Mechanism Design with Predictions
Improved Mechanisms for Facility Location

Guido Schäfer

Centrum Wiskunde & Informatica (CWI)
Institute for Logic, Language and Computation (UvA)
g.schaefer@cwi.nl

Course: Game Theory 2025
Lecture May 21, 2025

Plan for Today

Previous Lectures:

- Incentive Compatibility and Vickrey Auction
- Direct-Revelation Mechanisms with Money
- Combinatorial Auctions and VCG
- Mechanism Design Without Money

Today's Lecture: Approximate Mechanism Design

- Facility Location on the Line
- Design strategyproof mechanism with best-possible approximation guarantee
- Impossibility Barrier: strategyproofness vs. approximate efficiency
- New Paradigm: Mechanism Design With Predictions
- Derive improved facility location mechanism using predictions

Facility Location on the Line

Facility Location on the Line

Setting:

- set of agents $N = \{1, \ldots, n\}$ and a single facility (no opening cost)
- each agent $i \in N$ has a true location $v_i \in \mathbb{R}$ (private information)
- each agent $i \in N$ declares a (possibly false) location $x_i \in \mathbb{R}$

denotes loc. chosen by mech. w.r.t. x

Mechanism $\mathcal{M}$:

- collects declared locations of all agents: $\boldsymbol{x} = (x_1, \ldots, x_n)$ (location profile)
- determines a location $z = z(\boldsymbol{x}) \in \mathbb{R}$ where the facility is opened

$\rightarrow$ direct-revelation mechanism without money!

Agents' costs: each agent i wants to minimize their (true) distance to the facility

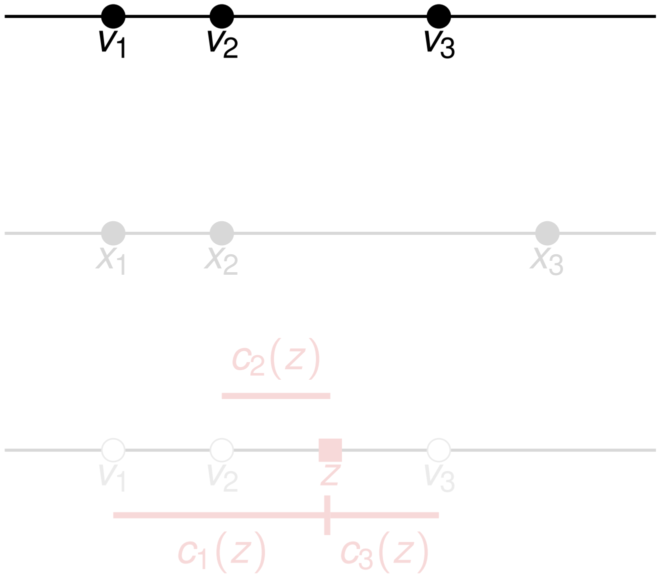
$$c_i(z) = c_i(z, v_i) = |z - v_i|$$

Illustration

true locations

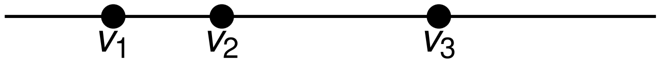
reported locations

agents' costs

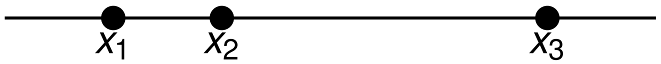


Illustration

true locations



reported locations



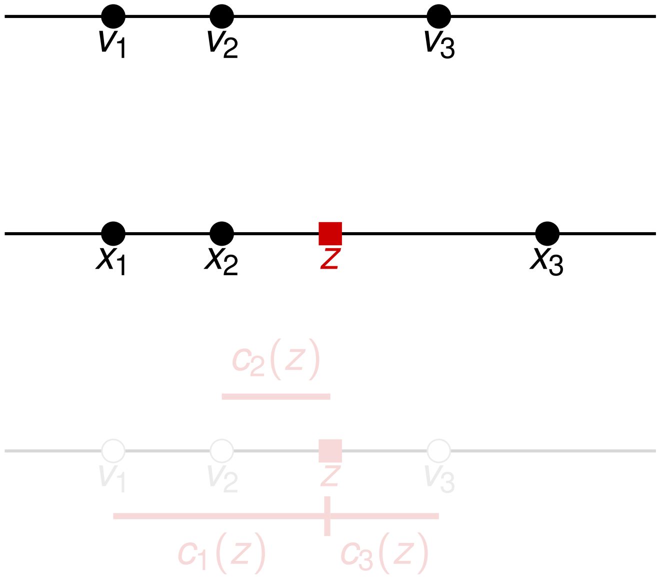
agents' costs

Illustration

true locations

reported locations

agents' costs

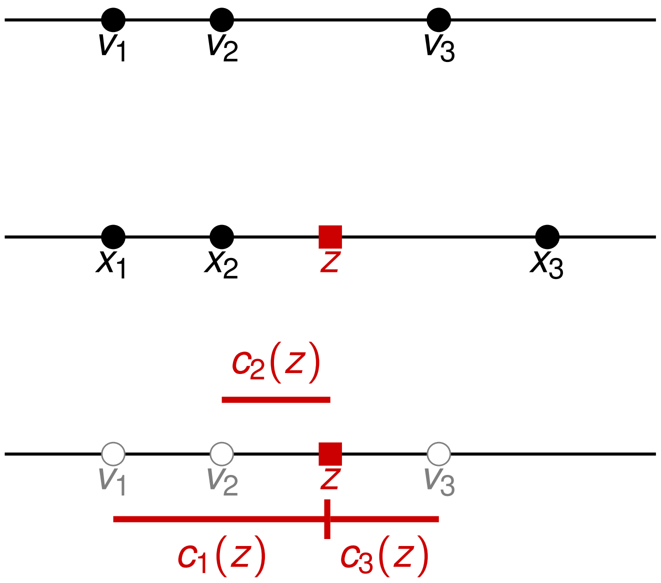


Illustration

true locations

reported locations

agents' costs



1 **Strategyproofness (SP):** mechanism ensures that for each agent i, it is a dominant strategy to report their true location:

true loc. any report

$$\forall i \in N \quad \forall \mathbf{x}_{-i} \in \mathbb{R}^{n-1}: \qquad c_i(z(v_i, \mathbf{x}_{-i})) \leq c_i(z(x_i, \mathbf{x}_{-i})) \quad \forall x_i \in \mathbb{R}$$

2 **Efficiency (EFF):** mechanism minimizes the maximum cost over all agents:

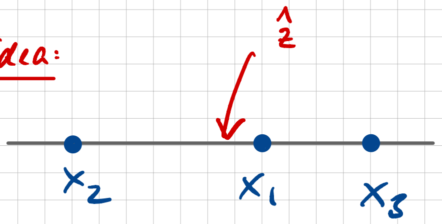
$$\forall \mathbf{x} \in \mathbb{R}^n: \quad z(\mathbf{x}) = \arg\min_{z \in \mathbb{R}} SC(z, \mathbf{x}) \quad \text{where} \quad SC(z, \mathbf{x}) = \max_{i \in N} c_i(z, x_i)$$

Egalitarian Social Cost: choose location that minimizes cost of furthest-away agent

$\rightarrow$ use $z^*(\mathbf{x})$ / $OPT(\mathbf{x})$ to refer to optimal location / egalitarian social cost

Question: How can we design a mechanism satisfying SP and EFF?

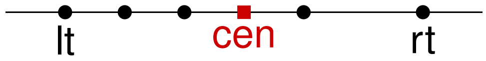
Idea:



choose z as center interval $[x_2, x_3]$

Idea: Center Mechanism

Notation: given location profile $\boldsymbol{x} = (x_1, \ldots, x_n)$, define



- $\mathrm{lt}(\boldsymbol{x}) = \min_{i \in N} x_i$ (leftmost location)
- $\mathrm{rt}(\boldsymbol{x}) = \max_{i \in N} x_i$ (rightmost location)
- $\mathrm{cen}(\boldsymbol{x}) = \frac{1}{2}(\mathrm{lt}(\boldsymbol{x}) + \mathrm{rt}(\boldsymbol{x}))$ (center of $[\mathrm{lt}(\boldsymbol{x}), \mathrm{rt}(\boldsymbol{x})]$)

  $= \ell t(x) + \frac{1}{2}(rt(x) - \ell t(x))$

Mechanism: Center

---

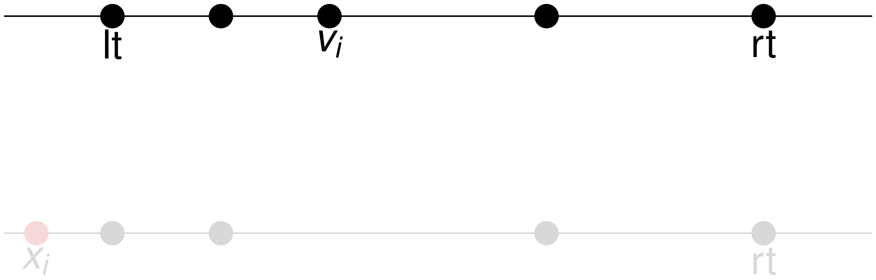
1. Collect location profile $\boldsymbol{x} = (x_1, \ldots, x_n)$ of the agents' reports
2. Choose the center: $z(\boldsymbol{x}) = \mathrm{cen}(\boldsymbol{x})$

Can prove: Center satisfies EFF.

Question: What about strategyproofness?

But: Strategyproofness Fails!

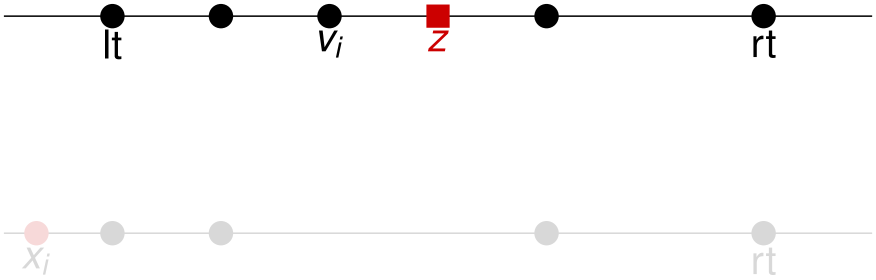
Example:



Question: Can we fix this?

But: Strategyproofness Fails!

Example:



Question: Can we fix this?

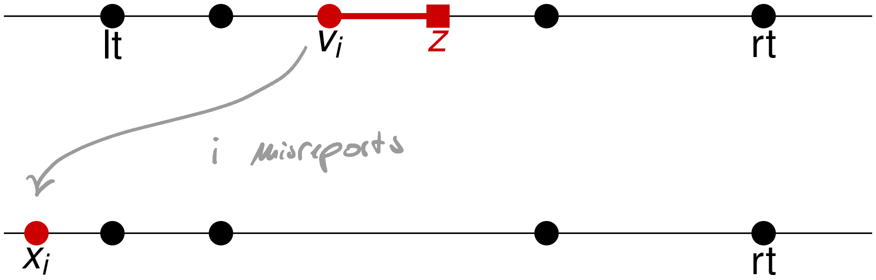
But: Strategyproofness Fails!

Example:

Question: Can we fix this?

But: Strategyproofness Fails!

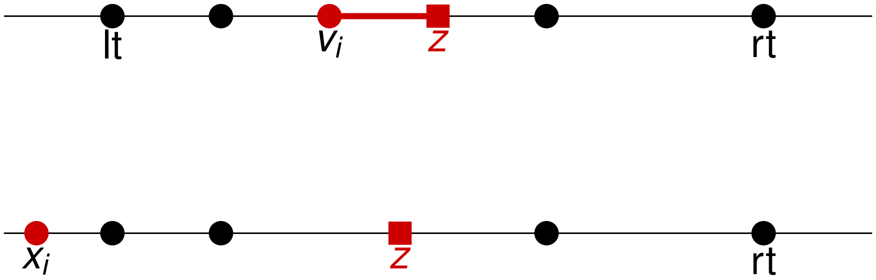
Example:



Question: Can we fix this?

But: Strategyproofness Fails!

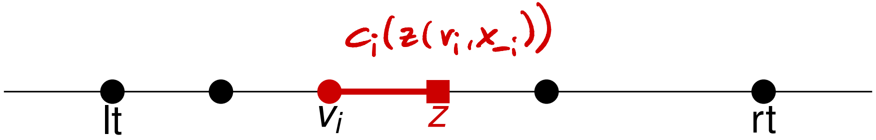
Example:

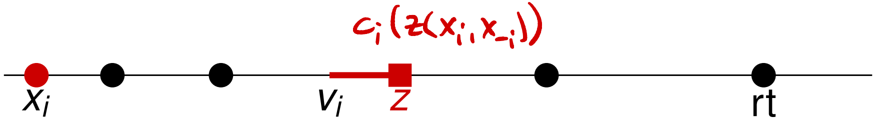


Question: Can we fix this?

But: Strategyproofness Fails!

Example:

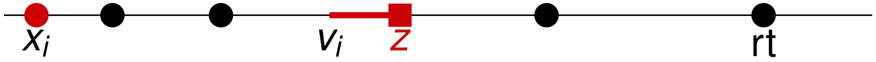




contradicts SP!

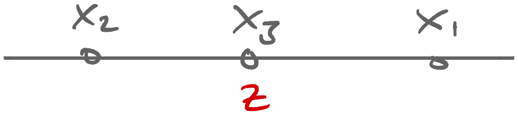
Question: Can we fix this?

Example:



Question: Can we fix this?

Idea: MEDIAN Mechanism



Notation:

- Let σ be permutation of $[n]$ such that $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \cdots \leq x_{\sigma(n)}$
- Define median as $\text{med}(\boldsymbol{x}) = x_{\sigma(k+1)}$ for $n = \{2k \mid 2k+1\}$

Mechanism: MEDIAN

---

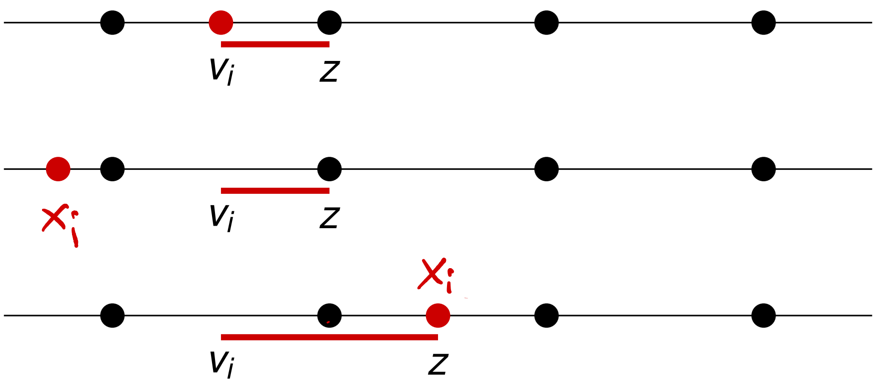
1 Collect location profile $\boldsymbol{x} = (x_1, \ldots, x_n)$ of the agents' reports
2 Choose the median: $z(\boldsymbol{x}) = \text{med}(\boldsymbol{x})$

Can prove: MEDIAN is strategyproof!

Strategyproofness

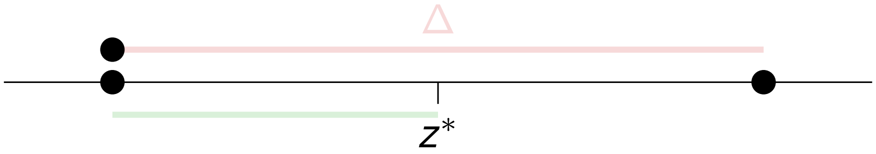
Lemma 1: MEDIAN *satisfies strategyproofness.*

Proof (by picture):



But: Efficiency Fails!

Example: $n = 3$



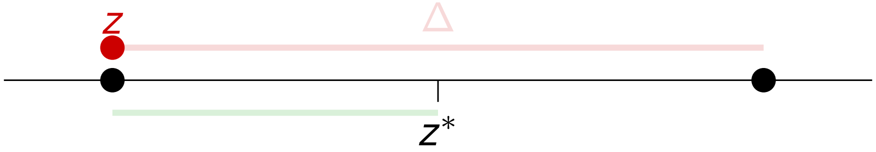
Define $\Delta = |\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x})|$. Optimal cost is $OPT(\boldsymbol{x}) = \frac{1}{2}\Delta$.
But MEDIAN outputs $z = z(\boldsymbol{x})$ and has social cost $SC(z, \boldsymbol{x}) = \Delta$.
$\rightarrow$ we lose a factor of 2 here!

α-Efficiency (α-EFF): mechanism α-approximates egalitarian social cost ($\alpha \geq 1$):

$$\forall \boldsymbol{x} \in \mathbb{R}^n : \quad SC(z(\boldsymbol{x}), \boldsymbol{x}) \leq \alpha \cdot OPT(\boldsymbol{x})$$

But: Efficiency Fails!

Example: $n = 3$



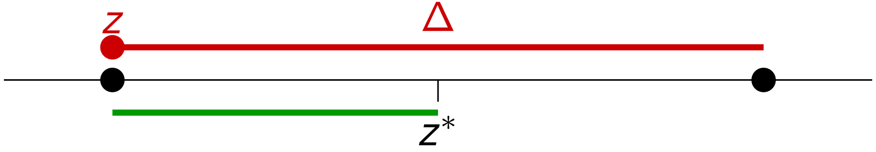
Define $\Delta = |\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x})|$. Optimal cost is $OPT(\boldsymbol{x}) = \frac{1}{2}\Delta$.
But MEDIAN outputs $z = z(\boldsymbol{x})$ and has social cost $SC(z, \boldsymbol{x}) = \Delta$.
$\rightarrow$ we lose a factor of 2 here!

α-Efficiency (α-EFF): mechanism α-approximates egalitarian social cost ($\alpha \geq 1$):

$$\forall \boldsymbol{x} \in \mathbb{R}^n : \quad SC(z(\boldsymbol{x}), \boldsymbol{x}) \leq \alpha \cdot OPT(\boldsymbol{x})$$

But: Efficiency Fails!

Example: $n = 3$



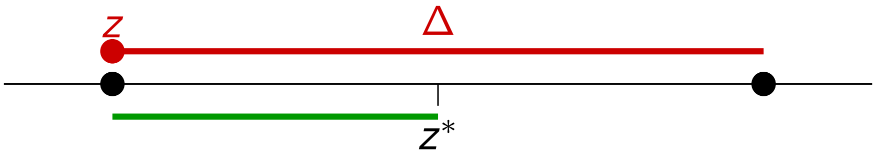
Define $\Delta = |\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x})|$. Optimal cost is $OPT(\boldsymbol{x}) = \frac{1}{2}\Delta$.
But MEDIAN outputs $z = z(\boldsymbol{x})$ and has social cost $SC(z, \boldsymbol{x}) = \Delta$.
$\rightarrow$ we lose a factor of 2 here!

α-Efficiency (α-EFF): mechanism α-approximates egalitarian social cost ($\alpha \geq 1$):

$$\forall \boldsymbol{x} \in \mathbb{R}^n : \quad SC(z(\boldsymbol{x}), \boldsymbol{x}) \leq \alpha \cdot OPT(\boldsymbol{x})$$

But: Efficiency Fails!

Example: $n = 3$



Define $\Delta = |\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x})|$. Optimal cost is $OPT(\boldsymbol{x}) = \frac{1}{2}\Delta$.
But MEDIAN outputs $z = z(\boldsymbol{x})$ and has social cost $SC(z, \boldsymbol{x}) = \Delta$.
$\rightarrow$ we lose a factor of 2 here!

α-Efficiency (α-EFF): mechanism α-approximates egalitarian social cost ($\alpha \geq 1$):

$$\forall \boldsymbol{x} \in \mathbb{R}^n : \quad SC(z(\boldsymbol{x}), \boldsymbol{x}) \leq \alpha \cdot OPT(\boldsymbol{x})$$

MEDIAN Mechanism

Theorem 1: MEDIAN *is strategyproof and* 2*-efficient.*

Proof (2-EFF). Let $\boldsymbol{x}$ be given. The optimal cost is

$$OPT(\boldsymbol{x}) = |\mathrm{rt}(\boldsymbol{x}) - \mathrm{cen}(\boldsymbol{x})| = \mathrm{rt}(\boldsymbol{x}) - \tfrac{1}{2}(\mathrm{lt}(\boldsymbol{x}) + \mathrm{rt}(\boldsymbol{x})) = \tfrac{1}{2}(\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x}))$$

MEDIAN returns location $z = z(\boldsymbol{x})$ of cost

$$SC(z, \boldsymbol{x}) \leq \mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x}) = 2 \cdot OPT(\boldsymbol{x}). \qquad \square$$

Remarks:

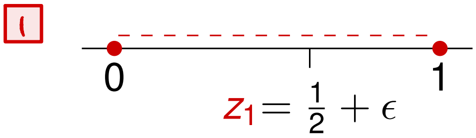
- Choosing any k-th order statistic $x_{\sigma(k)}$ with $k \in \{1, \dots, n\}$ works as well!
- Generalizes to single-peaked preferences and group-strategyproofness

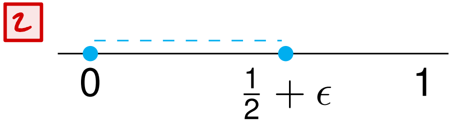
Impossibility Barrier

Theorem 2: *There is no strategyproof mechanism that is α-efficient with $\alpha < 2$.*

strict ineq. important!

Proof. Suppose $\mathcal{M}$ is strategyproof and α-efficient with $\alpha < 2$. Let $N = \{1, 2\}$.





1 Consider $\mathbf{x}^1 = (0, 1)$. Then $z_1^* = \frac{1}{2}$ has cost $\frac{1}{2}$. Assume wlog that $\mathcal{M}$ returns $z_1 = \frac{1}{2} + \epsilon$ for some $\epsilon \in [0, \frac{1}{2})$. open!

2 Consider $\mathbf{x}^2 = (0, \frac{1}{2} + \epsilon)$. Then $z_2^* = \frac{1}{4} + \frac{\epsilon}{2}$ has cost $\frac{1}{4} + \frac{\epsilon}{2}$. Since $\mathcal{M}$ is α-efficient with $\alpha < 2$, we have $z_2 \in (0, \frac{1}{2} + \epsilon)$.

But then: Suppose true locations are $\mathbf{x}^2$ as in **2**. We have $c_2(z_2) > 0$. If agent 2 declares $x_2^2 = 1$, we are in **1** and $c_2(z_1) = 0$, contradicting strategyproofness. $\square$

Wrap-Up

1. MEDIAN Mechanism: derived a strategyproof mechanism that is 2-efficient for facility location on the line
   $\rightarrow$ MEDIAN can be implemented to run in polynomial time

2. Impossibility Barrier: there is no strategyproof mechanism that achieves $(2-\varepsilon)$-efficiency for any $\varepsilon > 0$
   $\rightarrow$ holds because of conflicting objectives SP and α-EFF, independently of any computational constraints (e.g., even for exponential time mechanisms!)

3. Contributions 1 and 2 together imply that our mechanism is best-possible: no better mechanism exists in terms of SP vs. α-EFF tradeoff

4. Glimpse only: more complex facility location problems, different incentive compatibility notions, social cost objectives, etc. have been studied in the literature

Towards Mechanism Design with Predictions

Traditional Mechanism Design

profile $\boldsymbol{x}$	$\rightarrow$	mechanism $\mathcal{M}(\boldsymbol{x})$	$\rightarrow$	output z

Objectives:

- strategyproofness $\quad \mathcal{M}$ incentivizes truthful reports $x_i = v_i$
- α-efficiency $\quad \mathcal{M}$ computes α-approximate solution
- ...

Note: we provide guarantees by proving that these objectives are always achieved

Mechanisms with Predictions

prediction $\hat{p}$ profile $\boldsymbol{x}$	$\rightarrow$	mechanism $\mathcal{M}(\boldsymbol{x}, \hat{p})$	$\rightarrow$	output z

Prediction: $\hat{p}$ is prediction of some problem-relevant parameters

- predictions might be obtained from actual data via machine-learning techniques
- **Question:** Can we leverage predictions to develop improved mechanisms?
- **NB:** $\hat{p}$ might be erroneous, but we still care about provable guarantees!

Objectives:

- strategyproofness — $\mathcal{M}$ incentivizes truthful reports $x_i = v_i$
- α-consistency — $\mathcal{M}$ computes α-approximate solution if prediction is accurate
- β-robustness — $\mathcal{M}$ computes β-approximate solution even if prediction is off
- γ-approximability — $\mathcal{M}$ computes $\gamma(\eta)$-approximate solution if prediction error is η

Mechanisms with Predictions

What are suitable predictions?

- prediction of the true location profile $\boldsymbol{v} = (v_1, \ldots, v_n)$
- prediction of the optimal facility location $\hat{z}$ (aggregated information)

Facility Location Mechanism with Predictions:

1. Obtain prediction $\hat{z}$ of the optimal facility location
2. Collect location profile $\boldsymbol{x} = (x_1, \ldots, x_n)$ of the agents' reports
3. Choose facility location $z = z(\boldsymbol{x}, \hat{z})$

(**Crucial:** $\hat{z}$ does not depend on reports!)

Question: How can we exploit the prediction $\hat{z}$ to design improved mechanisms?

Consistency, Robustness, Approximability

- α-consistent: mechanism is α-approximate if the prediction is accurate

$$\forall \boldsymbol{x},\ \hat{z} = z^*(\boldsymbol{x}) : \qquad SC(z(\boldsymbol{x}, \hat{z}), \boldsymbol{x}) \leq \alpha \cdot OPT(\boldsymbol{x})$$

- β-robust: mechanism is always β-approximate even if the prediction is off

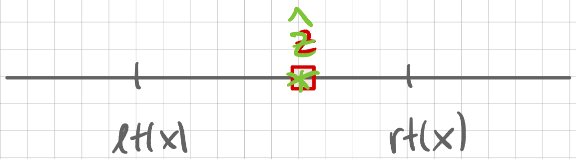
$$\forall \boldsymbol{x}\ \ \forall \hat{z} : \qquad SC(z(\boldsymbol{x}, \hat{z}), \boldsymbol{x}) \leq \beta \cdot OPT(\boldsymbol{x})$$

- γ-approximate: mechanism is $\gamma(\eta)$-approximate if prediction error is bounded by η

$$\forall \boldsymbol{x}\ \ \forall \hat{z} : \eta(\boldsymbol{x}, \hat{z}) \leq \eta : \qquad SC(z(\boldsymbol{x}, \hat{z}), \boldsymbol{x}) \leq \gamma(\eta) \cdot OPT(\boldsymbol{x})$$

Facility Location with Predictions

Q: How to use prediction $\hat{z}$?



Mechanism WITHINBOUNDARIES

Mechanism: WITHINBOUNDARIES

Obtain prediction $\hat{z}$ of the optimal facility location
Collect location profile $\boldsymbol{x} = (x_1, \dots, x_n)$ of the agents' reports
if $\hat{z} < \text{lt}(\boldsymbol{x})$ **then** choose leftmost location: $z(\boldsymbol{x}, \hat{z}) = \text{lt}(\boldsymbol{x})$
else if $\hat{z} > \text{rt}(\boldsymbol{x})$ **then** choose rightmost location: $z(\boldsymbol{x}, \hat{z}) = \text{rt}(\boldsymbol{x})$
else choose predicted location: $z(\boldsymbol{x}, \hat{z}) = \hat{z}$

Can prove: WITHINBOUNDARIES is 1-consistent and 2-robust!

Implications:

- mechanism outputs optimal solution if prediction is accurate
- mechanism is never worse than 2-efficient (same guarantee as before)
- by using predictions, we can break the impossibility barrier!

Best of Both Worlds

Theorem 3: WITHINBOUNDARIES *is strategyproof,* 1*-consistent and* 2*-robust.*

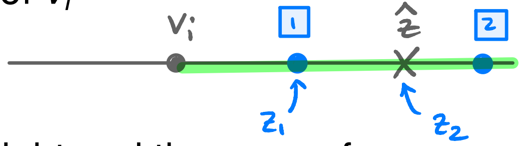
Proof: strategyproof

Fix some agent i and reports $\boldsymbol{x}_{-i}$. Assume $v_i \leq \hat{z}$ (analogous otherwise).

Case 1: all reported locations in $\boldsymbol{x}_{-i}$ are to the left of v_i

Then $z((v_i, \boldsymbol{x}_{-i}), \hat{z}) = v_i$ and i does not want to deviate.

Case 2: at least one reported location in $\boldsymbol{x}_{-i}$ is to the right of v_i



Then $z((v_i, \boldsymbol{x}_{-i}), \hat{z}) \geq v_i$.
If i misreports $x_i \leq v_i$: no change in outcome.
If i misreports $x_i > v_i$: $z((x_i, \boldsymbol{x}_{-i}), \hat{z})$ can only move to the right and thus away from v_i.

Best of Both Worlds

Theorem 3: WITHINBOUNDARIES *is strategyproof,* 1*-consistent and* 2*-robust.*

Proof: 1-consistent

Suppose the prediction is accurate, i.e., $\hat{z} = z^*(\boldsymbol{x})$. Then $\hat{z} = z^*(\boldsymbol{x}) \in [\mathrm{lt}(\boldsymbol{x}), \mathrm{rt}(\boldsymbol{x})]$ and the mechanism thus outputs $\hat{z}$.

Proof: 2-robust

The mechanism always outputs a location in $[\mathrm{lt}(\boldsymbol{x}), \mathrm{rt}(\boldsymbol{x})]$. Cost of any agent is thus at most $\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x})$. Optimal solution has egalitarian social cost $\frac{1}{2}(\mathrm{rt}(\boldsymbol{x}) - \mathrm{lt}(\boldsymbol{x}))$.

Outlook: How about Facility Location in $\mathbb{R}^2$?

Mechanism: MinBoundingBox

Obtain prediction $\hat{z} = (\hat{x}, \hat{y})$ of the optimal facility location

Collect reported location profile $\boldsymbol{x} = ((x_1, y_1), \dots, (x_n, y_n))$ of the agents

$x_z = \text{WithinBoundaries}((x_1, \dots, x_n), \hat{x})$

$y_z = \text{WithinBoundaries}((y_1, \dots, y_n), \hat{y})$

Choose $z(\boldsymbol{x}, \hat{z}) = (x_z, y_z)$

Skipped!

Given prediction $\hat{z}$, define error parameter $\eta(\boldsymbol{x}, \hat{z}) = \dfrac{|\hat{z} - z^*(\boldsymbol{x})|}{OPT(\boldsymbol{x})}$

Theorem 4: MinBoundingBox *is* $\min\{1 + \eta, 1 + \sqrt{2}\}$*-approximate.*

Conclusions

1. WITHINBOUNDARIES Mechanism: simple predictions lead to mechanism with improved guarantees for facility location

2. Best of Both Worlds: here we can improve best-case guarantee (1-consistent) without worsening the worst-case guarantee (2-robust). **But:** not always possible!

3. Power of Predictions:
   - might overcome traditional impossibility barriers
   - overarching theme: beyond worst-case analysis
   - game-changer: advent of ML techniques
   - results might make it into actual applications (Google, Meta, etc.)

join AGT
in Sept!

4. Mechanism Design with Predictions is a just emerging research field!

References

- Ariel D. Procaccia and Moshe Tennenholtz. 2013. *Approximate Mechanism Design without Money*. *ACM Transactions on Economics and Computation*, 1(4), Article 18, 26 pages. https://doi.org/10.1145/2542174.2542175.

- Priyank Agrawal, Eric Balkanski, Vasilis Gkatzelis, Tingting Ou, and Xizhi Tan. 2022. *Learning-Augmented Mechanism Design: Leveraging Predictions for Facility Location*. arXiv preprint arXiv:2204.01120. https://arxiv.org/abs/2204.01120.