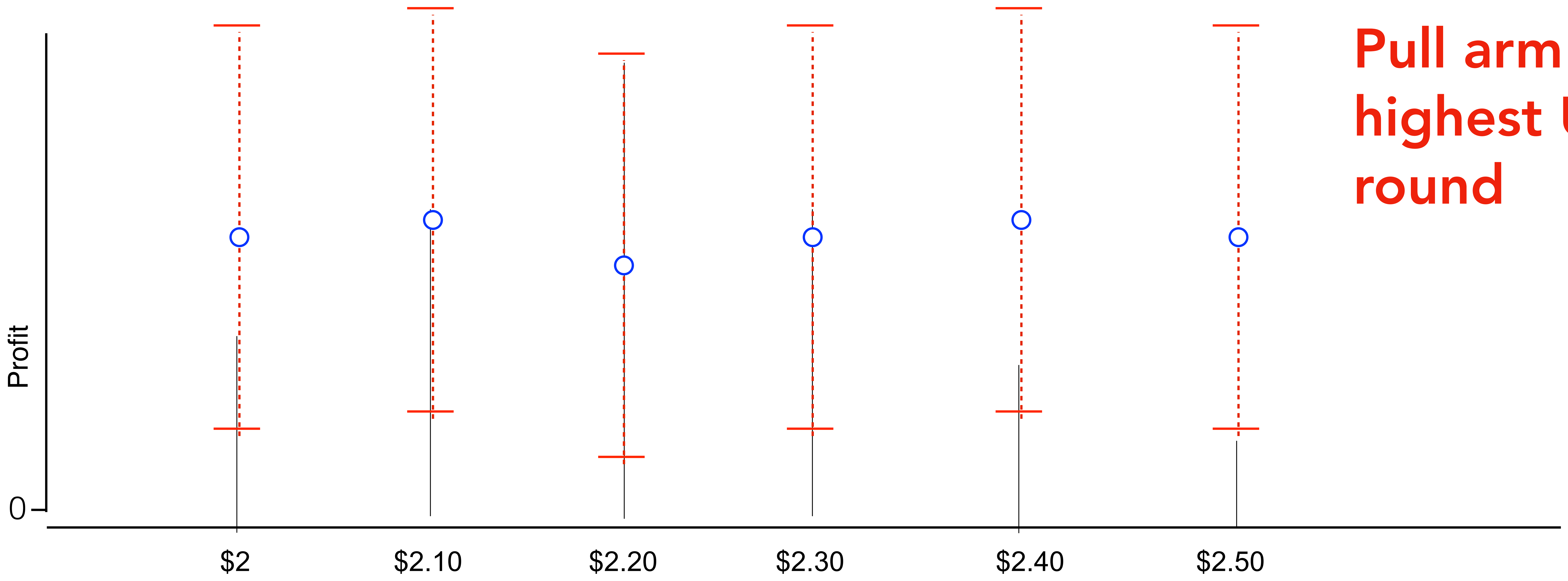
- A/B/n Testing:



- Necessarily $O(T)$ regret!

Review of Single Product Pricing: A/B/N testing

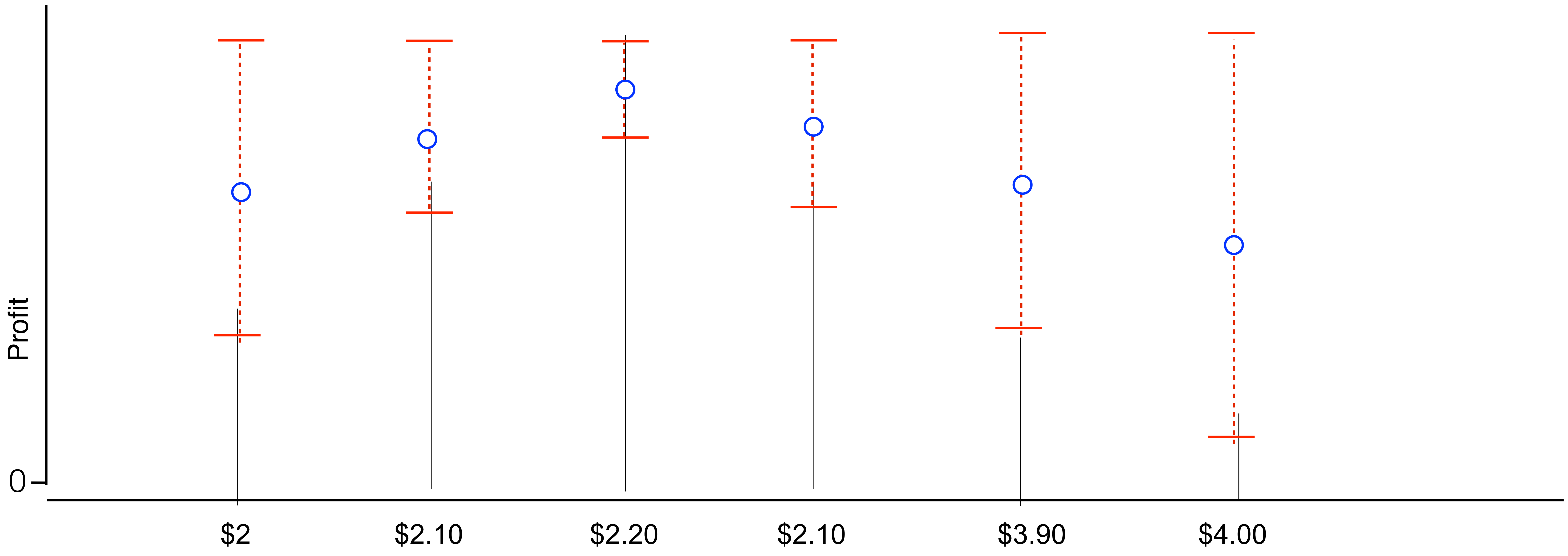
- **UCB Approach**: Multi-Armed-Bandits on discrete set of prices [KleinbergLeighton '03, MisraSchwarzAbernethy'19]



Pull arm with highest UCB each round

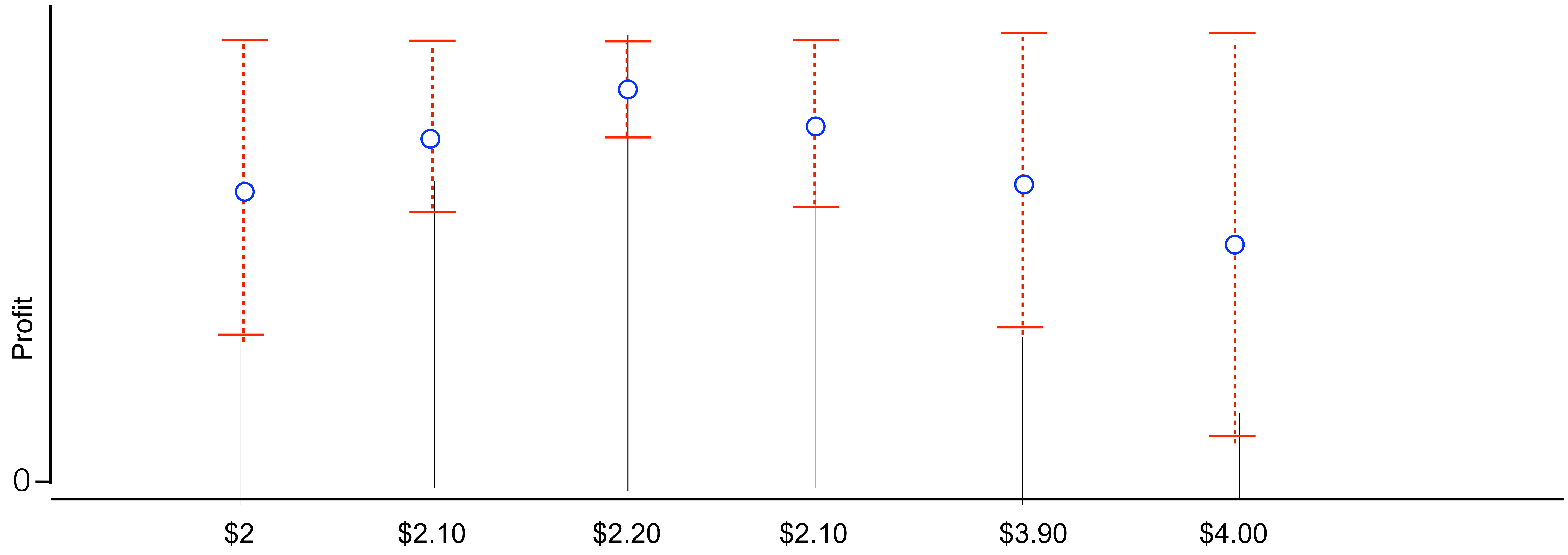
Review of Single Product Pricing: A/B/N testing

- **UCB Approach**: Multi-Armed-Bandits on discrete set of prices
  [KleinbergLeighton '03, MisraSchwarzAbernethy'19]



Review of Single Product Pricing: A/B/N testing

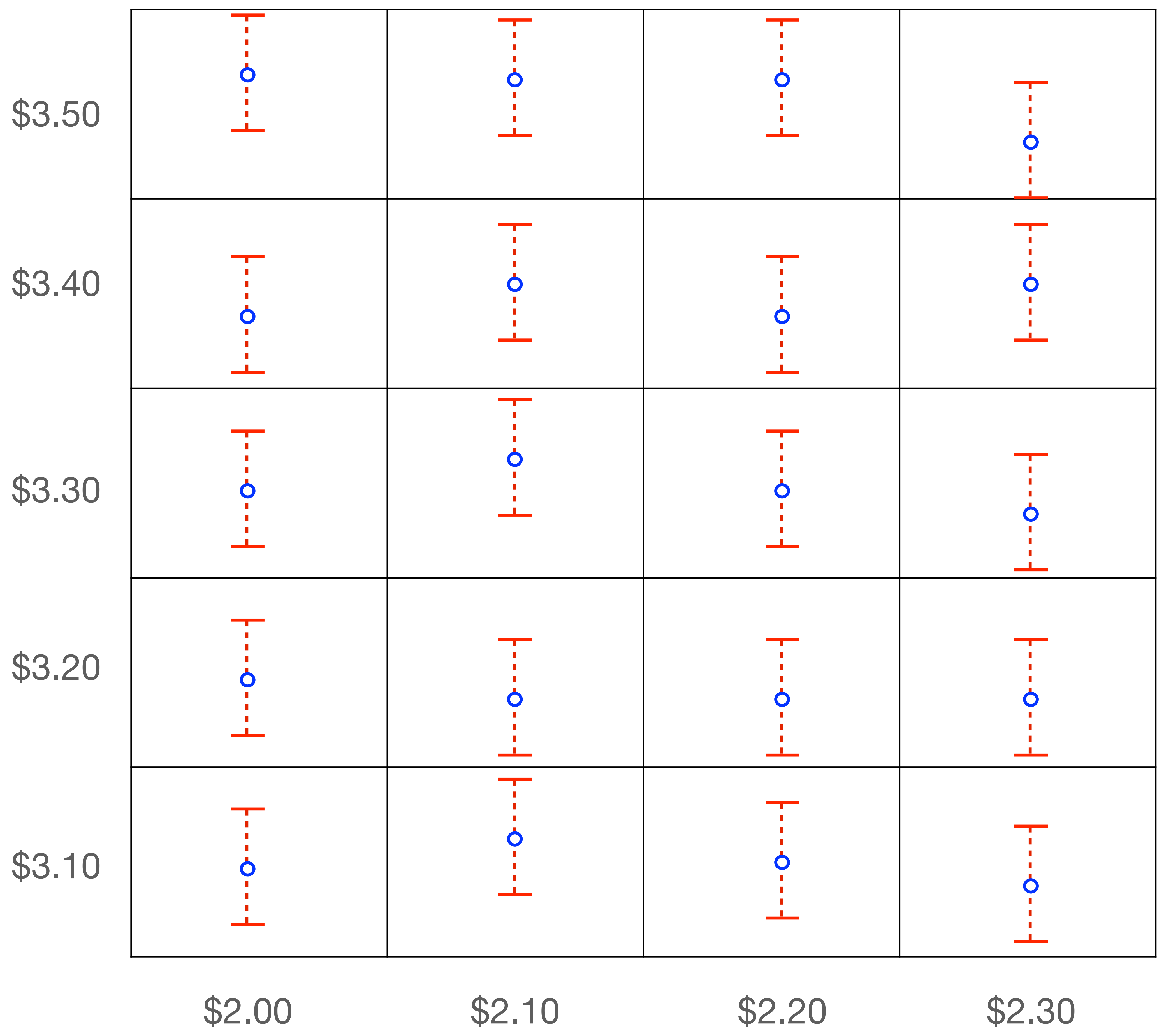
- **UCB Approach**: Multi-Armed-Bandits on discrete set of prices [KleinbergLeighton '03, MisraSchwarzAbernethy'19]



- Can guarantee $O(\sqrt{DT} + \epsilon T)$ in general
  - If profit function is "strongly concave" can choose D so regret is $O(\sqrt{T})$

Extending to Multiple Products

- Natural Approach: Multi-Armed-Bandits on a discrete set of prices



- Number of price combinations grows exponentially with number of products!

Shortfalls of Discretized Approaches

- Number of price combinations grows exponentially with number of products!
- Difficult to add promotions to the model
- Can't handle customer heterogeneity
- Not exploiting the "smoothness" of the problem

Fundamentally, a totally non-parametric approach is difficult to scale!*

Outline

1. Introduction

2. Adaptively Setting Prices and Promotions

3. Incorporating Context

Adaptive Pricing: Our Approach

Adaptive Pricing: Our Approach

- **Strategic Exploration**
  - Explore to learn the demand curve while Exploiting current information

Adaptive Pricing: Our Approach

- **Strategic Exploration**
  - Explore to learn the demand curve while Exploiting current information
- **Exploit demand curve to reduce experimentation cost**
  - Random utility choice model

Adaptive Pricing: Our Approach

- **Strategic Exploration**
  - Explore to learn the demand curve while Exploiting current information
- **Exploit demand curve to reduce experimentation cost**
  - Random utility choice model
- **Flexible**
  - Can accommodate both prices and promotions
  - Can incorporate customer heterogeneity

Model (McFadden '77)

K products, marginal costs $m \in \mathbb{R}^K_{\geq 0}$

Price Variable $\in [\ell, u]$

Promotion Variable $\in [0,1]$

$$u_{tk}(\mathbf{p}, \mathbf{x}) = \alpha_k - \beta_k p_k + \gamma_k x_k + \epsilon_{tk}$$

Type-1 GEV

Product k utility for user t

Model Parameters
$\theta = [(\alpha_k, \beta_k, \gamma_k)]_{k=1}^K \in \mathbb{R}^{3K}$

Model (McFadden '77)

K products, marginal costs $m \in \mathbb{R}^K_{\geq 0}$

Price Variable $\in [\ell, u]$

Promotion Variable $\in [0,1]$

$$u_{tk}(\mathbf{p}, \mathbf{x}) = \alpha_k - \beta_k p_k + \gamma_k x_k + \epsilon_{tk}$$

Type-1 GEV

Product k utility for user t

Model Parameters

$\theta = [(\alpha_k, \beta_k, \gamma_k)]_{k=1}^K \in \mathbb{R}^{3K}$

Demand

$$Q_k(\mathbf{p}, \mathbf{x}) := \mathbb{P}_\theta(I = k \,|\, \mathbf{p}, \mathbf{x}) = \frac{e^{u_k(p_k, x_k)}}{1 + \sum_{j=1}^K e^{u_j(p_j, x_j)}}$$

Model (McFadden '77)

K products, marginal costs $m \in \mathbb{R}^K_{\geq 0}$

Price Variable $\in [\ell, u]$

Promotion Variable $\in [0,1]$

$$u_{tk}(\mathbf{p}, \mathbf{x}) = \alpha_k - \beta_k p_k + \gamma_k x_k + \epsilon_{tk}$$

Type-1 GEV

Product k utility for user t

Model Parameters

$\theta = [(\alpha_k, \beta_k, \gamma_k)]_{k=1}^K \in \mathbb{R}^{3K}$

Demand

$$Q_k(\mathbf{p}, \mathbf{x}) := \mathbb{P}_\theta(I = k \,|\, \mathbf{p}, \mathbf{x}) = \frac{e^{u_k(p_k, x_k)}}{1 + \sum_{j=1}^K e^{u_j(p_j, x_j)}}$$

Expected Profit

$$R_\theta(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^K (p_k - m_k) Q_k(\mathbf{p}, \mathbf{x})$$

Model: Bayesian Approach

$$Q_k(\mathbf{p}, \mathbf{x}) := \mathbb{P}_\theta(\text{choose } k \,|\, \mathbf{p}, \mathbf{x}) = \frac{e^{u_k(p_k, x_k)}}{1 + \sum_{j=1}^{K} e^{u_j(p_j, x_j)}}$$

$$R_\theta(\mathbf{p}, \mathbf{x}) := \sum_{k=1}^{K} (p_k - m_k) Q_k(\mathbf{p}, \mathbf{x})$$

Bayesian Approach

Assume a prior Π_0, and $\theta \sim \Pi_0$

Objective: Minimize Bayesian Regret

Let $\mathbf{p}_\star, \mathbf{x}_\star$ be the optimal price and promotion.

$$Reg_T = \mathbb{E}_{\theta \sim \Pi_0}\left[\sum_{t=1}^{T} R_\theta(\mathbf{p}_\star, \mathbf{x}_\star) - R_\theta(\mathbf{p}_t, \mathbf{x}_t)\right]$$

Dynamic Pricing Literature

- **Parametric Generalized Linear Settings:** [KeskinZeevi'14],[BoerZwart'14], …
- **Non-Parametric:** [BesbesZeevi'09],…, [MisraSchwarzAbernethy'19]
- **Choice Models:** [JavanmardNazerzedahShao'19, MiaoChao'21]
- **Assortment Selection:** in retail settings, impossible to know choice set.

Our work is the first to consider:

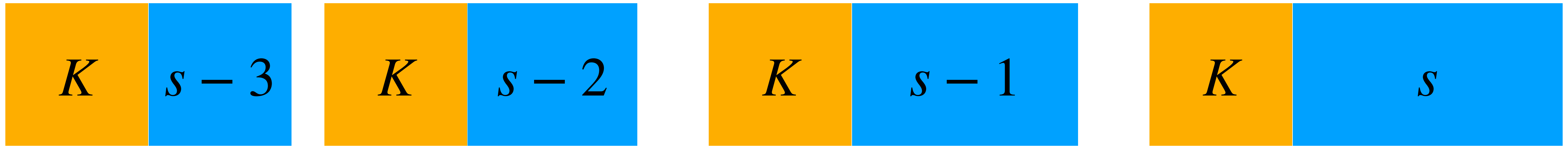
a) Choice Models

b) Promotion variables

c) Consumer Heterogeneity

Existing Work: Forced Exploration

Epoch: $s = 1,2,\cdots$

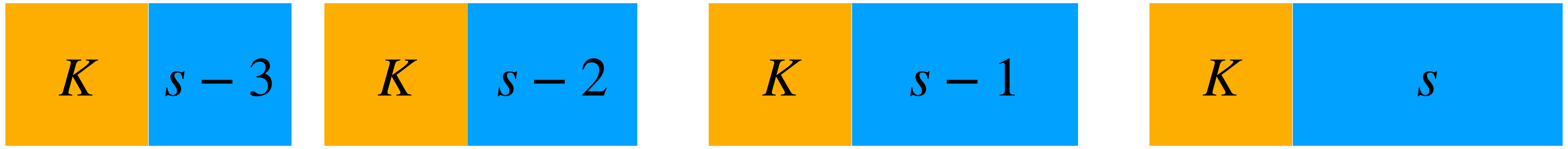


Pure Exploration Phases: Play K random prices, MLE estimate $\hat{\theta}_s$

Pure Exploitation Phases: Play $\hat{\mathbf{p}}_\mathbf{s} = \arg\max_{\mathbf{p}} R_{\hat{\theta}_s}(\mathbf{p})$

Existing Work: Forced Exploration

Epoch: $s = 1,2,\cdots$



Pure Exploration Phases: Play K random prices, MLE estimate $\hat{\theta}_s$

Pure Exploitation Phases: Play $\hat{\mathbf{p}}_\mathbf{s} = \arg\max_{\mathbf{p}} R_{\hat{\theta}_s}(\mathbf{p})$

Guarantee*: $Reg_T \leq O(K\sqrt{T})$

Playing randomized prices is not particularly feasible in practice!

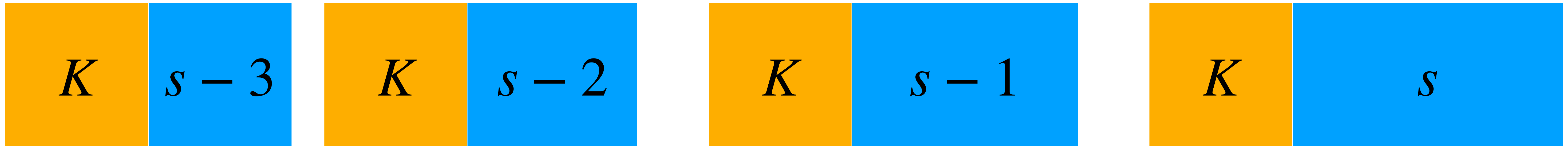
Can't handle promotions easily!

A form of ϵ-Greedy

[BroderRusmevichientong'12]
[´JavanmardNazerzedahShao'20]

Existing Work: Forced Exploration

Epoch: $s = 1,2,\cdots$



$t \approx Ks + s^2$

$K/\sqrt{t}$ in exploration

Pure Exploration Phases: Play K random prices, MLE estimate $\hat{\theta}_s$

Pure Exploitation Phases: Play $\hat{\mathbf{p}}_\mathbf{s} = \arg\max_{\mathbf{p}} R_{\hat{\theta}_s}(\mathbf{p})$

Guarantee*: $Reg_T \leq O(K\sqrt{T})$

Playing randomized prices is not particularly feasible in practice!

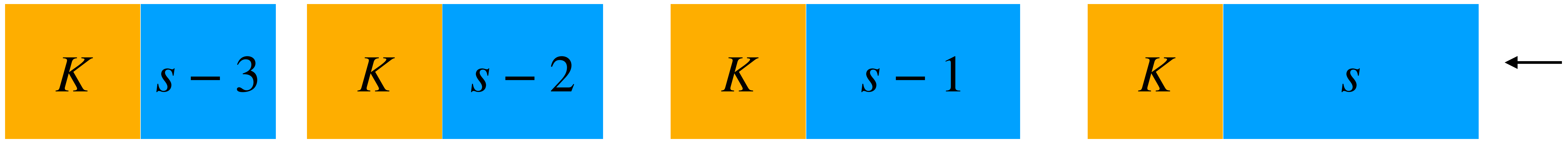
Can't handle promotions easily!

A form of ϵ-Greedy

[BroderRusmevichientong'12]
[´JavanmardNazerzedahShao'20]

Existing Work: Forced Exploration

Epoch: $s = 1,2,\cdots$



$t \approx Ks + s^2$

$K/\sqrt{t}$ in exploration

Pure Exploration Phases: Play K random prices, MLE estimate $\hat{\theta}_s$

Pure Exploitation Phases: Play $\hat{\mathbf{p}}_\mathbf{s} = \arg\max_{\mathbf{p}} R_{\hat{\theta}_s}(\mathbf{p})$

Guarantee*: $Reg_T \leq O(K\sqrt{T})$

Playing randomized prices is not particularly feasible in practice!

Can't handle promotions easily!

A form of ϵ-Greedy

[BroderRusmevichientong'12]
[´JavanmardNazerzedahShao'20]

Our Approach: Thompson/Posterior Sampling

Model Based Exploration ←

Model Based Optimization ←

Input: K products, X promotion set

Initialize: Π_0 as some prior distribution over θ

For $t = 1,2,3,\cdots,T$:

1. **Sample** $\theta_t \sim \Pi_t$
2. **Set** best price/promotion for θ_t:

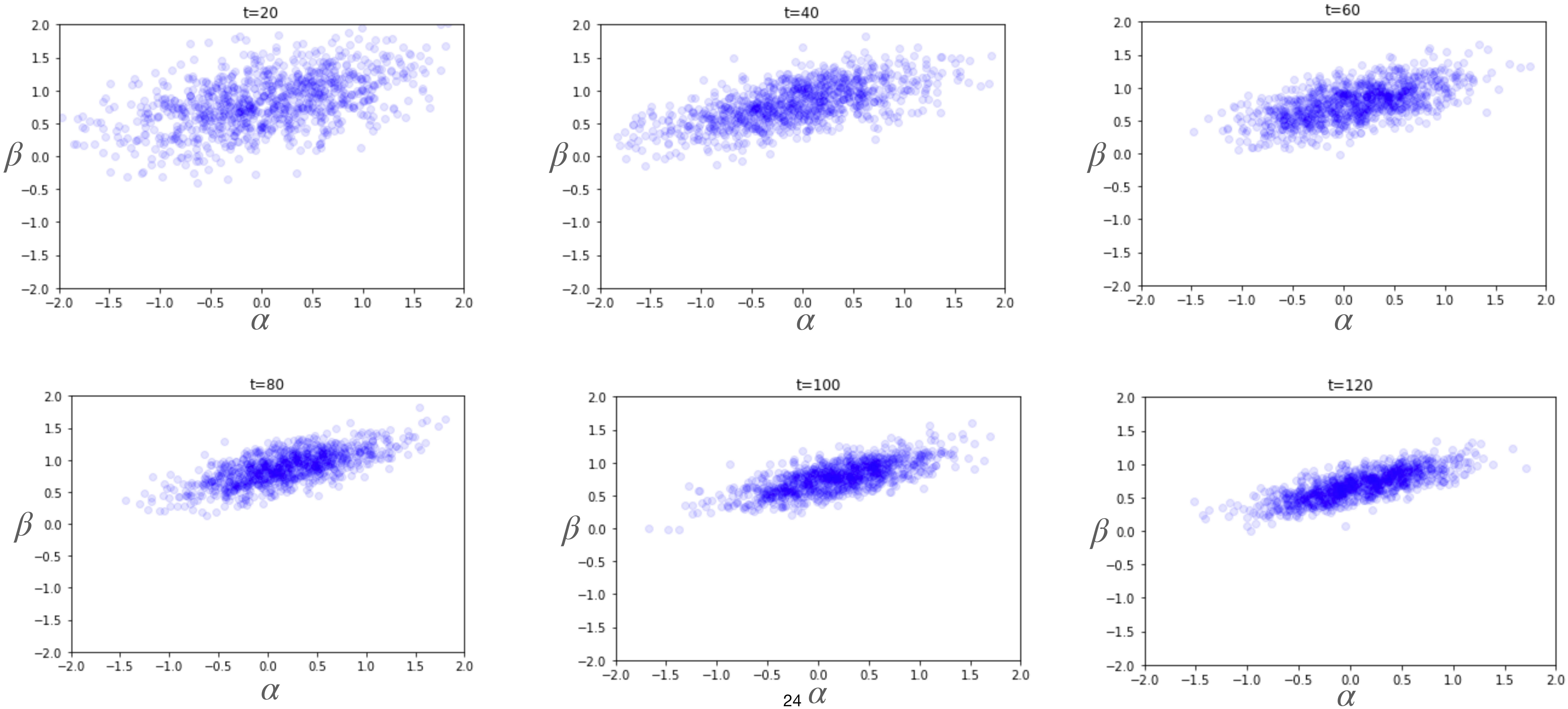
$$\mathbf{p_t}, \mathbf{x_t} = \arg\max_{\mathbf{p},\mathbf{x}} R_{\theta_t}(\mathbf{p}, \mathbf{x})$$

4. **Observe** $I_t \sim Q_t(\mathbf{p}_t, \mathbf{x}_t)$, collect revenue p_{t,I_t}
5. **Update** $\Pi_{t+1} =$ Posterior(Π_t, θ_{t+1})

Our Approach: Thompson Sampling

- **Model Based Exploration and Pricing:** Exploration is driven by the model, not by playing random prices
- **Computational Advantages:** Easily implemented if you can *sample* from the posterior. Maintaining the posterior is impossible in many settings, but sampling is straightforward.
- **Easily Extended:** Can easily incorporate additional features to the model

Thompson Sampling: Intuition



Implementation Challenges

1. Optimizing over p and x at the same time is non-convex and high dimensional

$$\mathbf{p_t}, \mathbf{x_t} = \arg\max_{\mathbf{p},\mathbf{x}} R_{\theta_t}(\mathbf{p}, \mathbf{x})$$

2. Posterior Computation Π_t

Optimize Prices with Fixed Promotions

Revenue

$$R_\theta(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} (p_k - m_k) \frac{e^{\alpha_k - \beta_k p_k + \gamma_k x_k}}{1 + \sum_{j=1}^{K} e^{\alpha_k - \beta_k p_k + \gamma_k x_k}}$$

$$\mathbf{p_t}, \mathbf{x_t} = \arg\max_{\mathbf{p},\mathbf{x}} R_{\theta_t}(\mathbf{p}, \mathbf{x})$$

Unfortunately non-convex in $\mathbf{p}$

Lemma (Aydin & Ryan '00) For a fixed value of $\mathbf{x}$,

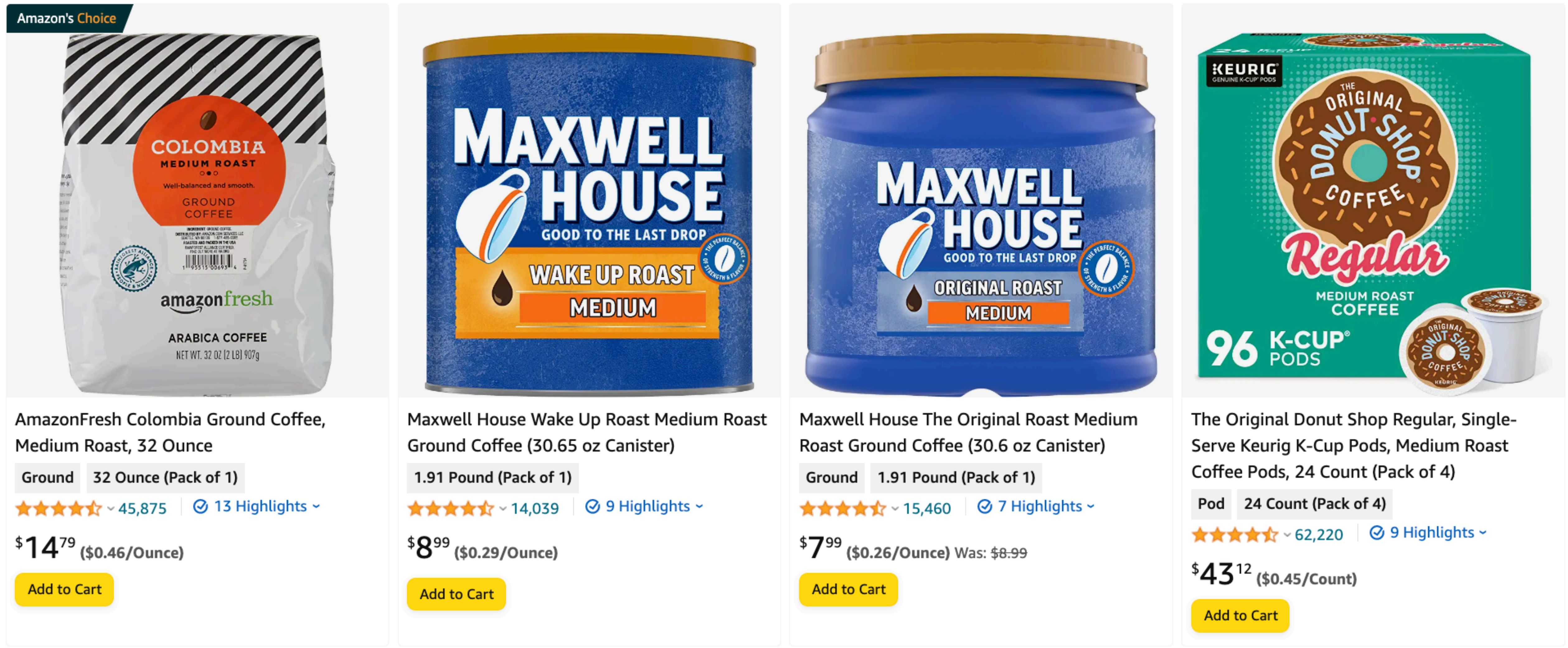
$$\mathbf{p}_{*,i} = \frac{1}{\beta_i} + R \qquad R = \sum_{i=1}^{K} \frac{1}{\beta_i} e^{-(1+\beta_i R)} e^{\alpha_i + \gamma_i x_i}$$

- A fairly fast binary search procedure works well

Optimizing Promotion at a Fixed Price

Easy Setting: X finite and combinatorial

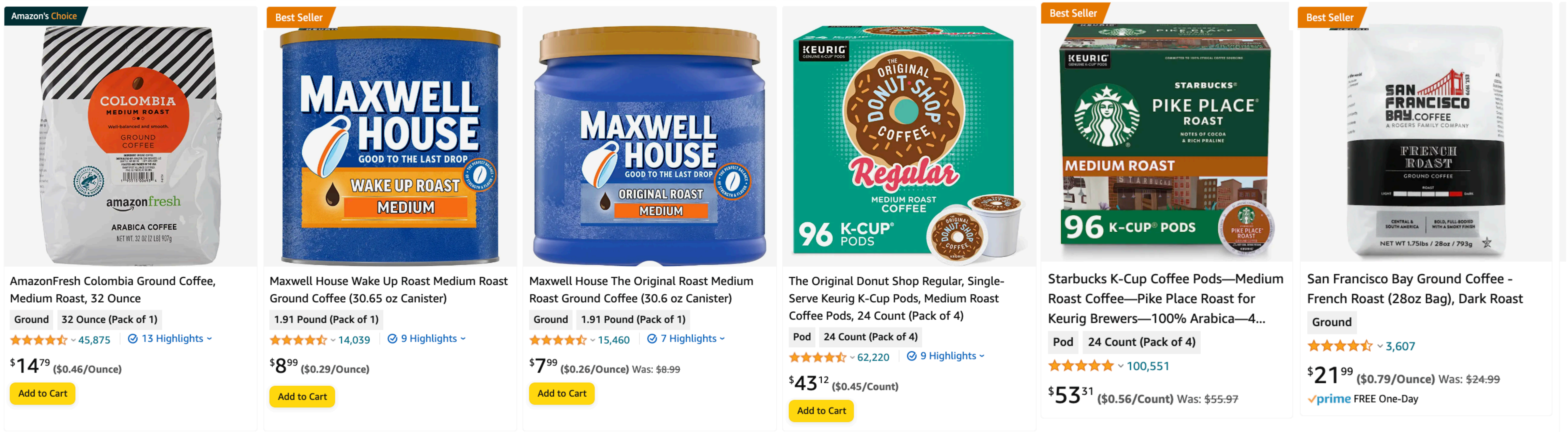
- e.g. $X = \{e_1, \cdots, e_K\}$ - we can promote at most one item



Optimizing Promotion at a Fixed Price

Easy Setting: X finite and combinatorial

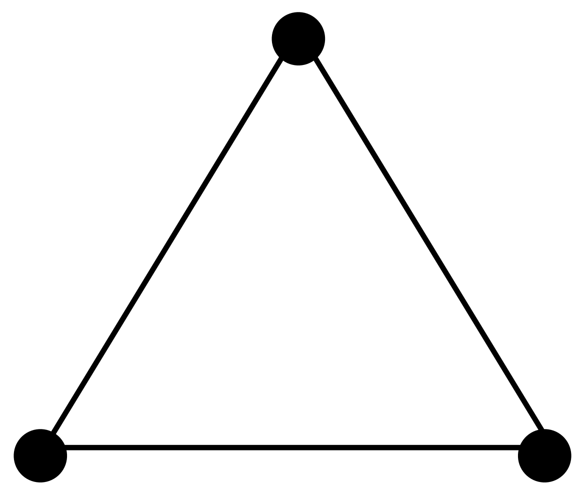
- e.g. $X = \{e_1, \cdots, e_K\}$ - we can promote at most one item
- e.g. $X = \{0,1\}^K$ - we can promote a subset of items



But discrete settings don't capture magnitude of the promotion.

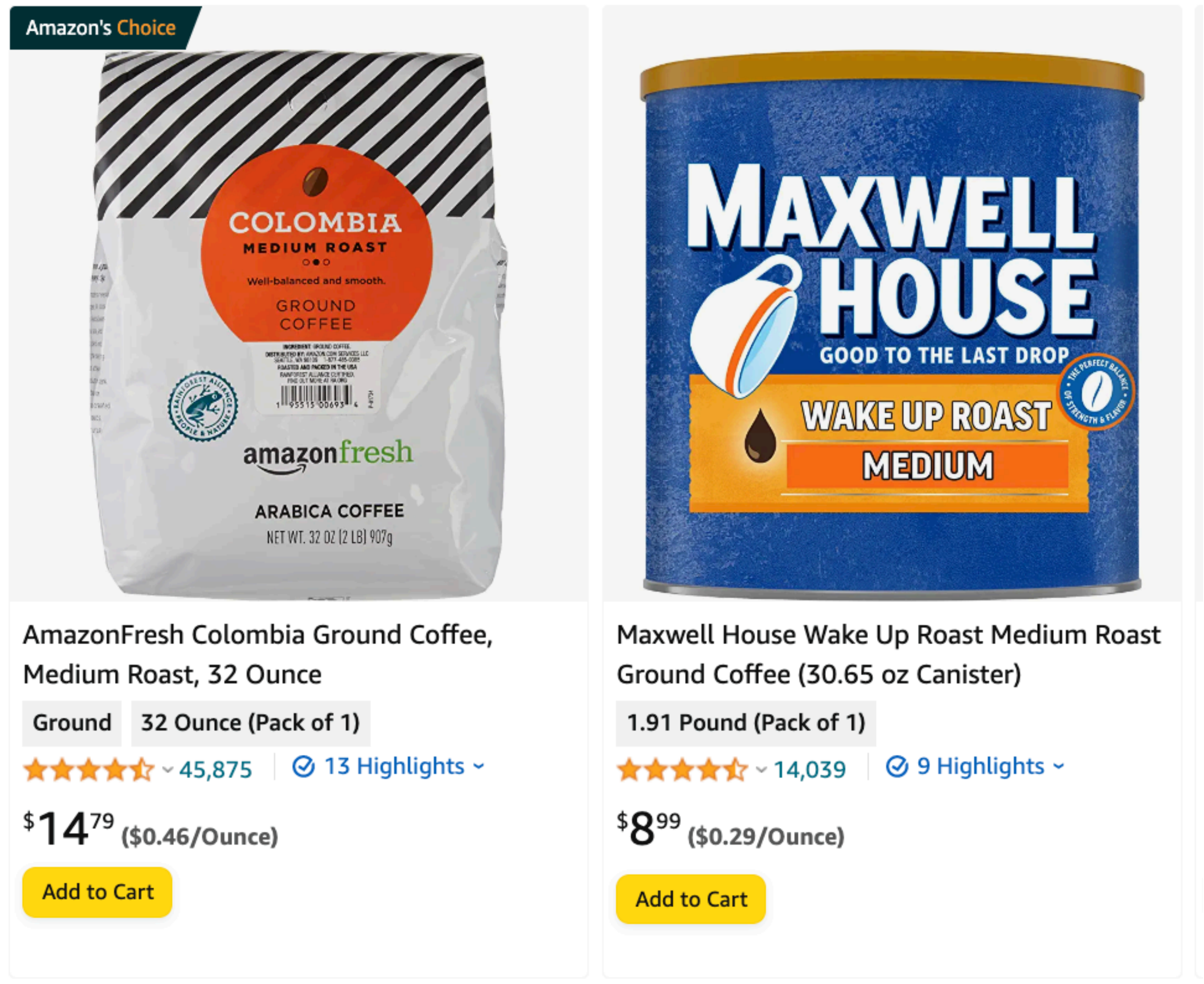
Optimize Promotions Fixing Prices

Simplex Constraint: $X = \Delta^K := \{\mathbf{x} \in \mathbb{R}^K_{\geq 0} : \sum_{i=1}^{K} x_i = 1\}$



Choosing amount of budget each item gets:

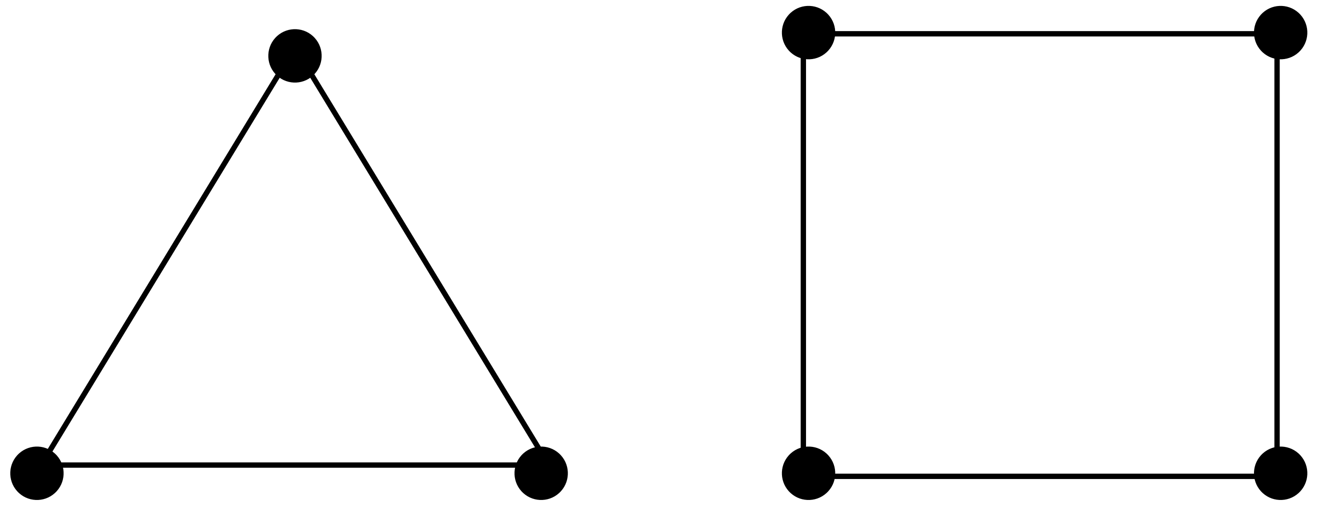
- e.g. some items get a larger amount of screen space



Optimize Promotions Fixing Prices

Optimal Promotion Lemma. The optimal promotion is a vertex of X:

1. $X = \Delta^K$: $\mathbf{x}_\star \in \{e_1, \cdots, e_K\}$
2. $X = [0,1]^K$: $\mathbf{x}_\star \in \{0,1\}^K$



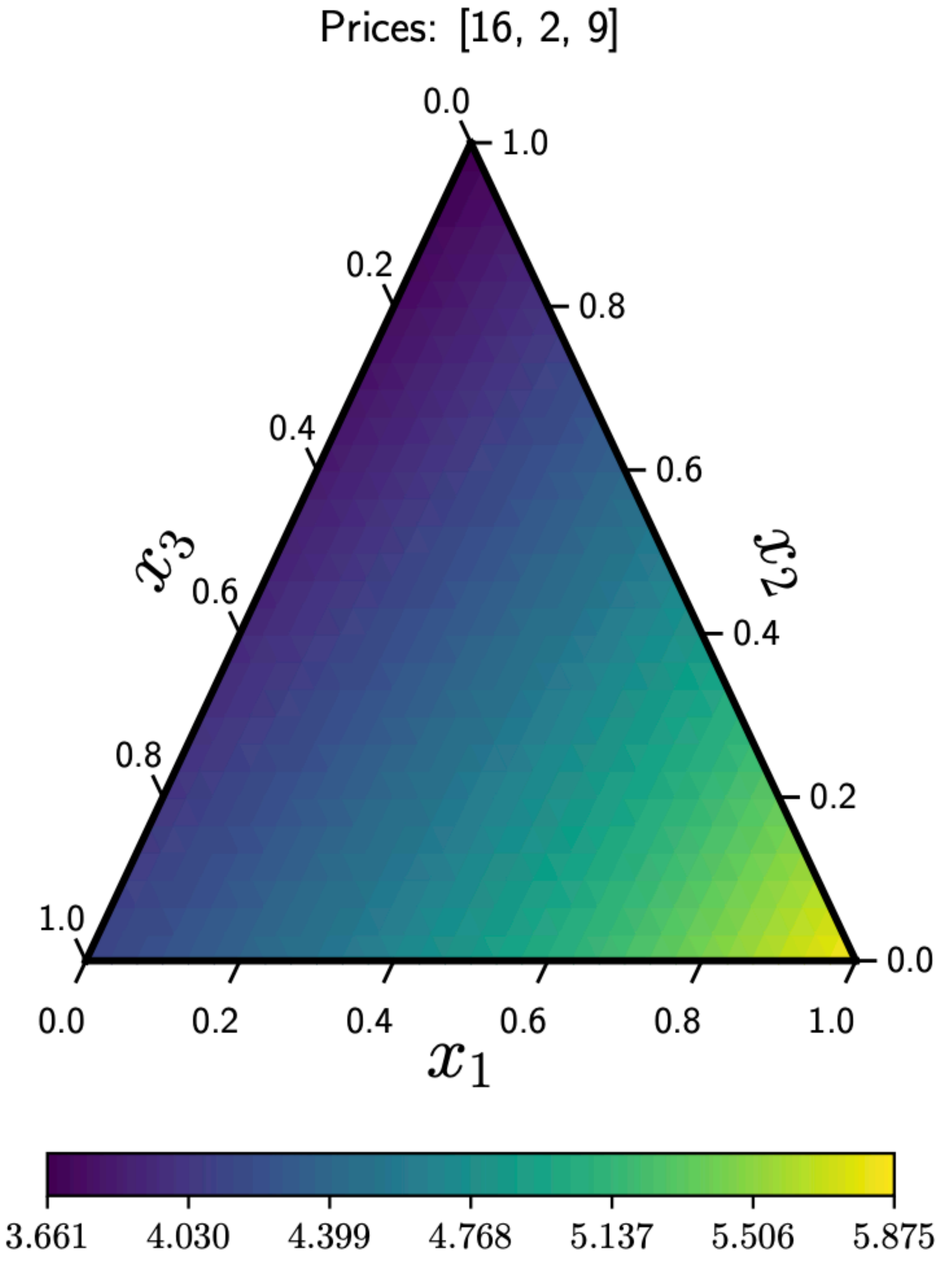
The optimal marketing mix is an all or nothing strategy!

Optimal Promotion: Intuition

$$R_\theta(\mathbf{p}, \mathbf{x}) = \sum_{k=1}^{K} (p_k - m_k) \frac{e^{\alpha_k - \beta_k p_k + \gamma_k x_k}}{1 + \sum_{j=1}^{K} e^{\alpha_j - \beta_j}}$$

However, optimal promotion may not always align with the highest price item!

	Parameters	Product 1	Product 2	Product 3
	α	1	1	1
	β	.1	.2	.3
	γ	.8	.3	.5
Case 1	$\mathbf{p}$	\$16	\$2	\$9
	$\mathbf{x}_*$	1	0	0



Challenges

1. Optimizing over p and x at the same time is non-convex and high dimensional

$$\mathbf{p_t}, \mathbf{x_t} = \arg\max_{\mathbf{p},\mathbf{x}} R_{\theta_t}(\mathbf{p}, \mathbf{x})$$

Solution: Can assume X is finite, find optimal price for each $x \in$ X

2. Posterior Computation Π_t

Posterior Computation

Likelihood Function:

$$\mathbf{L}(\theta) := \mathbf{L}(\theta \,|\, \{(\mathbf{p}_s, \mathbf{x}_s, I_s)\}_{s=1}^t) = \prod_{s=1}^t Q_{I_s}(\mathbf{p}_s, \mathbf{x}_s)$$

Posterior Distribution:

$$p(\theta \,|\, \{(\mathbf{p}_s, \mathbf{x}_s, I_s)\}_{s=1}^t) \propto \mathbf{L}(\theta)\Pi_0(\theta)$$

$$\propto \exp(\log \mathbf{L}(\theta) + \log \Pi_0(\theta))$$

Posterior Computation

Langevin Dynamics:

for $r = 1,2,\cdots,R$

$$\theta_{r+1,t} = \theta_{r,t} + \epsilon_t \nabla_\theta[\log \mathbf{L}(\theta_{r,t}) + \log \Pi_0] + \sqrt{2\epsilon_t}\eta_r$$

$\eta_k \sim N(0,I)$

Posterior Sampling:

$$\text{Langevin}(\theta_{r,t}) \overset{R\to\infty}{\Rightarrow} \exp(\log \mathbf{L}(\theta) + \log \Pi_0(\theta))$$

↓

MCMC method which converges to posterior sampling [WellingYeh'15]

Generally take $\epsilon_t = O(1/t)$

Posterior Computation

Langevin Dynamics:

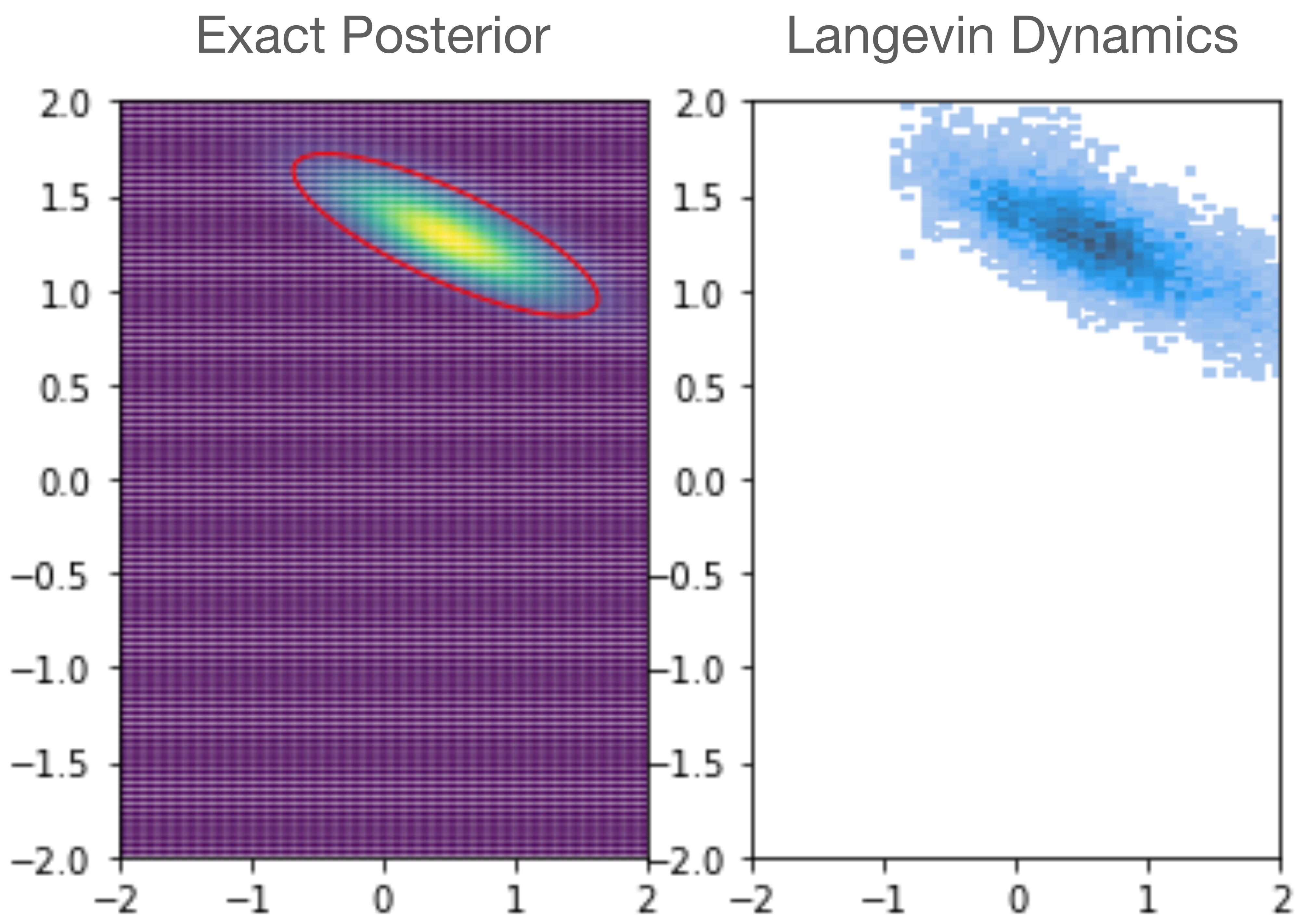
for $r = 1,2,\cdots,R$

$$\theta_{r+1,t} = \theta_{r,t} + \epsilon_t \nabla_\theta[\log \mathbf{L}(\theta_{r,t}) + \log \Pi_0] + \sqrt{2\epsilon_t}\eta_r$$

$$\eta_k \sim N(0,I)$$

Example of 1 product, $\alpha_1 = 1, \beta_1 = 1.25$

1. Very fast updates in PyTorch
2. Take $\epsilon_k = O(1/k)$
3. Need a few dozen steps each iteration



Challenges

1. Optimizing over p and x at the same time is non-convex and high dimensional

$$\mathbf{p_t}, \mathbf{x_t} = \arg\max_{\mathbf{p},\mathbf{x}} R_{\theta_t}(\mathbf{p}, \mathbf{x})$$

Solution: Can assume X is finite, find optimal price for each $x \in$ X

2. Posterior Computation Π_t

Solution: Langevin Dynamics

Our Approach: Thompson Sampling

Input: K products, X promotion set

Initialize: Π_0 as some prior distribution over θ

For $t = 1,2,3,\cdots,T$:

1. **Sample** $\theta_t \sim \Pi_t$

   for $r = 1,2,\cdots,R$

   Sample $\eta_r \sim N(0,I)$

   $$\theta_{r+1,t} = \theta_{r,t} + \epsilon_t \nabla_\theta \log \mathbf{L}(\theta_{r,t}) + \sqrt{2\epsilon_t}\eta_r$$

2. **Set** best price/promotion for θ_t:

   for $x \in$ X

   find $\mathbf{p} = \arg\max_{\mathbf{p}} R_{\theta_{t,R}}(\mathbf{p}, \mathbf{x})$, take highest

4. **Observe** $I_t \sim Q_t(\mathbf{p}_t, \mathbf{x}_t)$, collect revenue p_{t,I_t}

5. **Update** $\Pi_{t+1} =$ Posterior(Π_t, θ_{t+1})

Regret Guarantees

Theorem: [**J**LMY] The Bayesian regret of the Thompson Sampling Procedure after a time horizon of T steps is

$$\approx K\sqrt{\kappa T}$$

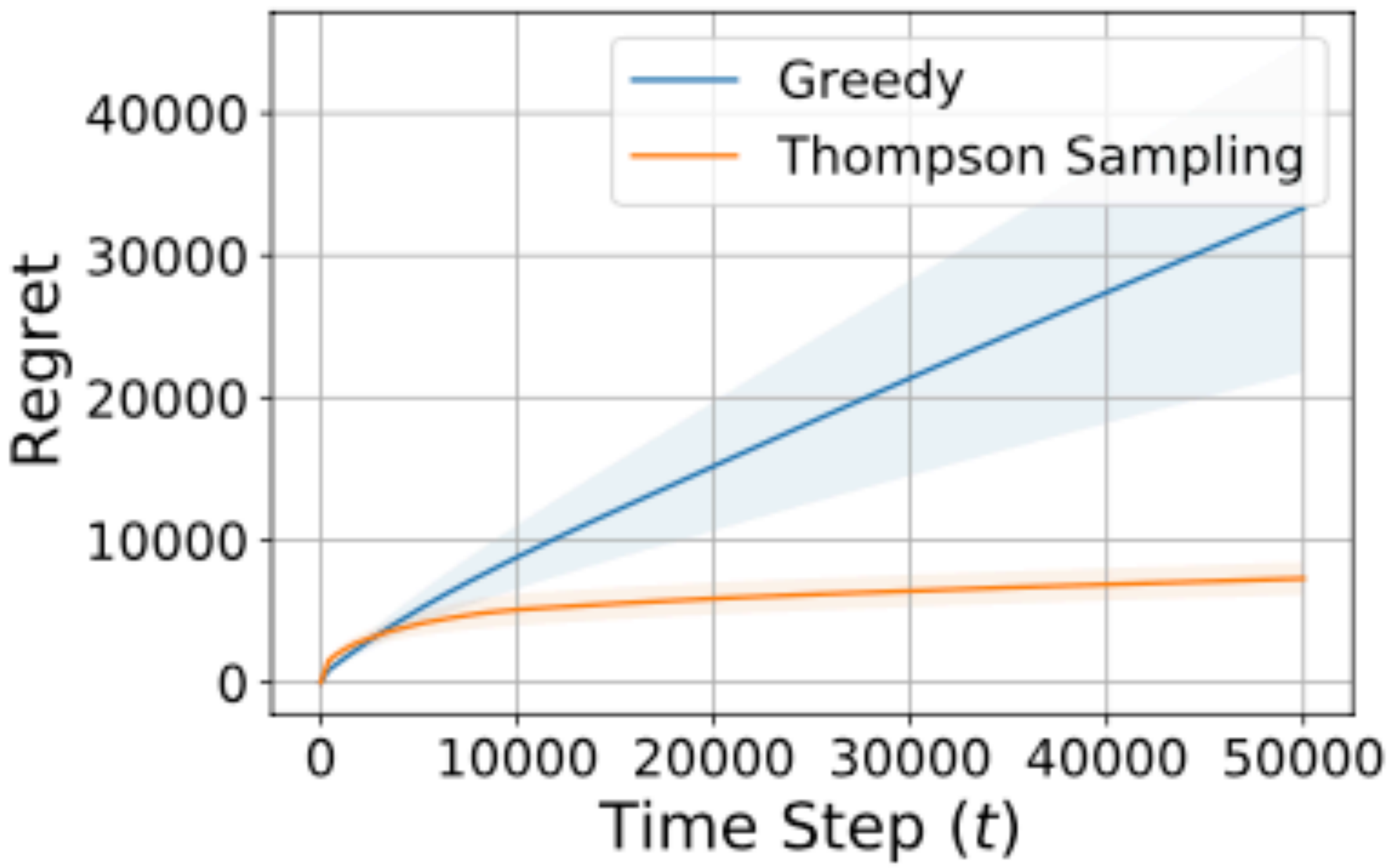
$$\kappa = \frac{1}{\min_{\mathbf{p}\in[\ell,u]^K, x\in\mathcal{X}} \dot{Q}(\mathbf{p}, \mathbf{x})}$$

Empirical Example

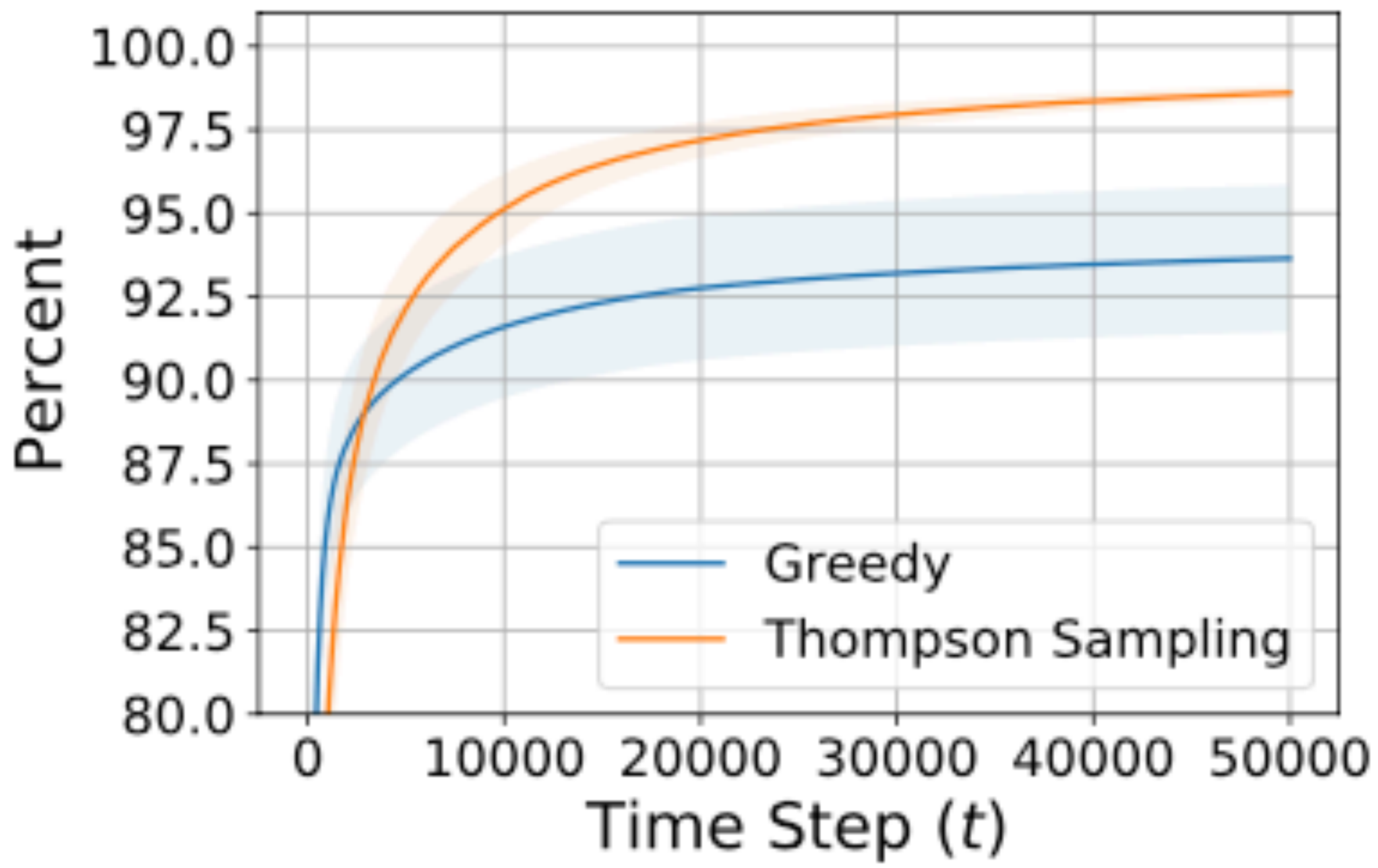
- **Greedy:** Solve the MLE at each time and play the optimal price
- **Thompson Sampling:** Implemented using Langevin Dynamics

Parameters	Product 1	Product 2	Product 3
α	1	1	1
β	.1	.2	.3
γ	.8	.3	.5
$\mathbf{p}_*$	\$20.50	\$15.50	\$13.83
$\mathbf{x}_*$	1	0	0

Cumulative Regret



Percent of Revenue Recovered

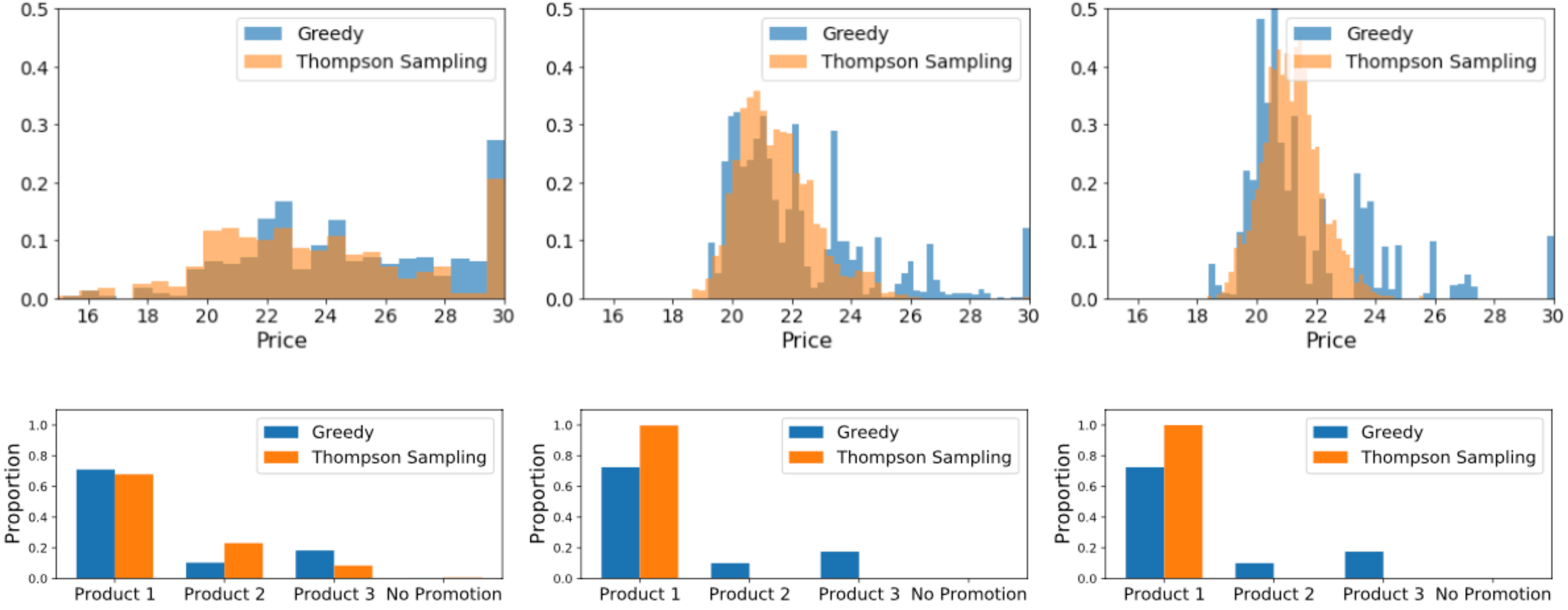


5% gap over all time!

40 MC repeats

What's going on?

$T \leq 2000$ | $10{,}000 \leq T \leq 20{,}000$ | $40{,}000 \leq T \leq 50{,}000$



Outline

1. Introduction

2. Adaptively Setting Prices and Promotions

3. Incorporating Context

Incorporating Heterogeneity

- **Demand may depend on information about customers**
  - Past purchases, location, device, etc.
- **Demand may depend on environmental conditions**
  - Time of year, location
- **Demand may depend on other firms actions**
  - Promotions and Prices of other firms

Adaptive Pricing: Protocol

For $t = 1,2,3,\cdots,T$

Customer arrives at platform with context $\mathbf{c}_t \in \mathbb{R}^d$

Firm chooses price $\mathbf{p}_t \in [\ell, u]^K, \mathbf{x}_t \in \mathrm{X}$

Observe purchase decision $I_t \in \{0,1,\cdots,K\}$, and collect revenue p_{I_t}

Incorporating Heterogeneity: Model

- Utility parameters depend on the context

$$u_k(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k$$

$$\alpha_k : \mathbb{R}^d \to \mathbb{R}$$
$$\beta_k : \mathbb{R}^d \to \mathbb{R}$$
$$\gamma_k : \mathbb{R}^d \to \mathbb{R}$$

$$R_\theta(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \sum_{k=1}^{K} p_k \frac{e^{\alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k}}{1 + \sum_{j=1}^{K} e^{\alpha_j(\mathbf{c}) - \beta_j(\mathbf{c})p_j + \gamma_j(\mathbf{c})x_j}}$$

- Captures user level elasticities in an economically motivated model.
- Easily compute the maximal price and promotion using previous methods.
- Can estimate using MLE

Incorporating Heterogeneity: Linear Case

$$u_{ik}(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k$$

$$\alpha_k(\mathbf{c}) = \langle \alpha_k, \mathbf{c} \rangle$$

$$\beta_k(\mathbf{c}) = \langle \beta_k, \mathbf{c} \rangle$$

$$\gamma_k(\mathbf{c}) = \langle \gamma_k, \mathbf{c} \rangle$$

$$\alpha_k, \beta_k, \gamma_k \in \mathbb{R}^d$$

for all $1 \leq k \leq K$

Incorporating Heterogeneity: Linear Case

$$u_{ik}(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k$$

$$\alpha_k(\mathbf{c}) = \langle \alpha_k, \mathbf{c} \rangle$$

$$\beta_k(\mathbf{c}) = \langle \beta_k, \mathbf{c} \rangle$$

$$\gamma_k(\mathbf{c}) = \langle \gamma_k, \mathbf{c} \rangle$$

$$\alpha_k, \beta_k, \gamma_k \in \mathbb{R}^d$$

for all $1 \leq k \leq K$

- **Ban+Keskin '21, Javanmard+Nazerzedah '20, Javanmard+Nazerzedah+Shao '21, Qiang+Bayati '16, Dube+Misra '17**
- Forced exploration methods tend to consider the linear Gaussian case
- Require the context distribution to be *fixed and stochastic*

Thompson Sampling

Input: K products, X promotion set

Initialize: Π_0 as some prior distribution over θ

For $t = 1,2,3,\cdots,T$:

1. Receive $\mathbf{c}_t \in \mathbb{R}^d$

2. **Sample** $\theta_t \sim \Pi_t$

   for $r = 1,2,\cdots,R$

   Sample $\eta_r \sim N(0,I)$

   $$\theta_{r+1,t} = \theta_{r,t} + \epsilon_t \nabla_\theta \log \mathbf{L}(\theta_{r,t}) + \sqrt{2\epsilon_t}\eta_r$$

3. **Set** best price/promotion for θ_t:

   for $x \in$ X

   find $\mathbf{p} = \arg\max_{\mathbf{p}} R_{\theta_{t,R}}(\mathbf{p}, \mathbf{x}, \mathbf{c_t})$, take highest

5. **Observe** $I_t \sim Q_t(\mathbf{p}_t, \mathbf{x}_t, \mathbf{c}_t)$, collect revenue p_{t,I_t}

6. **Update** $\Pi_{t+1} = \text{Posterior}(\Pi_t, \theta_{t+1})$

Experiment: Real Life Setting

Considered two large supermarket on the category of ground coffee

- Have access to price/oz of 9 different brands
- Considered a year of data with market share aggregated weekly
- Have price and promotion variables for each brand at a weekly level
- **Fit** a choice model using Berry Inversion
- **Simulated** using this data

NielsenIQ Nielsen

Experiment: Simulation

Store Dummy Variables

$$c = (Q_1, Q_2, Q_3, Q_4, S1, S2)$$

Quarter Dummy Variables

- 40000 Purchase Decisions
- Split into Four Quarters
- 65% chance of store 1, 35% from store 2
- Compared Thompson Sampling, Greedy, M3P (Forced Exploration)

Experiments: Estimation

Utility Equation

$$u_i(\mathbf{p}_t, \mathbf{x}_t, \mathbf{c}_t) = \alpha_i(c_t) - \beta_i(c_t) + \gamma_i(c_t)x_{it}$$

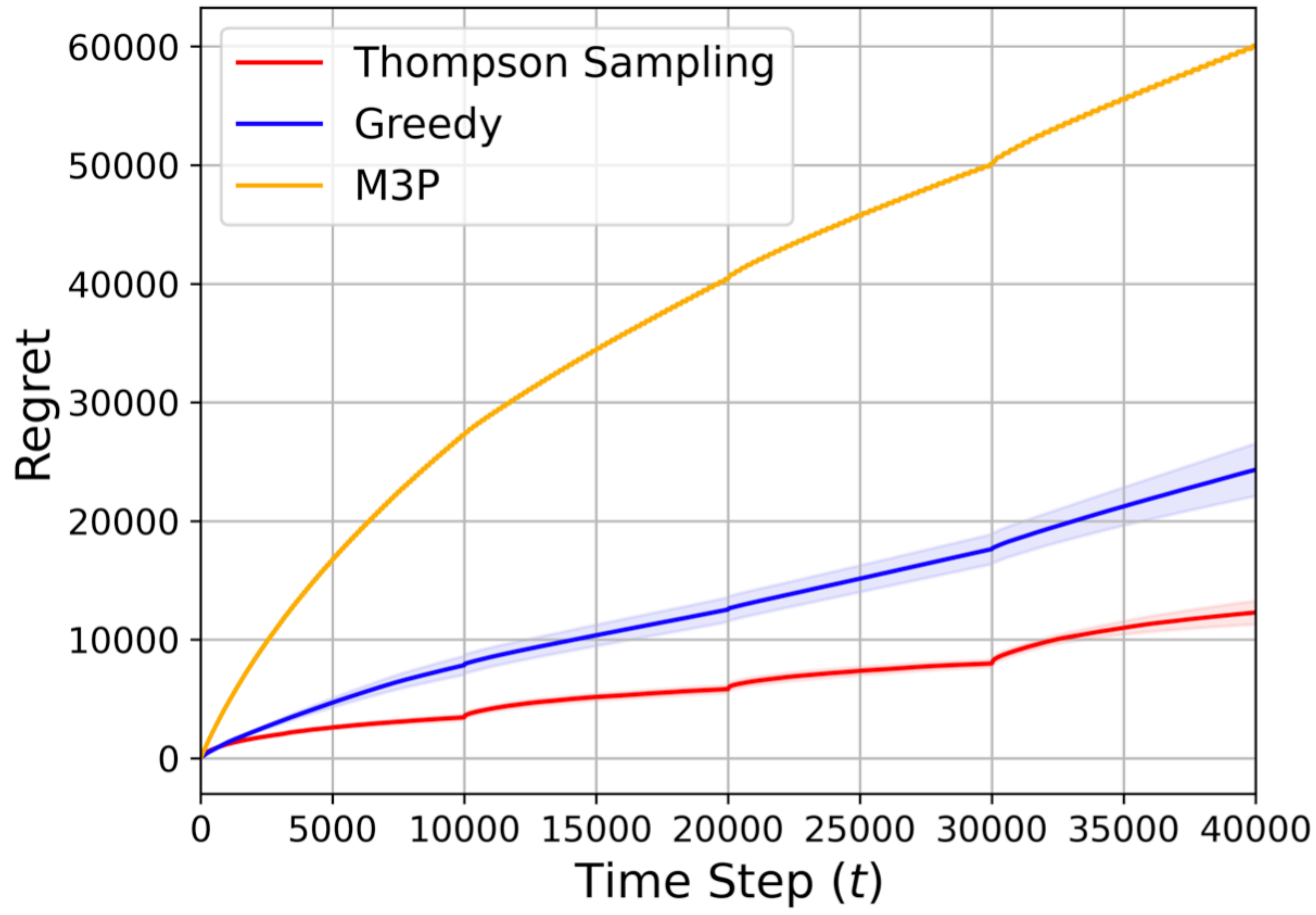
Where

$$\alpha_i(\mathbf{c}) = \alpha_{iQ_1}\mathbb{I}(Q_1) + \alpha_{iQ_2}\mathbb{I}(Q_2) + \alpha_{iQ_3}\mathbb{I}(Q_3) + \alpha_{iQ_4}\mathbb{I}(Q_4) + \alpha_{iS_1}\mathbb{I}(store = 1) + \alpha_{iS_2}\mathbb{I}(store = 2)$$

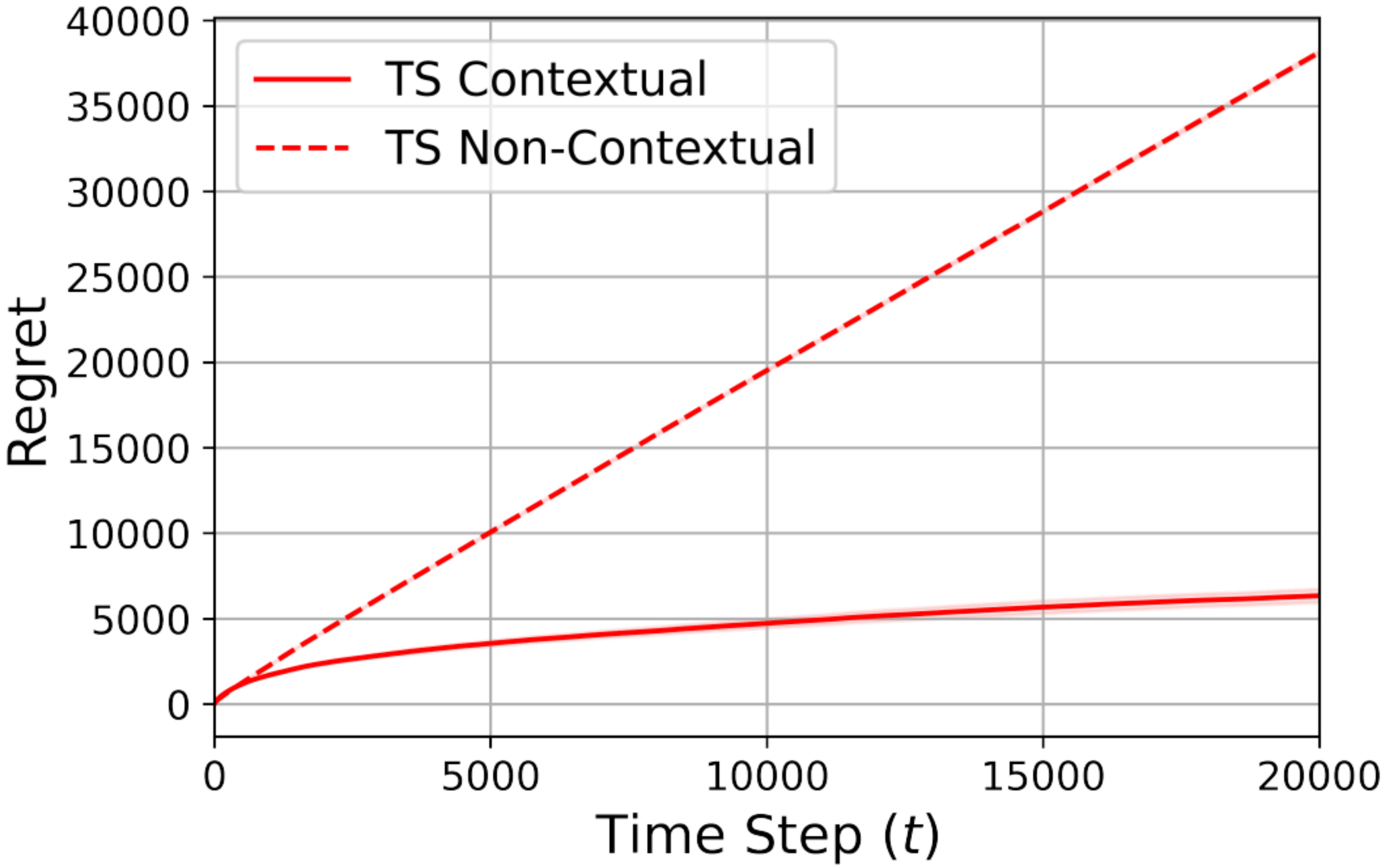
$$\beta_i(\mathbf{c}) = \beta_{iQ_1}\mathbb{I}(Q_1) + \beta_{iQ_2}\mathbb{I}(Q_2) + \beta_{iQ_3}\mathbb{I}(Q_3) + \beta_{iQ_4}\mathbb{I}(Q_4) + \beta_{iS_1}\mathbb{I}(store = 1) + \beta_{iS_2}\mathbb{I}(store = 2)$$

$$\gamma_i(\mathbf{c}) = \gamma_{iQ_1}\mathbb{I}(Q_1) + \gamma_{iQ_2}\mathbb{I}(Q_2) + \gamma_{iQ_3}\mathbb{I}(Q_3) + \gamma_{iQ_4}\mathbb{I}(Q_4) + \gamma_{iS_1}\mathbb{I}(store = 1) + \gamma_{iS_2}\mathbb{I}(store = 2).$$

Experiment: Simulation Results



Model Misspecification



Regret Guarantees in Linear Setting

Theorem: [**J**LMY] The Bayesian regret of the Thompson Sampling Procedure after a time horizon of T steps is

$$\approx Kd\sqrt{\kappa T}$$

With no assumption on the context distribution!

Nonlinear Pricing Experiments

$$u_{ik}(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k$$

What if $\alpha_k, \beta_k, \gamma_k$ are non-linear functions of the context $\mathbf{c}$?

- Gradient Boosted Trees
- Neural Networks
- Gaussian Process Methods

Nonlinear Pricing Experiments

$$u_{ik}(\mathbf{p}, \mathbf{x}, \mathbf{c}) = \alpha_k(\mathbf{c}) - \beta_k(\mathbf{c})p_k + \gamma_k(\mathbf{c})x_k$$

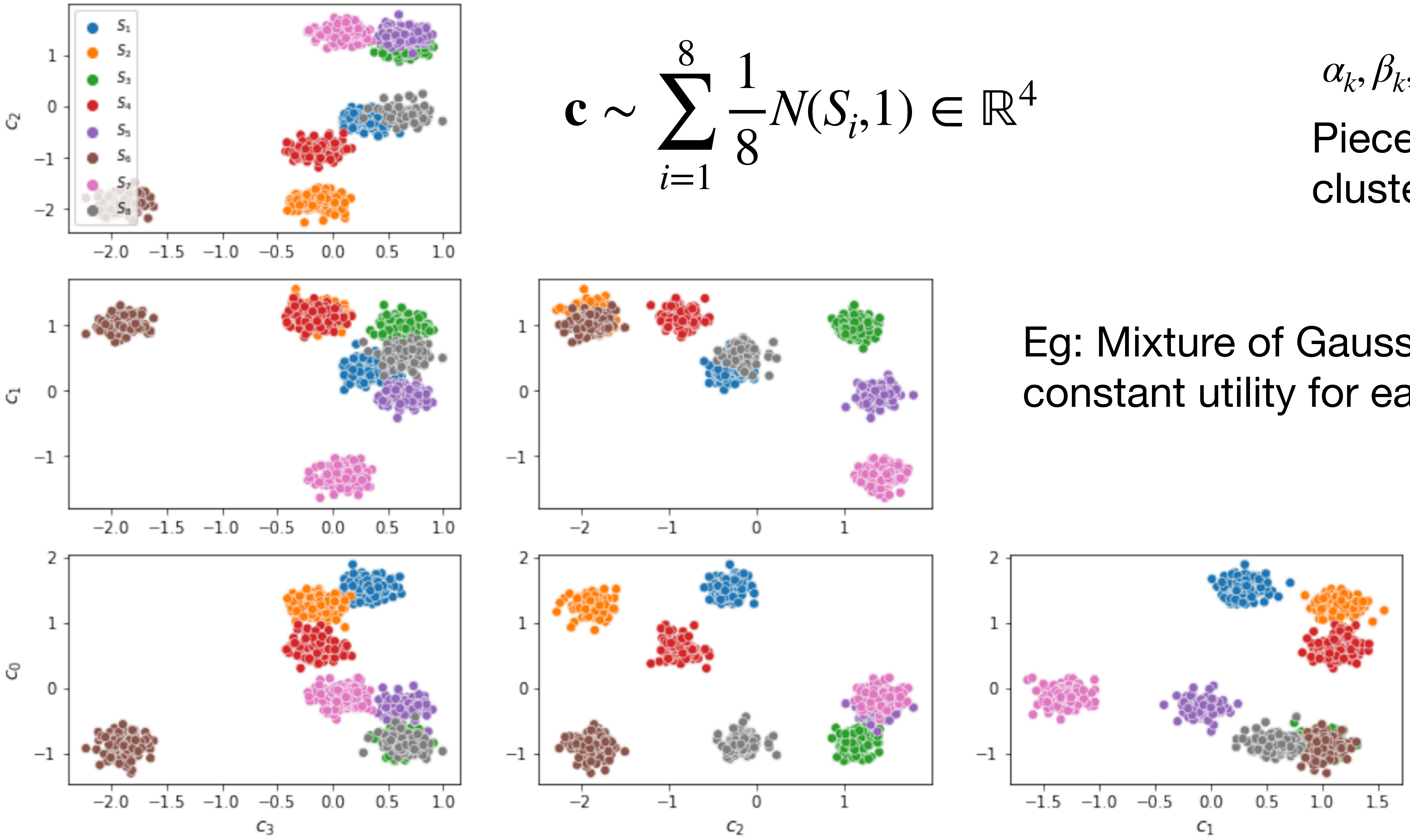
What if $\alpha_k, \beta_k, \gamma_k$ are non-linear functions of the context $\mathbf{c}$?

- Gradient Boosted Trees
- Neural Networks
- Gaussian Process Methods

How do we adopt Posterior sampling to more general classes?

Answer: Langevin Dynamics as Deep Bayesian Posterior Approximation

Experiment: Clustered Customer Preferences



Context Distribution

$$\mathbf{c} \sim \sum_{i=1}^{8} \frac{1}{8} N(S_i, 1) \in \mathbb{R}^4$$

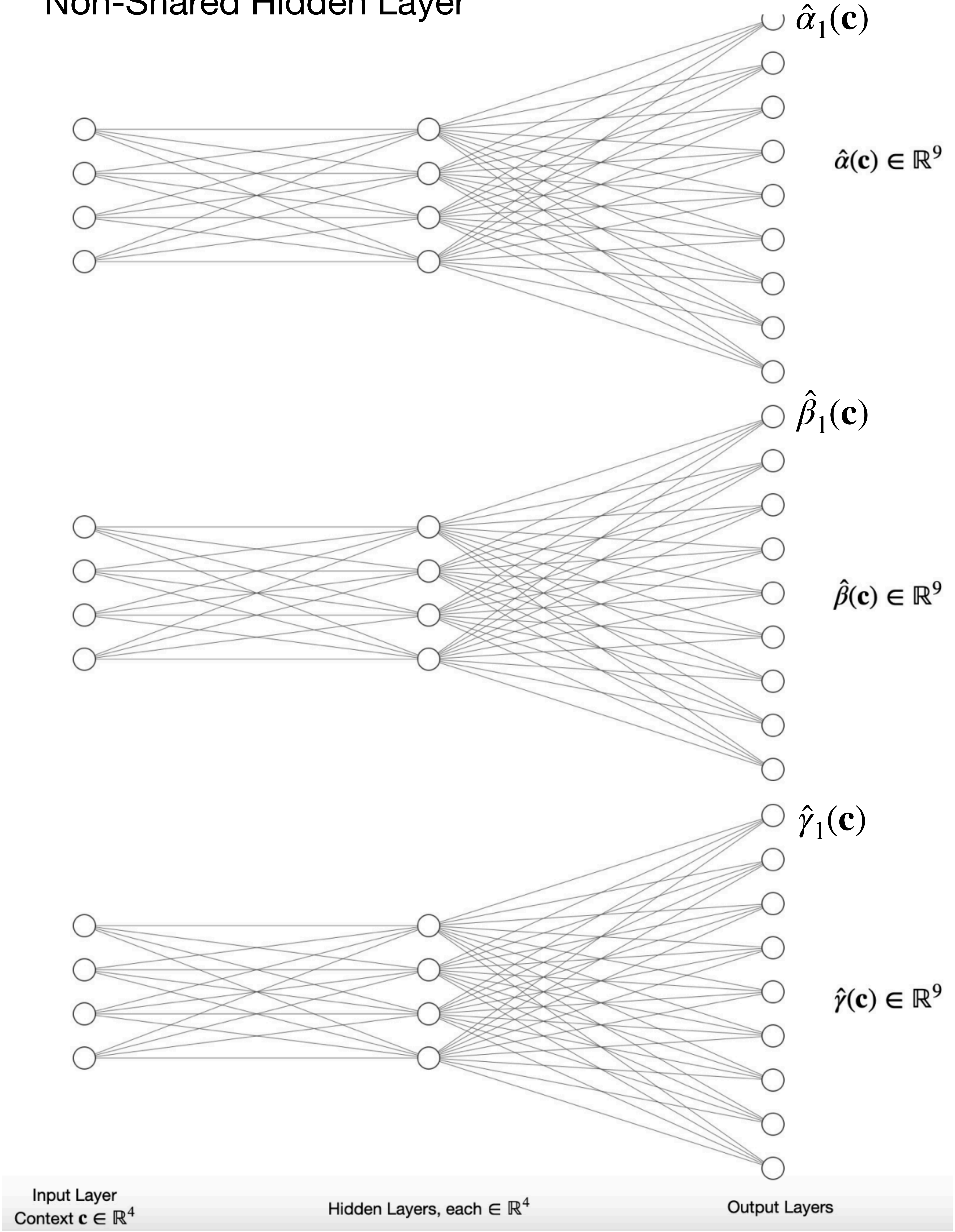
Utility Parameters

$$\alpha_k, \beta_k, \gamma_k : \mathbb{R}^4 \to \mathbb{R}, k \leq 9$$

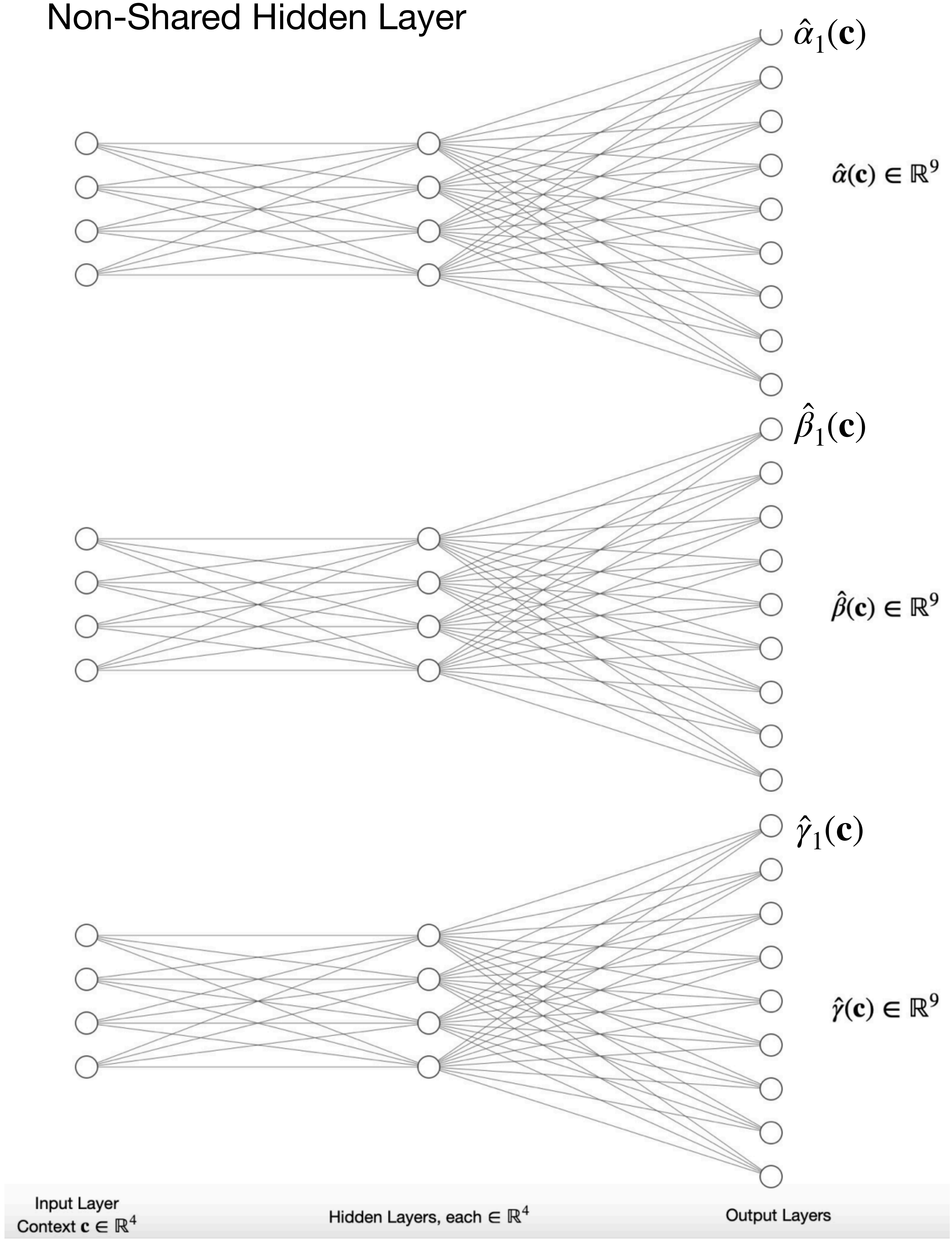
Piecewise constant on each cluster - 27 parameters/cluster

Eg: Mixture of Gaussians, with piecewise constant utility for each cluster

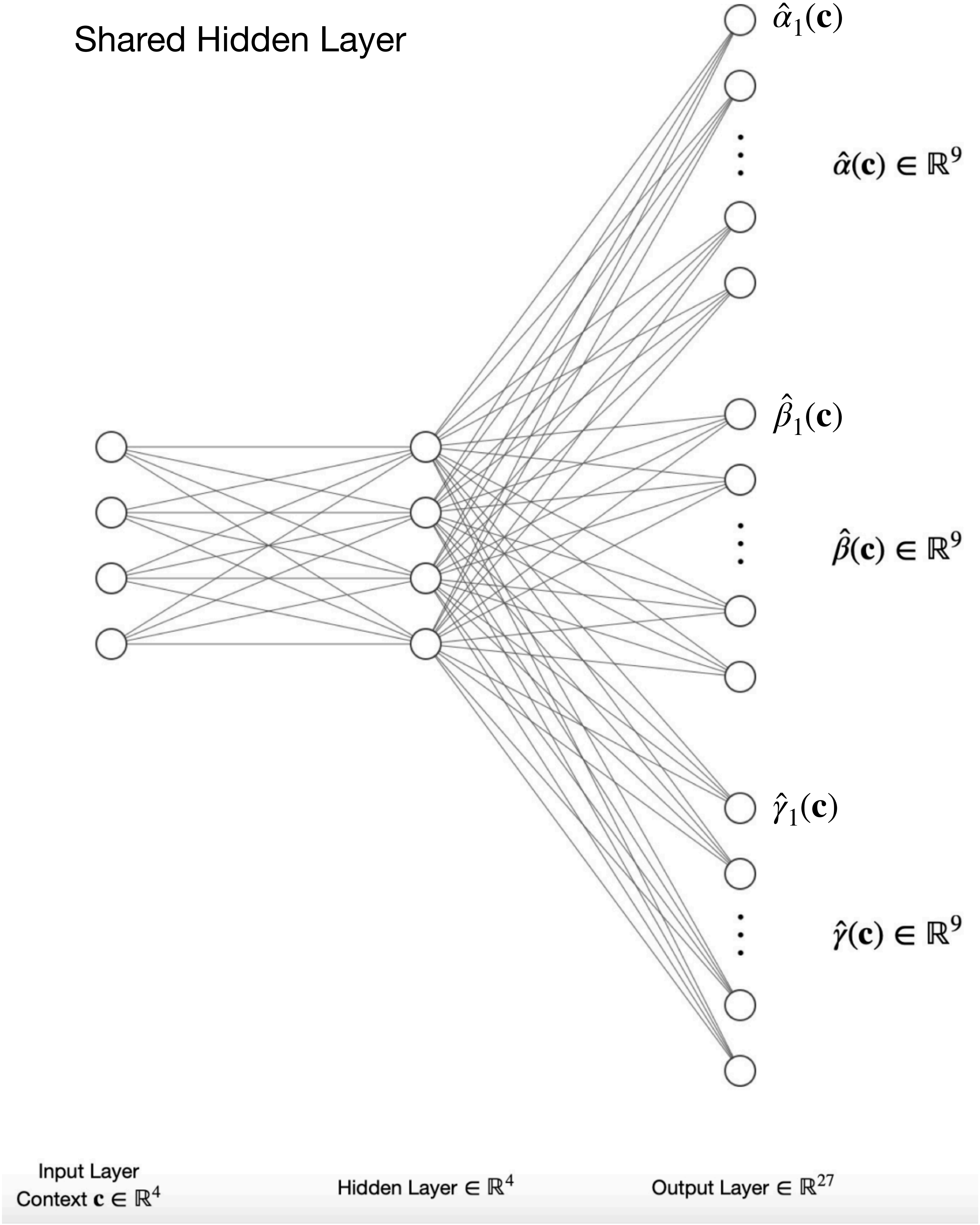
Non-Shared Hidden Layer



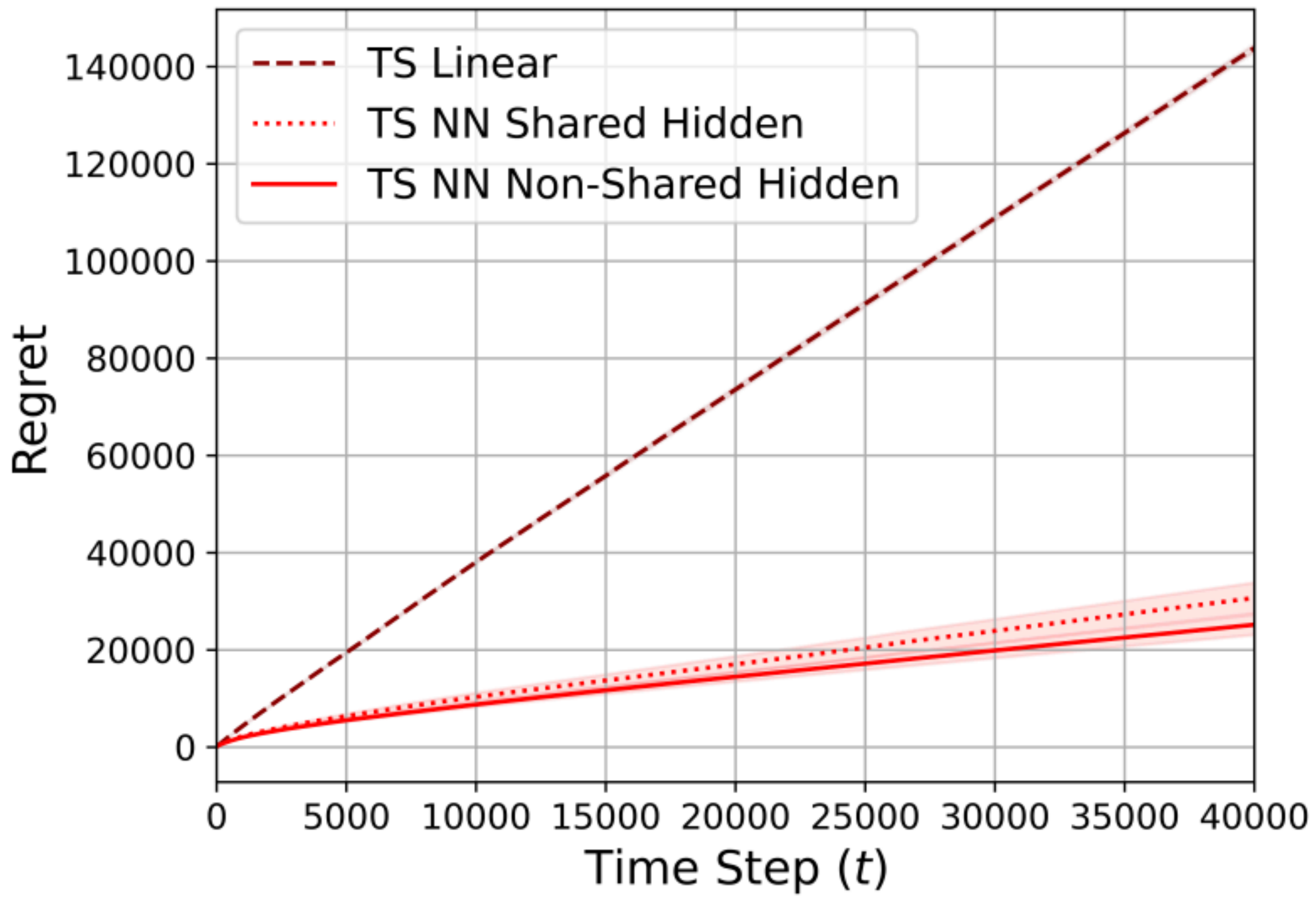
Non-Shared Hidden Layer

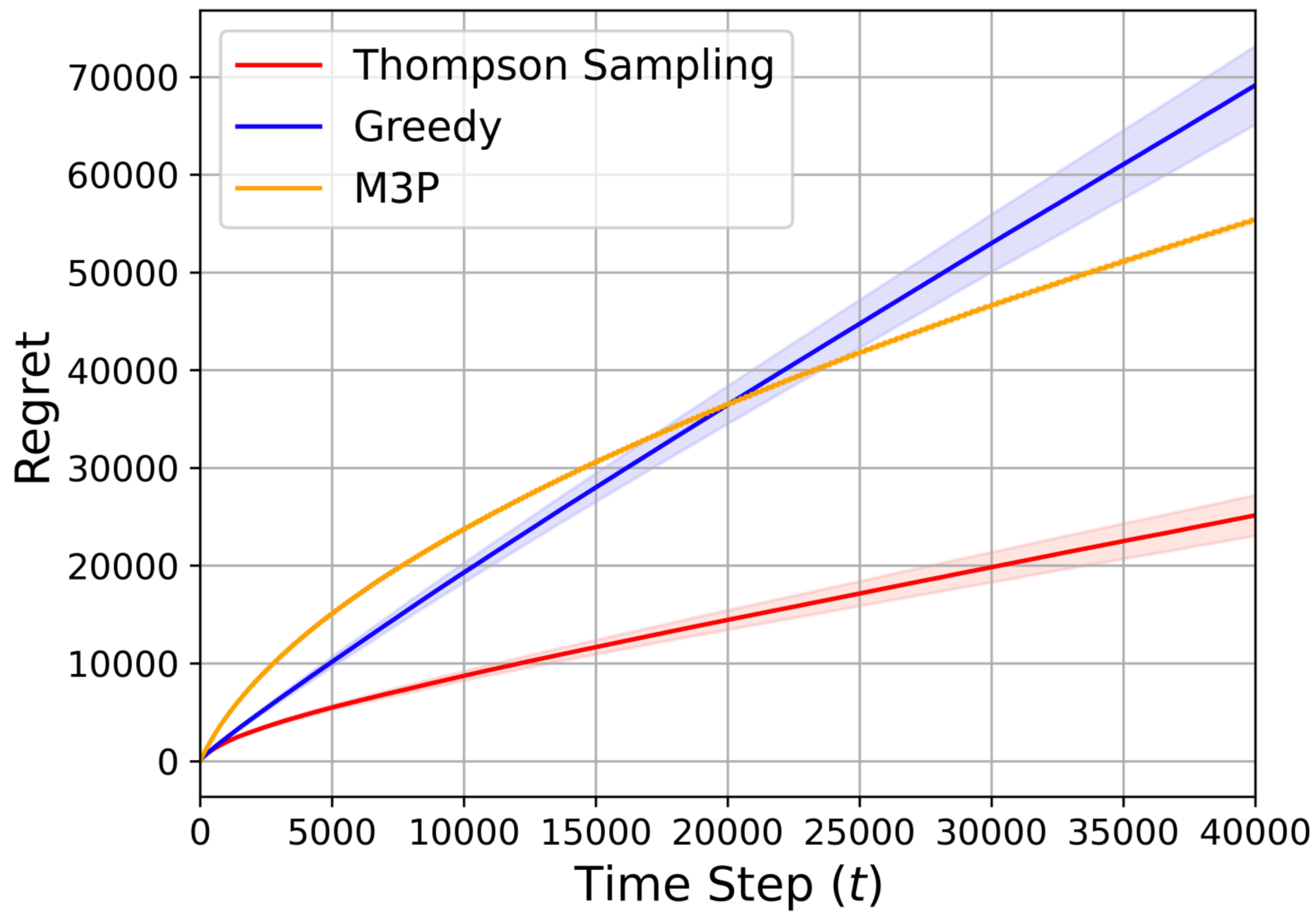


Shared Hidden Layer



Experiment: Results





Summary

- Introduced the new setting of adaptive pricing with promotions
- Bounded the regret of a Thompson Sampling procedure.
- Extended to settings with context and non-linear utility.
- Demonstrated the viability of this methodology on real-life inspired datasets.

Thank you!! Questions?