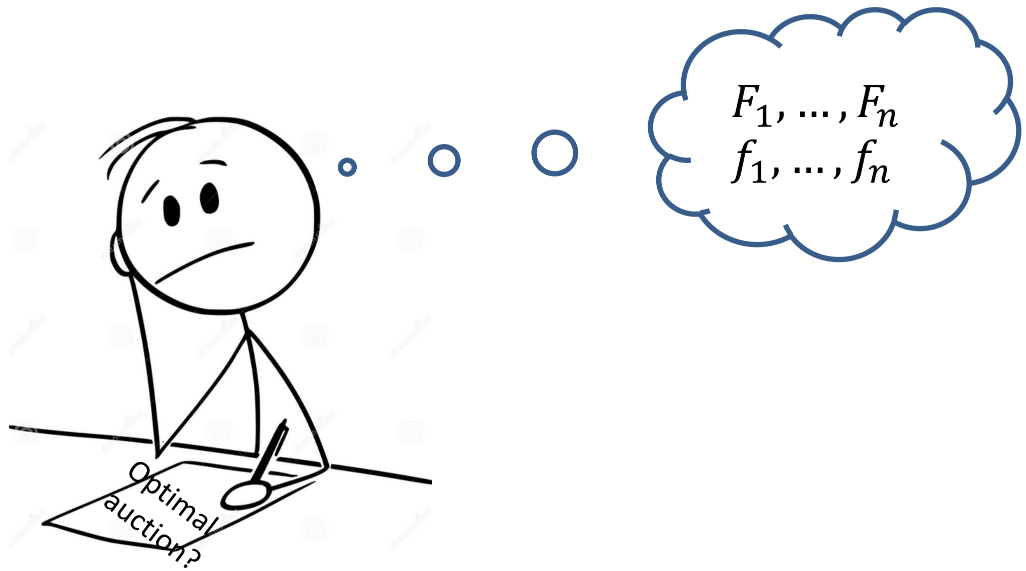
Learning Across and Within Instances for Mechanism Design

Siddharth Prasad

Joint work with Nina Balcan and Tuomas Sandholm

Background and motivation

- Classical mechanism design
  - 1981: Myerson showed how to sell a single item to maximize revenue (uses details of the distribution of buyers’ values for the item)
  - Today: don’t know how to sell two items optimally

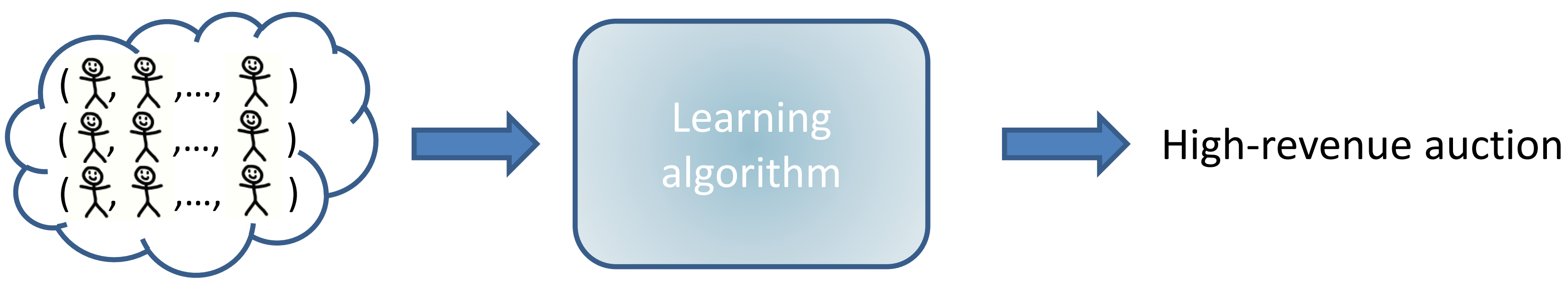


Background and motivation

- Automated mechanism design [Conitzer and Sandholm UAI '03]:
  - why struggle with the hard economics problem of designing explicit mechanisms when a computer program can do it for you?
  - requiring details of distributions -> computational hardness in many settings

Background and motivation

- Sample-based automated mechanism design
  - use machine learning, don’t need details about distribution [Sandholm and Likhodedov AAAI ‘04, ’05, Operations Research ’15, Mohri and Medina ICML ‘14, Morgenstern and Roughgarden NIPS ‘15, Balcan, Sandholm, and Vitercik EC’18, Duetting, Feng, Narasimhan, Parkes, Ravindranath ICML’19]



Outline

Part 1

Efficiently learning high-revenue two-part tariffs from samples/across instances.

Part 2a

Seller faces a single instance of buyers. Can he "learn within an instance" to find a high-revenue auction?

Part 2b

Seller faces a fixed population of buyers. Can he learn an auction that extracts high revenue from a shrinking market?

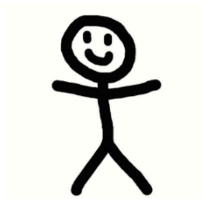
Part 1
Efficiently learning high-revenue two-part tariffs from samples

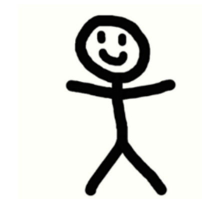
Two-part tariffs

Up-front fee p_1 + #units · (per-unit fee p_2)

How to choose prices?

- Seller has K units to sell, wants to maximize revenue
- Buyer has values $v(1), v(2), \dots, v(K)$





Seller: sets two-part tariff (p_1, p_2)

Buyer: purchases q units maximizing

$$v(q) - (p_1 + q \cdot p_2)$$

Seller's revenue: $p_1 + q \cdot p_2$

Challenge: Seller does not see $v(1), \dots, v(K)$. Buyer's values are drawn from an unknown distribution from which seller sees samples.

Goal: efficiently find (p_1, p_2) that yields close-to-optimal expected revenue.

Empirical revenue maximization

Seller sees N IID samples from unknown distribution D.

Chooses (p_1, p_2) to maximize empirical revenue:

$v^1(1), \dots, v^1(K) \sim D$

$\vdots$

$v^N(1), \dots, v^N(K) \sim D$

$$\widehat{Rev}_S(\boldsymbol{p}) = \frac{1}{N}\sum_{i=1}^{N} Rev_{v^i}(\boldsymbol{p})$$

[Balcan, Sandholm, Vitercik EC'18] $N = O(\log K)$ samples suffice to guarantee that with high probability, the empirical-revenue-maximizing two-part tariff is near-optimal on a new buyer.

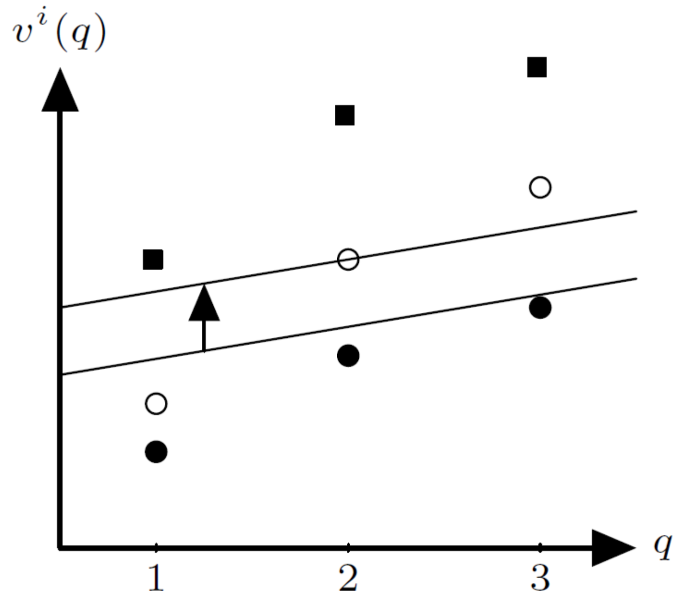
Algorithm for a single two-part tariff

Theorem (Balcan, **P.**, Sandholm)**:** *The empirical-revenue maximizing single tariff over a sample set of size N can be found in $O(N^3K^3)$ time.*

- Space of two-part tariffs (p_1, p_2) is infinite, but two crucial insights reduce search space to N^2K^2 two-part tariffs.

(1) If (p_1, p_2) is a two-part tariff that maximizes empirical revenue over S, the line with y-intercept p_1 and slope p_2 passes through a point $(q, v^i(q))$ for some quantity q and sample i.

This is because a vertical shift of a two-part tariff until it passes through such a point only increases revenue.



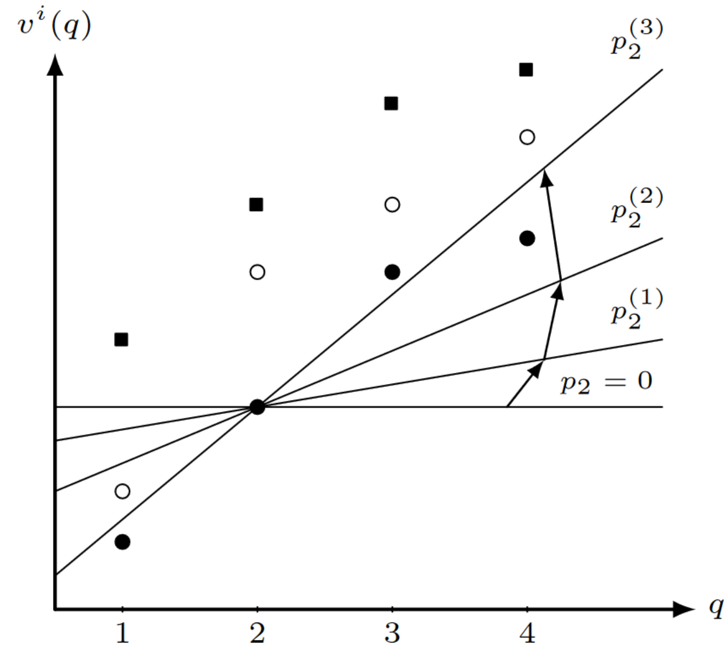
Algorithm for a single two-part tariff (cont.)

Theorem (Balcan, **P.**, Sandholm)**:** *The empirical-revenue maximizing single tariff over a sample set of size N can be found in $O(N^3K^3)$ time.*

- Space of two-part tariffs (p_1, p_2) is infinite, but two crucial insights reduce search space to N^2K^2 two-part tariffs.

(2) Hinge a two-part tariff at $(q, v^i(q))$ and continuously rotate it, starting with slope $p_2 = 0$. Each buyer's most-preferred quantity changes at most K times, with revenue varying linearly between each change.

Can compute the "change" points. NK "change" points for each of the NK "hinge" points -> N^2K^2 two-part tariffs to search over.



Menus of two-part tariffs

- Seller can offer a *menu* $(p_1^1, p_2^1), \ldots, (p_1^L, p_2^L)$ of L two-part tariffs to extract more revenue
- Buyer with values $v(1), \ldots, v(K)$ purchases quantity q of two-part tariff r to maximize

$$v(q) - (p_1^r + q \cdot p_2^r)$$

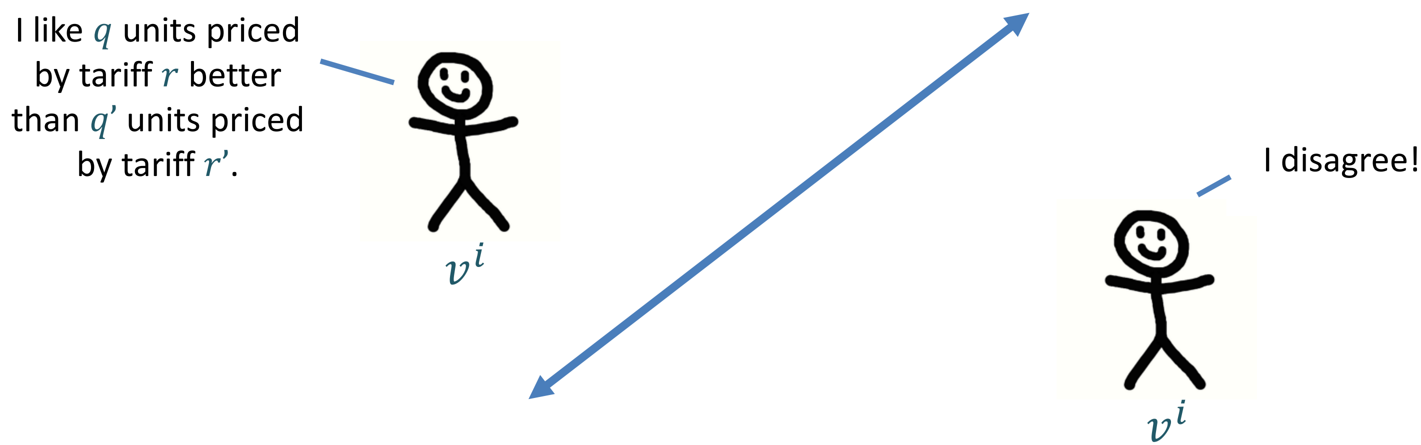
[Balcan, Sandholm, Vitercik EC'18] $N = O(L \log(K))$ samples suffice to guarantee that with high probability, the empirical-revenue-maximizing menu of two-part tariffs is near-optimal on a new buyer.

This work: an $(NK)^{O(L)}$-time algorithm to compute the empirical-revenue-maximizing menu of two-part tariffs.

Algorithm for a menu of two-part tariffs

- $N(KL)^2$ hyperplanes partitioning $\boldsymbol{R}^{2L}$ such that empirical revenue is linear over each region determined by the hyperplanes:

$$v^i(q) - (p_1^r + q \cdot p_2^r) = v^i(q') - \left(p_1^{r'} + q' \cdot p_2^{r'}\right)$$



for each sample $i = 1, \dots, N$, each pair of quantities q, q', and each pair of two-part tariffs r, r'.

Algorithm for a menu of two-part tariffs (cont.)

- $N(KL)^2$ hyperplanes partition $\boldsymbol{R}^{2L}$ into at most $N^{2L}K^{4L}$ convex polytopes.
- The empirical-revenue maximizing menu of two-part tariffs *within a single region* is the solution to a linear program with $2L$ variables and $N(KL)^2$ constraints.
- There is a simple $poly(|H|)$-time algorithm that outputs representations of each convex polytope as a set of LP constraints.

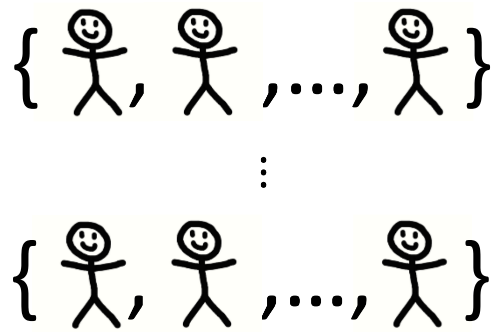
Theorem (Balcan, **P.**, Sandholm): *The empirical-revenue maximizing menu of L two-part tariffs over a sample set of size N can be found in $(NK)^{O(L)}$ time.*

In most applications, L is a small constant, typically 2 or 3.

Multiple buyers

Selling to n buyers.

See N samples, each consisting of n buyers.



$\sim D$

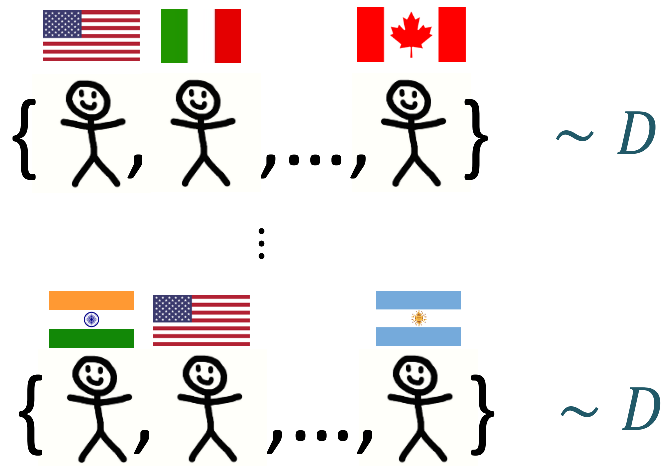
$\sim D$

Both algorithms generalize:

- Optimal single two-part tariff: $O(n^3 N^3 K^3)$-time algorithm.
- Optimal menu of two-part tariffs: $(nNK)^{O(L)}$-time algorithm.

Market segmentation

- Each buyer belongs to one of M markets
- Seller can set M different menus, one for each market



[Balcan, Sandholm, Vitercik ‘18] $N = O(ML\log(nK))$ samples suffice to guarantee that with high probability, the empirical-revenue-maximizing two-part tariff structure is near-optimal on a new buyer.

Market segmentation

- Seller must ensure that the total demand from each market does not exceed capacity K.

Theorem (Balcan, **P.**, Sandholm)**:** *Let* $N \geq O_\varepsilon(ML\log(nK))$*. With high probability over the draw of* N *samples, if* $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_M$ *are two-part tariff menus that are feasible over the samples,*

$$\Pr_{v\sim D}(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_M \text{ are feasible for } v) \geq 1-\varepsilon.$$

Market segmentation (computational hardness)

- Very simple case: buyers have additive valuations, i.e.

$$v(q_1 + q_2) = v(q_1) + v(q_2)$$

- Parameters L and K vanish, seller simply needs to choose a single price for each market.

Theorem (Balcan, **P.**, Sandholm)**:** *Finding the empirically-optimal collection of market prices is NP-hard.*

Market segmentation (IID buyers)

- If each buyer (across markets) is drawn independently from the same distribution, can compute the non-market-segmented solution and reuse it in each market.

- Reusing the non-market-segmented solution is nearly as good as the optimal market-segmented solution (with high probability).

- Can circumvent computational hardness and use either the single-tariff algorithm or the algorithm for a menu of two-part tariffs.

Part 2

Learning within an instance, combinatorial auctions, and shrinking markets

Outline

Part 1

Efficiently learning high-revenue two-part tariffs from samples/across instances.

Part 2a

Seller faces a single instance of buyers. Can he "learn within an instance" to find a high-revenue auction?

Part 2b

Seller faces a fixed population of buyers. Can he learn an auction that extracts high revenue from a shrinking market?

Combinatorial auctions crash course

- Seller has m indivisible items to sell among set S of n bidders.
- Bidders have combinatorial valuations $v_i: 2^{\{1,\ldots,m\}} \to \mathbf{R}_{\geq 0}$.
- For reported valuations $v_1, \ldots, v_n$, an auction M specifies an allocation $\alpha(v_1, \ldots, v_n)$ and payments $p_i(v_1, \ldots, v_n)$.
- Seller wants to design M that extracts high revenue in an incentive compatible manner, that is,

$$v_i\big(\alpha(v_1, \ldots, v_n)\big) - p_i(v_1, \ldots, v_n) \geq v_i\big(\alpha(v_1, \ldots, \hat{v}_i, \ldots, v_n)\big) - p_i(v_1, \ldots, \hat{v}_i, \ldots, v_n)$$

Combinatorial auctions crash course

Vickrey-Clarke-Groves (VCG) auction:

– use allocation α^* that maximizes welfare

$$W(\alpha) = \sum_{i=1}^{n} v_i(\alpha)$$

– bidder i pays

$$\max_{\alpha} \sum_{j \neq i} v_j(\alpha) - \sum_{j \neq i} v_j(\alpha^*)$$

– VCG is incentive compatible

Combinatorial auctions crash course

λ-auction: parameterized by $\lambda \in \mathbf{R}^{(n+1)^m}$

- use allocation α^* that maximizes welfare plus boost

$$\sum_{i=1}^{n} v_i(\alpha) + \lambda(\alpha)$$

- Bidder i pays

$$\max_{\alpha}\left[\sum_{j \neq i} v_j(\alpha) + \lambda(\alpha)\right] - \left[\sum_{j \neq i} v_j(\alpha^*) + \lambda(\alpha^*)\right]$$

- λ-auctions are incentive compatible
- Many other parameterized generalizations of VCG exist

Revenue benchmark

- We compare all our revenue guarantees to the very strong benchmark of

$$W(S) = \max_{\alpha} W(\alpha)$$

- No auction can obtain revenue greater than $W(S)$, since no bidder would pay more than her value for the bundle she receives.

Outline

Part 1

Efficiently learning high-revenue two-part tariffs from samples/across instances.

Part 2a

Seller faces a single instance of buyers. Can he "learn within an instance" to find a high-revenue auction?

Part 2b

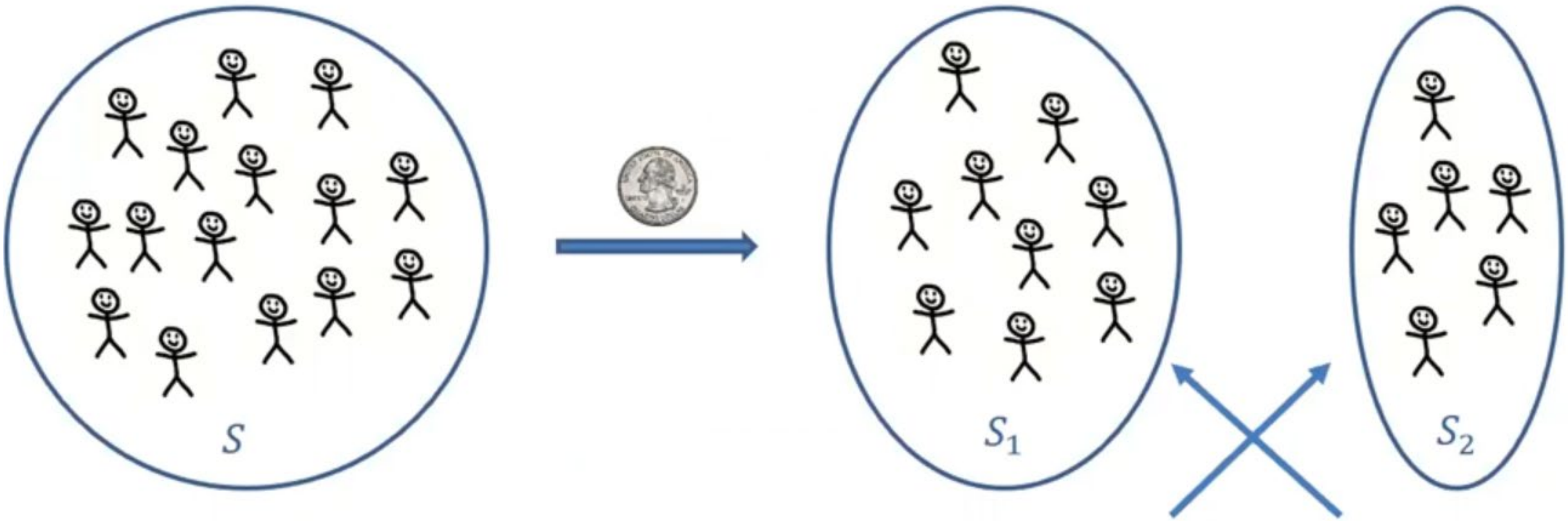
Seller faces a fixed population of buyers. Can he learn an auction that extracts high revenue from a shrinking market?

Setup and background

- We know how to learn high-revenue auctions for limited supply across instances/samples [Mohri and Medina ICML '14, Morgenstern and Roughgarden NIPS '15, Balcan, Sandholm, and Vitercik EC'18]
- What if the mechanism designer does not even have access to samples?
  - Set S of bidders shows up. How to design an auction that performs well on S with no prior information about S?
- Balcan, Blum, Hartline, Mansour FOCS'05 answer this in the *unlimited-supply* setting.
  - Random-sampling mechanism, can be viewed as "learning within an instance"
  - Several follow up works on random sampling for unlimited supply

Random-sampling mechanisms

Balcan, Blum, Hartline, Mansour (FOCS 2005)



Revenue = $\mathrm{Rev}_{M_1}(S_2) + \mathrm{Rev}_{M_2}(S_1)$

M_1 optimal mechanism for S_1

M_2 optimal mechanism for S_2

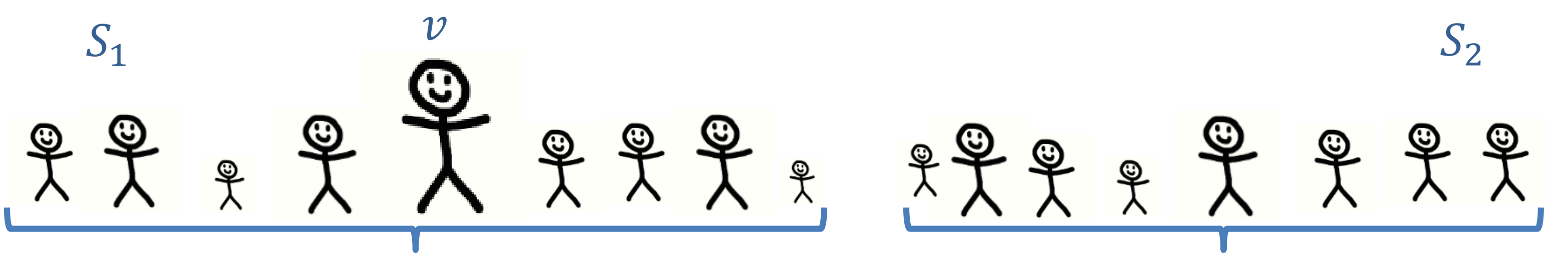
- Incentive compatibility: mechanism used on a bidder's group independent of report
- Sample complexity guarantee: as number of bidders grows, random sampling achieves a (1-ε)-approximation to the optimum.
- **Crucially relies on lack of supply constraints.**

Difficulties of limited supply

- Allocations must be feasible.
- Balcan, Blum, Hartline, Mansour set optimal prices for items, and let buyers purchase according to their demand functions.
- Not possible for limited supply without considering complicated benchmarks depending on the order of buyers purchasing, etc.
- We apply variants of λ-auctions that prescribe feasible allocations.

Difficulties of limited supply

- Try to adapt random sampling to limited supply
- Attempt 1: Partition S into S_1, S_2 as before, compute optimal mechanism for S_1, run on S_2
  - Single item, sold via second-price auction with reserve price
  - Item goes to highest bidder v_1, payment = $\max\{v_2, r\}$



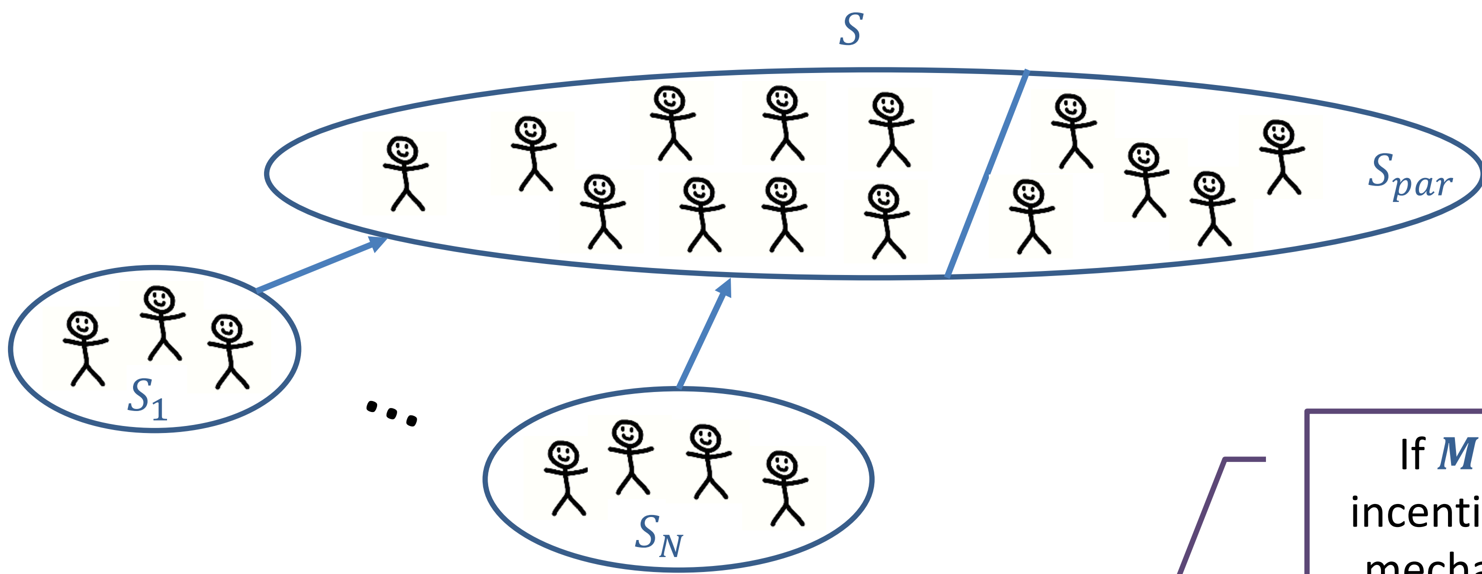
$r_{OPT} = v$ $\quad$ $\mathrm{Rev}_{r_{OPT}}(S_1) = v$ $\quad$ $Rev_{r_{OPT}}(S_2) = 0$

With probability ½, revenue = 0!

Difficulties of limited supply

- Try to adapt random sampling to limited supply
- Attempt 2: Partition S into $S_1,\ldots,S_N,S_{N+1}$. Compute empirically optimal mechanism for $S_1,\ldots,S_N$, run on S_{N+1}
  - If auction class is complex, N needs to be large for generalization guarantees to hold
  - S_{N+1} contains a tiny fraction of bidders, losing a lot of revenue

Learning Within an Instance (LWI) mechanism



If $\boldsymbol{M}$ is a class of incentive compatible mechanisms, LWI is incentive compatible

1. Sample participatory group $S_{par} \sim_p S$
2. Sample learning groups $S_1, \dots, S_N \sim_q S \setminus S_{par}$
3. Compute ERM mechanism $\widehat{M} \in \boldsymbol{M}$ over learning groups
4. Run mechanism $\widehat{M}$ on S_{par}

Basic general guarantee

Theorem (Balcan, **P.**, Sandholm)**.** For $N \geq N_{\boldsymbol{M}}(\varepsilon, \delta)$,

$$Rev_{\widehat{M}}(S_{par}) \geq W(S)\left(L_{\boldsymbol{M}}(S_{par}) - \varepsilon\right) - 2\tau_{\boldsymbol{M}}(q, S_{par})$$

with probability $\geq 1 - 2\delta$ over the draw of the learning groups.

- $N_{\boldsymbol{M}}(\varepsilon, \delta) = O\left(\frac{\mathrm{Pdim}(\boldsymbol{M}) \ln\left(\frac{1}{\delta}\right)}{\varepsilon^2}\right)$ a standard sample complexity term

- $L_{\boldsymbol{M}}(S_{par}) = \frac{OPT_{\boldsymbol{M}}(S_{par})}{W(S)}$ measures how much revenue is lost by restricting to S_{par}

- $\tau_{\boldsymbol{M}}(q, S_{par})$ a complexity/uniformity measure we call *partition discrepancy*

$$\tau_M(q, S_{par}) = \sup_{M \in \boldsymbol{M}} \left| Rev_M(S_{par}) - \mathbf{E}_{S_0 \sim_q S \setminus S_{par}}[Rev_M(S_0)] \right|$$

Making the guarantee more concrete

Bounding partition discrepancy

Theorem (Balcan, **P.**, Sandholm). For $N \geq N_{\boldsymbol{M}}(\varepsilon, \delta), p = 1/2, q = 1/3,$ and $W(S)$ sufficiently large,

$$Rev_{\widehat{M}}(S_{par}) \geq W(S)\big(L_{\boldsymbol{M}}(S_{par}) - 2\varepsilon\big)$$

with probability $\geq 1 - 3\delta$ over the run of LWI.

Instantiation for specific auction class

Theorem (Balcan, **P.**, Sandholm). For $N \geq N_{\boldsymbol{M}}(\varepsilon, \delta),$ and $W(S)$ sufficiently large,

$$Rev_{\widehat{M}}(S_{par}) \geq W(S)\big((1-\eta)p - \varepsilon\big) - 2\tau_{\boldsymbol{M}}(q, S_{par})$$

with probability $\geq 1 - 2\delta$ over the draw of the learning groups. $\boldsymbol{M}$ is the class of “bundling-boosted auctions”.

Comparison to existing results

- Subadditive valuations
  - $O(2^{\sqrt{\log m \log\log m}})$-approximation [Balcan, Blum, Mansour EC 2008]
  - $O(\log^2 m)$-approximation [Chakraborty, Huang, Khanna FOCS 2009; SICOMP 2013]
- Additive valuations
  - $O(\log(h/l))$-approximation, h, l highest and lowest values for any bundle [Sandholm and Likhodedov AAAI 2005; Operations Research 2015]
- Our guarantees
  - No assumptions on valuation functions
  - Fine tuned to structure in the set of bidders

Overview of techniques

- High probability partition discrepancy upper bound
  - In terms of *pseudodimension* $\mathrm{Pdim}(\boldsymbol{M})$, analog of VC dimension for classes of real-valued functions
  - Leverages techniques from learning theory: draws connection between partition discrepancy and L_1 covering number
- Population-size-independent auctions
  - Introduce several new auction classes that can be parameterized in a way that does not depend on the number of bidders participating
  - Derive strong generalization guarantees that do not depend on the number of bidders in S (sample complexity of LWI does not depend on $|S|$)

Outline

Part 1

Efficiently learning high-revenue two-part tariffs from samples/across instances.

Part 2a

Seller faces a single instance of buyers. Can he "learn within an instance" to find a high-revenue auction?

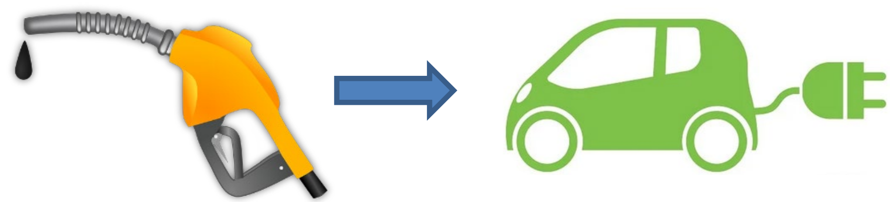
Part 2b

Seller faces a fixed population of buyers. Can he learn an auction that extracts high revenue from a shrinking market?

Examples of shrinking markets

Cord cutters

Fuel oil market



Labor markets among a shrinking population

www.ere.net › labor-market-where-is-everybody-the-sh...

Labor Market: Where Is Everybody? (The Shrinking Labor ...

Sep 24, 2020 — Simply put, the **labor force** participation rate has been falling. The rate for men has been trending downward for nearly 60 years, from 86.7% in ...

www.epi.org › news › shrinking-labor-force-explains-d...

Shrinking labor force explains drop in unemployment

In her analysis of the report, labor economist Heidi Shierholz explained that most of that decline can be explained by the drop in the **labor force** participation rate ...

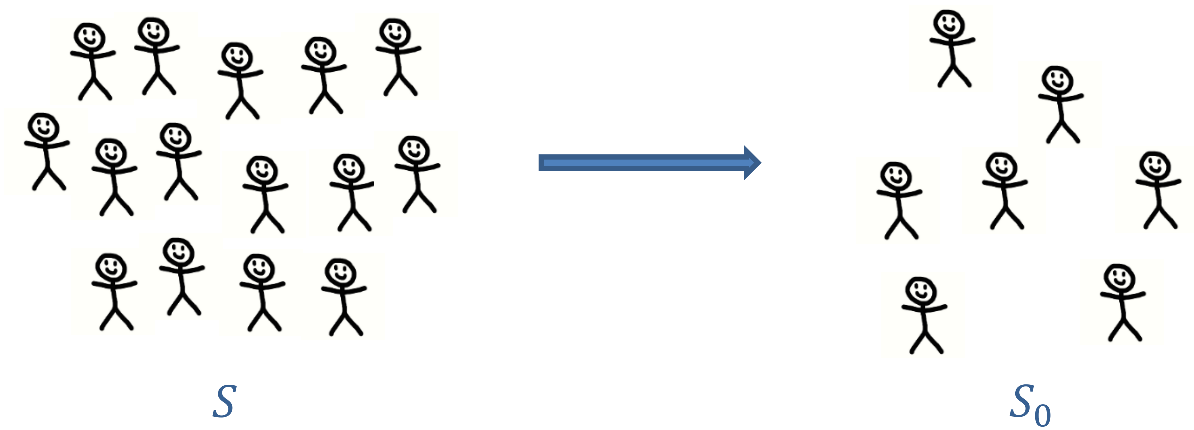
www.wsj.com › articles › covid-shrinks-the-labor-market-...

Covid Shrinks the Labor Market, Pushing Out Women and ...

Dec 3, 2020 — Nearly four million Americans have stopped working or looking for jobs, a 2.2% contraction of the U.S. work force. A smaller **labor market** leaves ...

Modeling a shrinking market

- Fixed set $S = \{v_1, \dots, v_n\}$ of bidder valuations
- Seller knows S
- Each bidder in S shows up independently with probability p



What fraction of revenue can the seller guarantee?

$$\sup_M \mathbf{E}[\mathrm{Rev}_M(S_0)] \geq (???) \cdot W(S)$$

Revenue loss can be drastic

- At first glance answer might appear to be p (or even higher, if revenue thought to have diminishing returns in number of buyers)
- Example 1: $\mathbf{E}[\mathrm{Rev}_{VCG}(S_0)] = p^2 \mathrm{Rev}_{VCG}(S) = p^2(W(S) - \varepsilon)$
  - Due to reduced competition among buyers

c	0	0
$c - \varepsilon/m$	0	0
0	c	0
0	$c - \varepsilon/m$	0
0	0	c
0	0	$c - \varepsilon/m$

VCG gets payment of $c - \varepsilon/m$ for each item so $\mathrm{Rev}_{VCG}(S) = mc - \varepsilon = W(S) - \varepsilon$

But

$$\mathbf{E}[\mathrm{Rev}_{VCG}(S_0)] = \sum_{\text{item } i} \mathbf{E}[\text{Rev from item } i]$$

$$= p^2(mc - \varepsilon)$$

Revenue loss can be drastic

Theorem (Balcan, **P.**, Sandholm). For any $\varepsilon > 0$ there exists a set S of bidders with allocational valuations such that

$$\sup \mathbf{E}[\mathrm{Rev}_M(S_0)] \leq p^{m/2} \cdot (\mathrm{Rev}_{VCG}(S) + 2\varepsilon) + \varepsilon$$

where the supremum is over all possible auctions M.

c	0	0
$c - 2\varepsilon/m$	0	0
0	c	0
0	$c - 2\varepsilon/m$	0
0	0	c
0	0	$c - 2\varepsilon/m$

“Strong” bidders have valuations as before, but with the caveat that the valuation is zero unless each “weak” bidder receives at least one item.

For any significant revenue to be generated, all weak bidders must be in S_0, which happens with probability $p^{m/2}$.

Bidders with values all $\leq 2\varepsilon/m$

How much revenue can be preserved?

- Set S of bidders such that if bidder i wins in VCG, and j leaves, i still wins in VCG (previous example did not satisfy this).

Theorem (Balcan, **P.**, Sandholm). There exists a λ-auction such that

$$\mathbf{E}[\text{Rev}_\lambda(S_0)] \geq \frac{p^2}{16k^{1+\log_{1/\gamma}(4/p)}} \cdot W(S)$$

γ a constant depending on S, $k \approx$ max number of winners in VCG

- Get families of bounds when mechanism designer places practically-motivated constraints on the auction setting
  - Limits the number of winners, to increase competition, or just to reduce hassle
  - Places bundling constraints on the items: e.g. stipulates that certain items ought to be sold together

Proof sketch

- For every $S_0 \subseteq S$ there is a λ-auction with revenue $W(S_0)$
  - Choosing one at random would yield expected revenue $\geq \frac{1}{2^n} \cdot W(S)$
- Equivalence relation: $S_1 \sim S_2$ if corresponding λ-auctions are the same
  - With constant probability, $W(S_0)$ is not too small (S_0 is "heavy")
  - Winner monotonicity -> number of "heavy" equivalence classes not too large (subsets of S obey a tree structure)
  - Use λ-auction corresponding to a uniformly random heavy class

How to choose an auction

- Learn from samples! Seller knows p, can draw samples himself and do ERM
- Unfortunately λ-auctions need exponentially many samples [Balcan, Sandholm, Vitercik 2018]
- Γ-boosted λ-auctions: positive boosts only to allocations in Γ
  - Can get similar guarantee for this class
  - We give practically-efficient algorithms that leverage routines for solving winner determination for learning a Γ-boosted λ-auction that is robust to market shrinkage
  - Runtime is exponential only in |Γ|, polynomial in all other parameters