GOOD ALLOCATIONS FROM BAD ESTIMATES

Sílvia Casacuberta

Stanford University

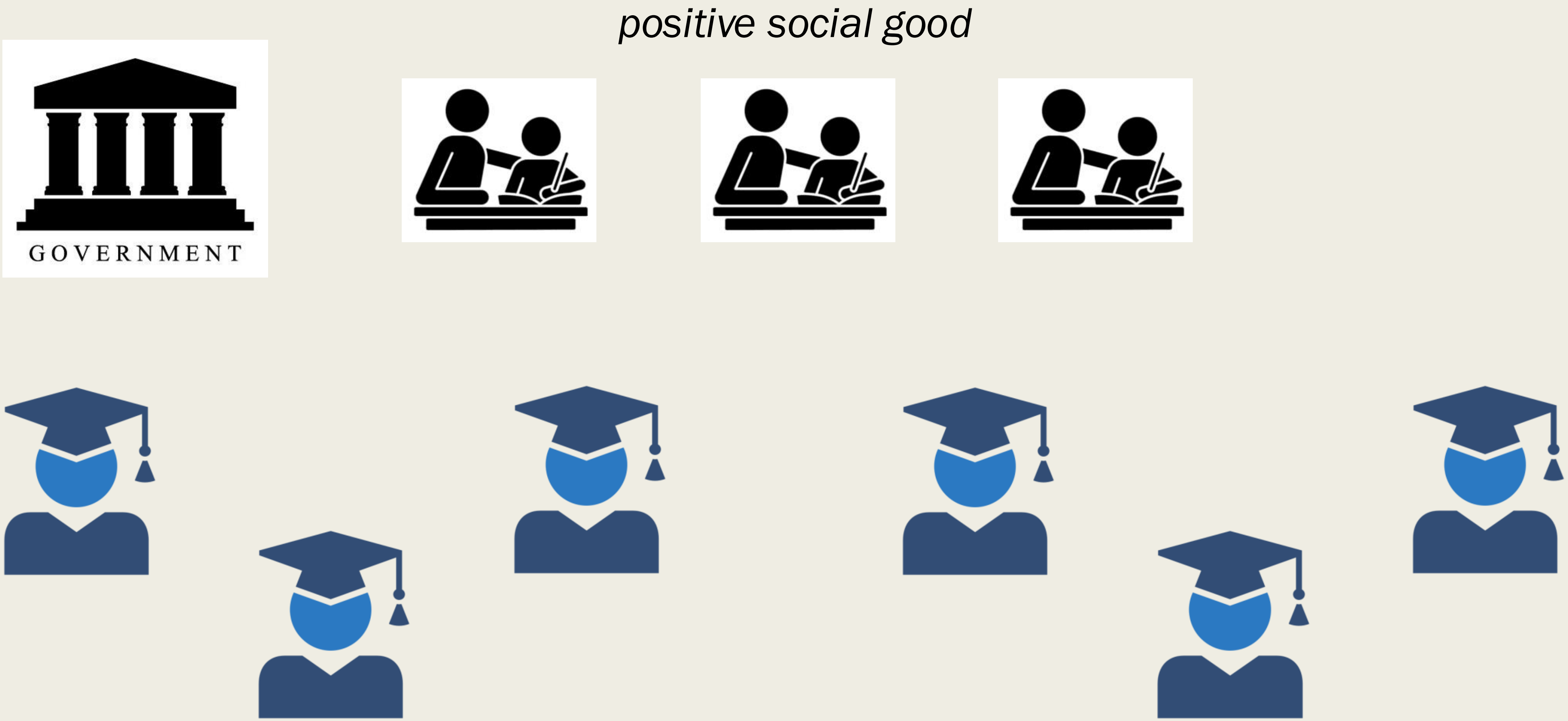
Moritz Hardt

Max Planck Institute for Intelligent Systems

ICLR 2026

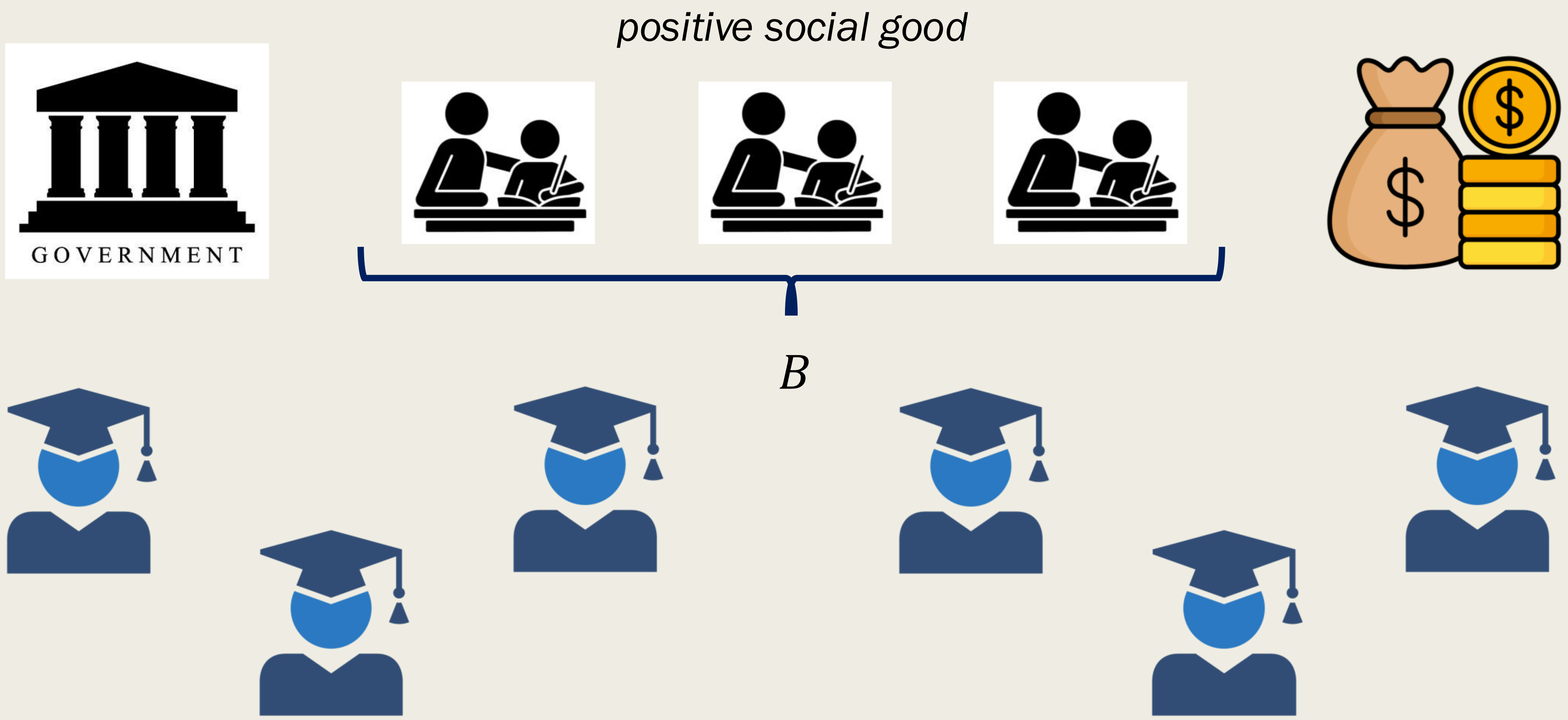
Allocation vs Prediction

- Usual paradigm: we allocate resources by performing individual predictions

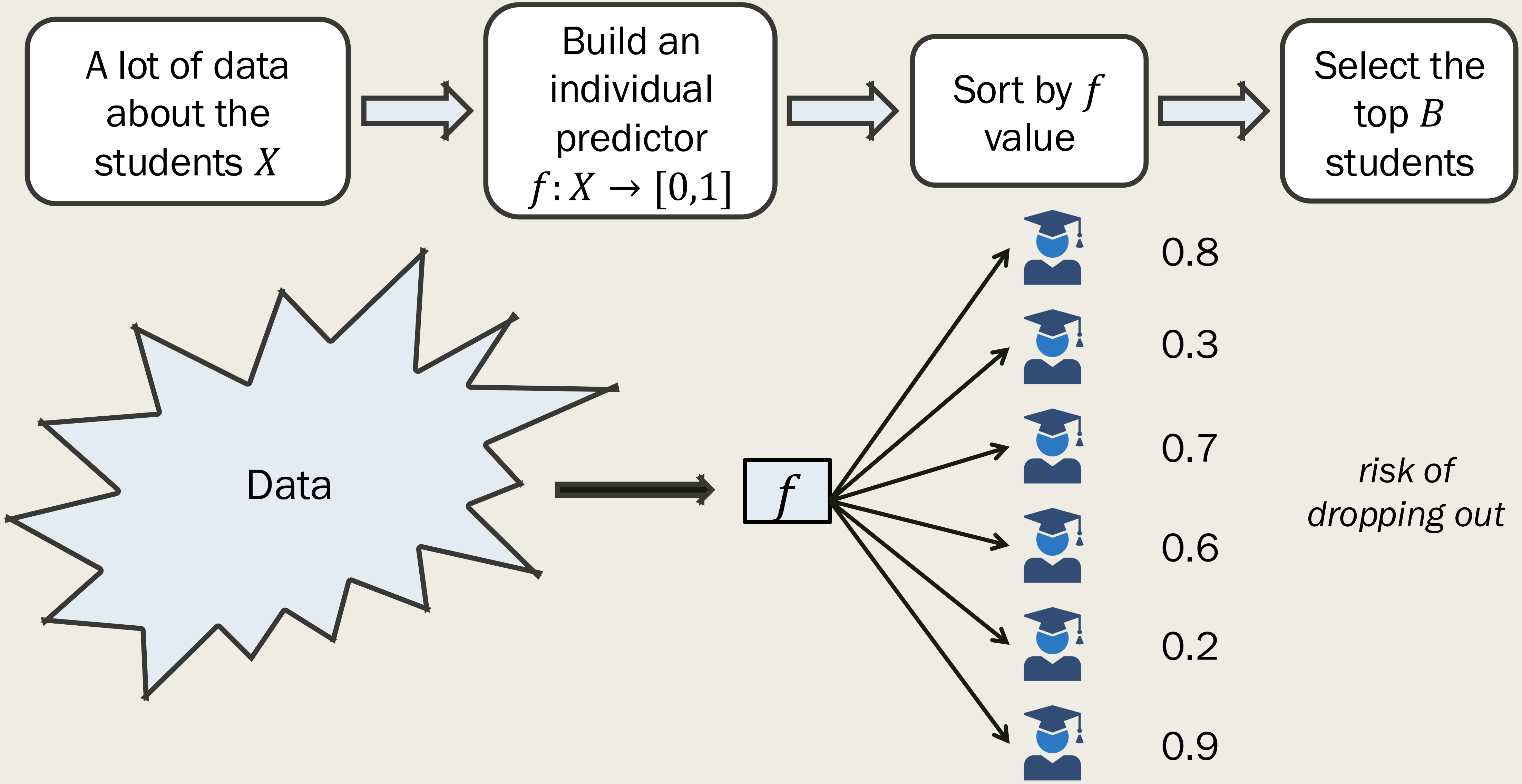


Allocation vs Prediction

- Usual paradigm: we allocate resources by performing individual predictions



Allocation vs Prediction



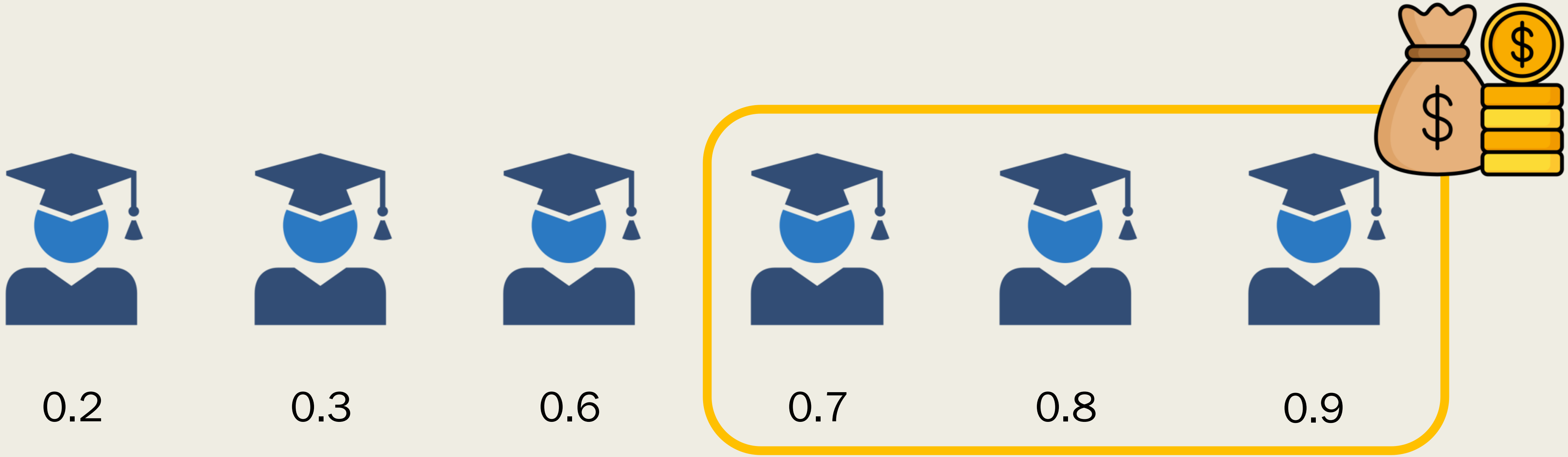
Allocation vs Prediction

A lot of data about the students X → Build an individual predictor $f: X \to [0,1]$ → Sort by f value → Select the top B students

0.2 0.3 0.6 0.7 0.8 0.9

Allocation vs Prediction

A lot of data about the students X → Build an individual predictor $f: X \to [0,1]$ → Sort by f value → Select the top B students



But do we really need to predict first?

A lot of data about the students X → Build an individual predictor $f: X \to [0,1]$ → Sort by f value → Select the top B students

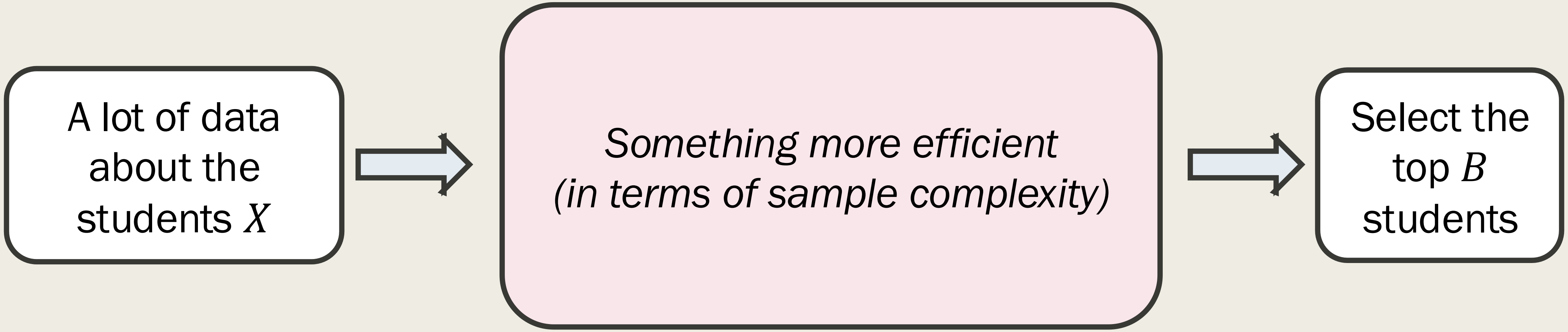
But do we really need to predict first?

A lot of data about the students X

Something more efficient
(in terms of sample complexity)

Select the top B students

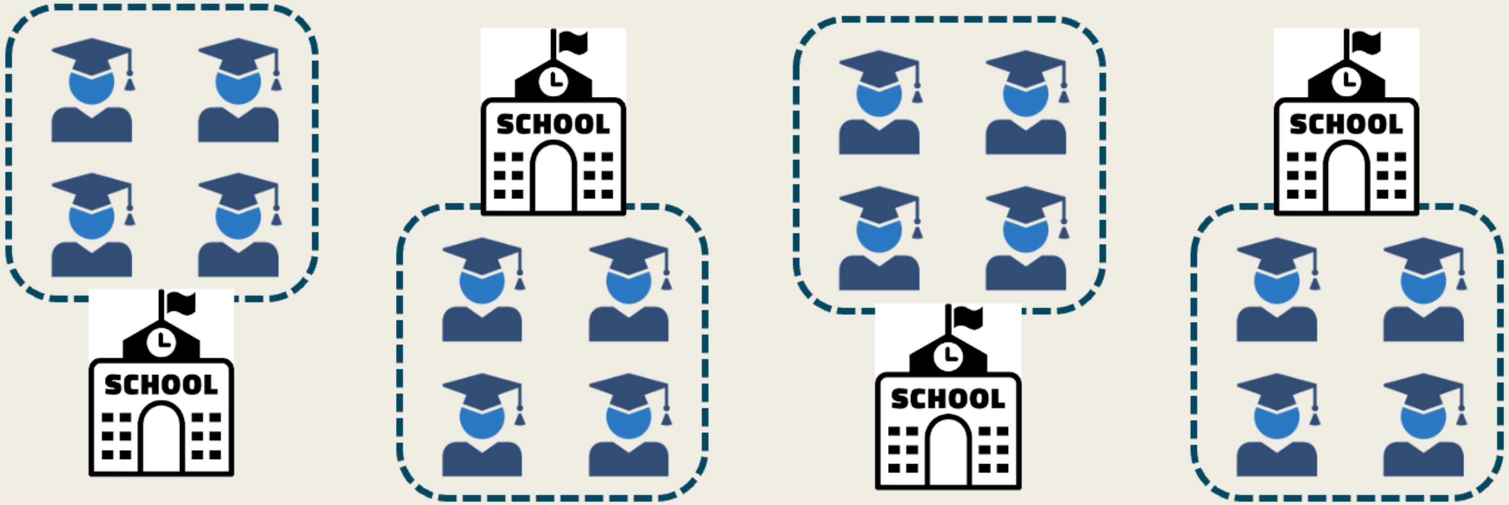
But do we really need to predict first?



- **Goal:** allocation, not prediction
- Individual prediction is the means, not the end goal

This premise is starting to be questioned

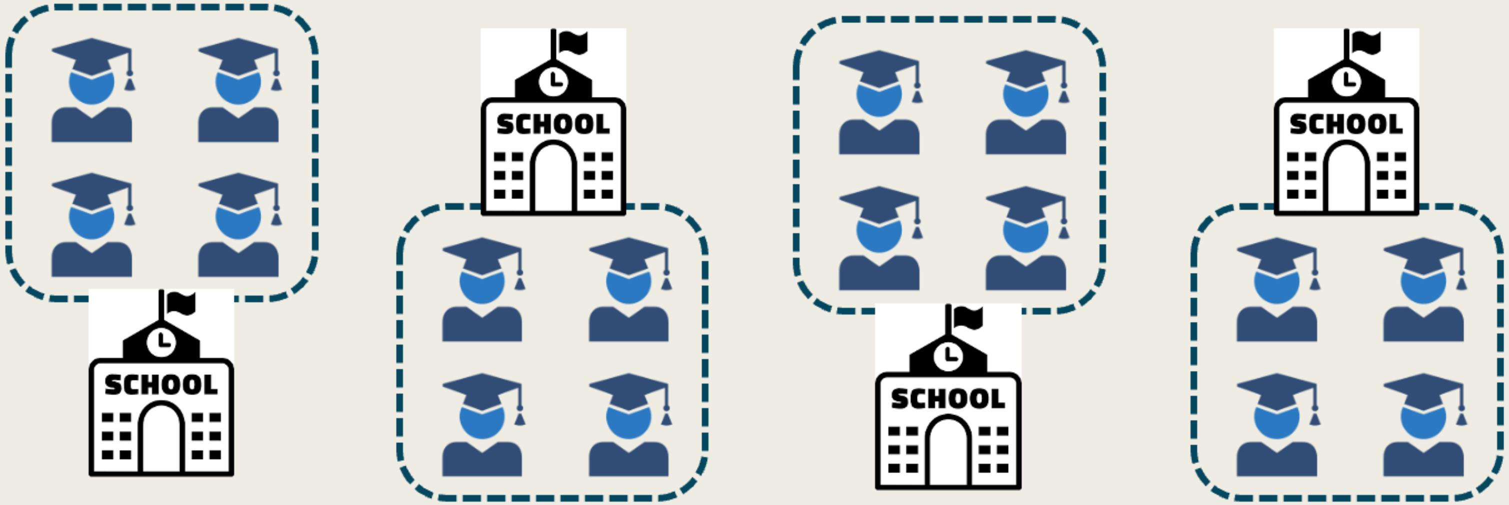
- *Difficult lessons on social prediction from Wisconsin Public Schools* [Perdomo, Britton, Hardt, Abebe, 2023]
- *Allocation requires prediction only if inequality is low* [Shirali, Abebe, Hardt, 2024]
- *The value of prediction in identifying the worst-off* [Fischer-Abaigar, Kern, Perdomo, 2025]

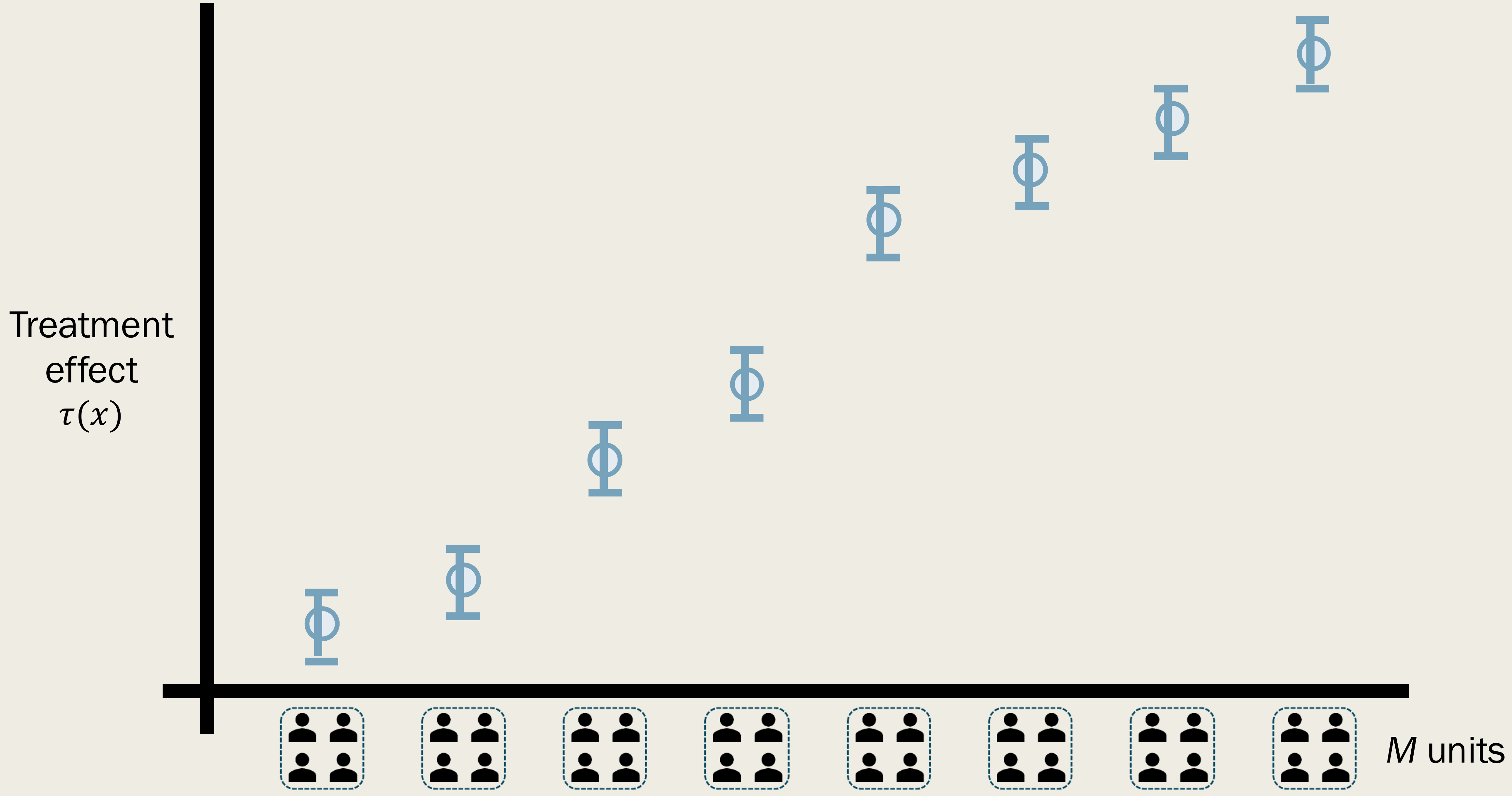


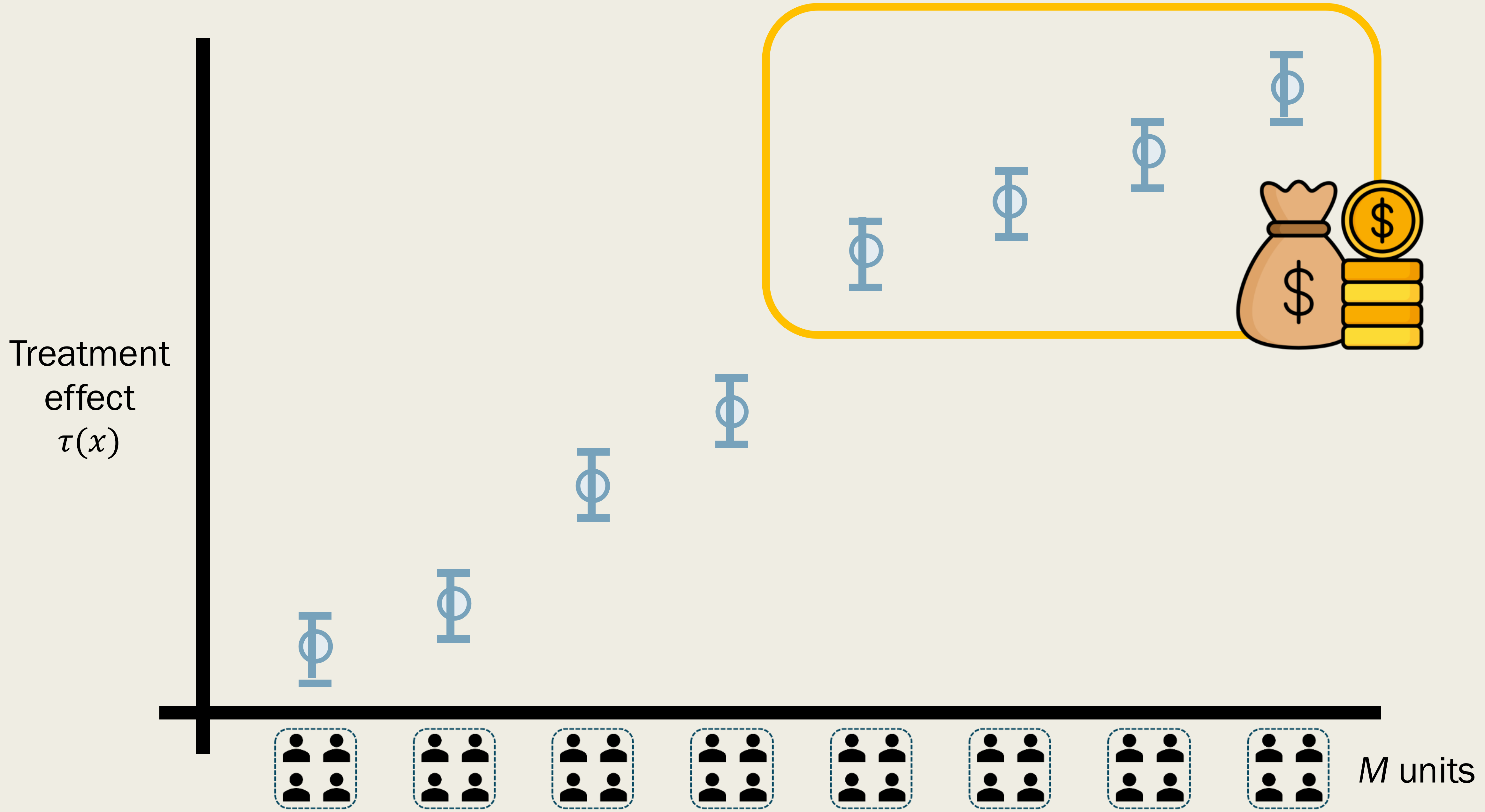
This premise is starting to be questioned

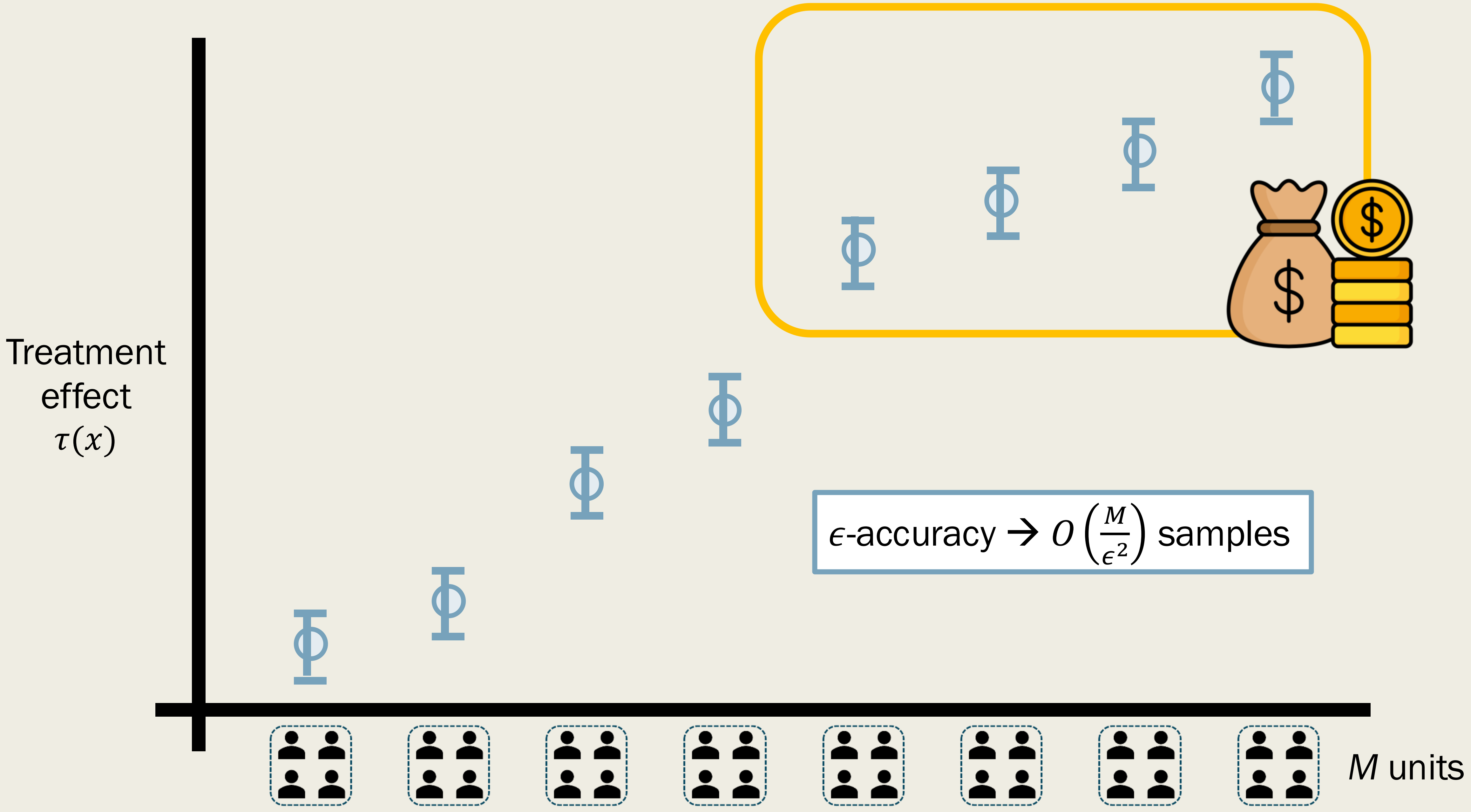
In this work: we focus on sample complexity

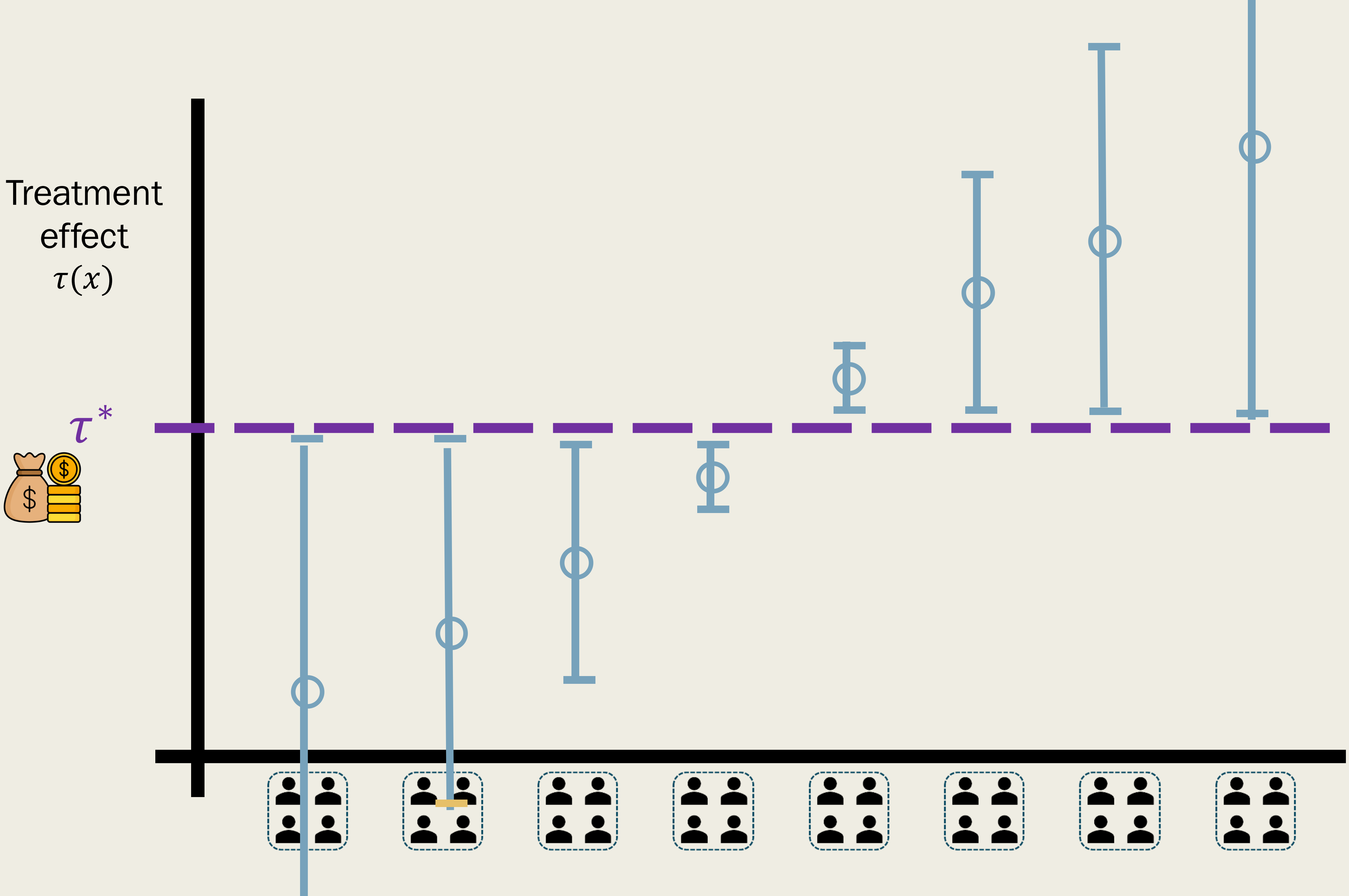
- *Difficult lessons on social prediction from Wisconsin Public Schools* [Perdomo, Britton, Hardt, Abebe, 2023]
- *Allocation requires prediction only if inequality is low* [Shirali, Abebe, Hardt, 2024]
- *The value of prediction in identifying the worst-off* [Fischer-Abaigar, Kern, Perdomo, 2025]

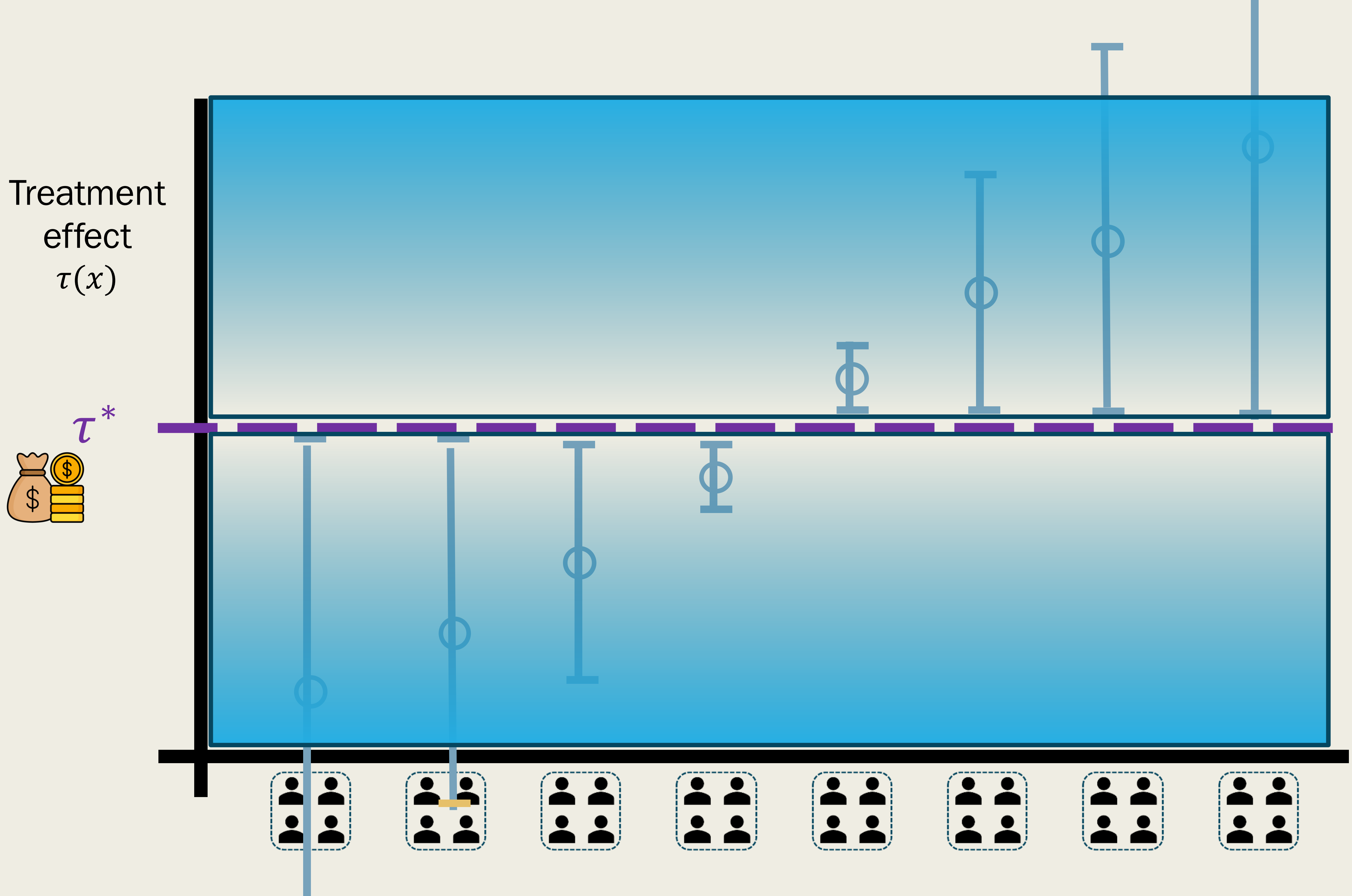



Treatment
effect
$\tau(x)$
M units

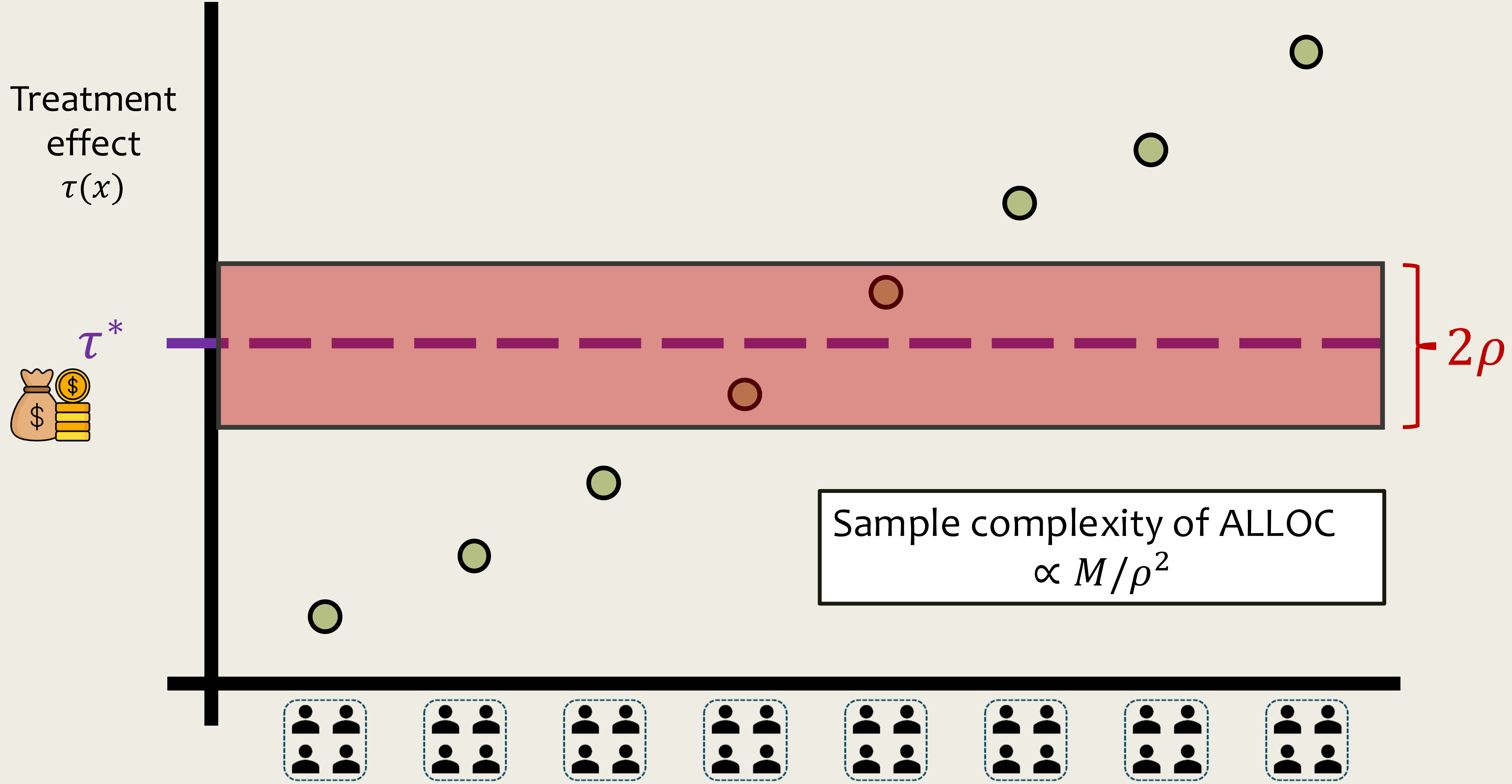

Treatment effect $\tau(x)$
M units


Treatment effect $\tau(x)$
ϵ-accuracy → $O\left(\frac{M}{\epsilon^2}\right)$ samples
M units


Treatment
effect
$\tau(x)$
τ^*



Treatment effect $\tau(x)$
τ^*
$


Treatment effect
$\tau(x)$
τ^*
2ρ
Sample complexity of ALLOC
$\propto M/\rho^2$

What should we set ρ to?

- Goal: achieve an ϵ-approximation of the optimal allocation U^* (in terms of value)

$$\frac{Value(ALLOC)}{Value(ALLOC^*)} = \frac{\sum_{u \in U} \tau(u)}{\sum_{u \in U^*} \tau(u)} \geq 1 - \epsilon$$

- Trade-off between:
  - Sample complexity: want a large ρ
  - Accuracy: want a small ρ

The sample complexity of allocation

- **Theorem.** If the distribution of treatment effect values is "smooth" (ρ-regular), then we can obtain an $(1-\epsilon)$-optimal allocation with $\boldsymbol{O(M/\epsilon)}$ many samples.

ρ-regular: Probability mass around τ^* is $\approx 2\rho$

- We can do with $\sqrt{\epsilon}$ **-accurate** estimates!

$$\frac{M}{\rho^2} = \frac{M}{(\sqrt{\epsilon})^2} = \frac{M}{\epsilon}$$

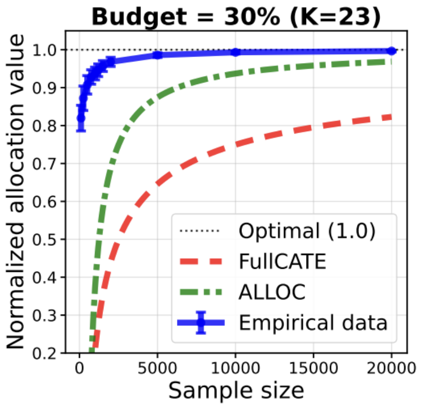
The sample complexity of allocation

- **Theorem.** For the uniform distribution of treatment effect values,

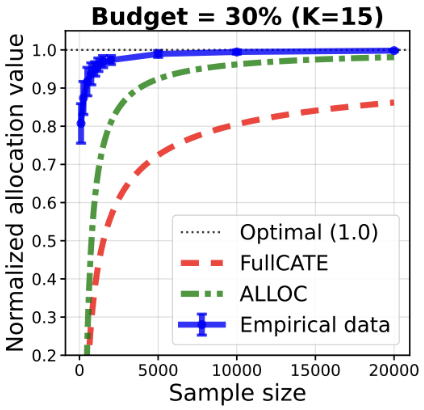
1. We can compute an $(1-\epsilon)$-optimal allocation with $O(M/\epsilon)$ samples.
2. FullCATE requires $\Omega(M/\epsilon^2)$ samples.

- Proof: minimax lower bound

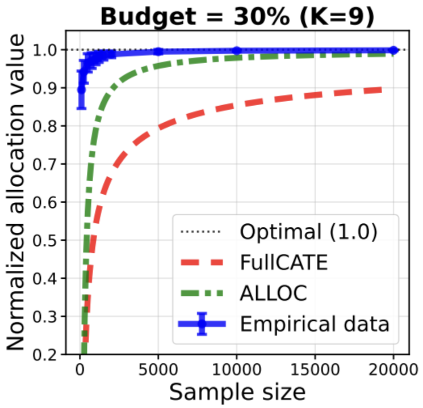
RCT experiments



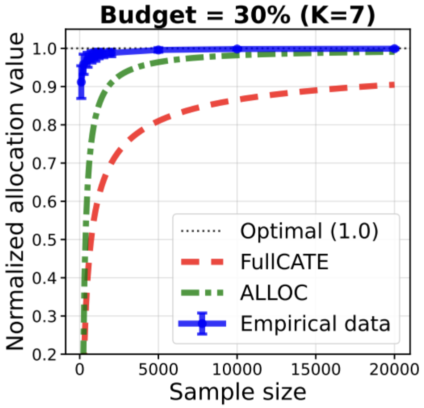
(a) STAR, schools.



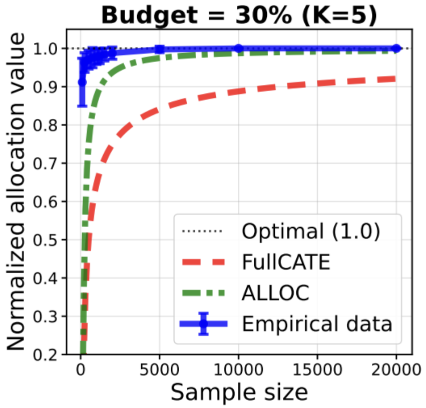
(b) TUP, baseline poverty.



(c) NSW, baseline earnings.



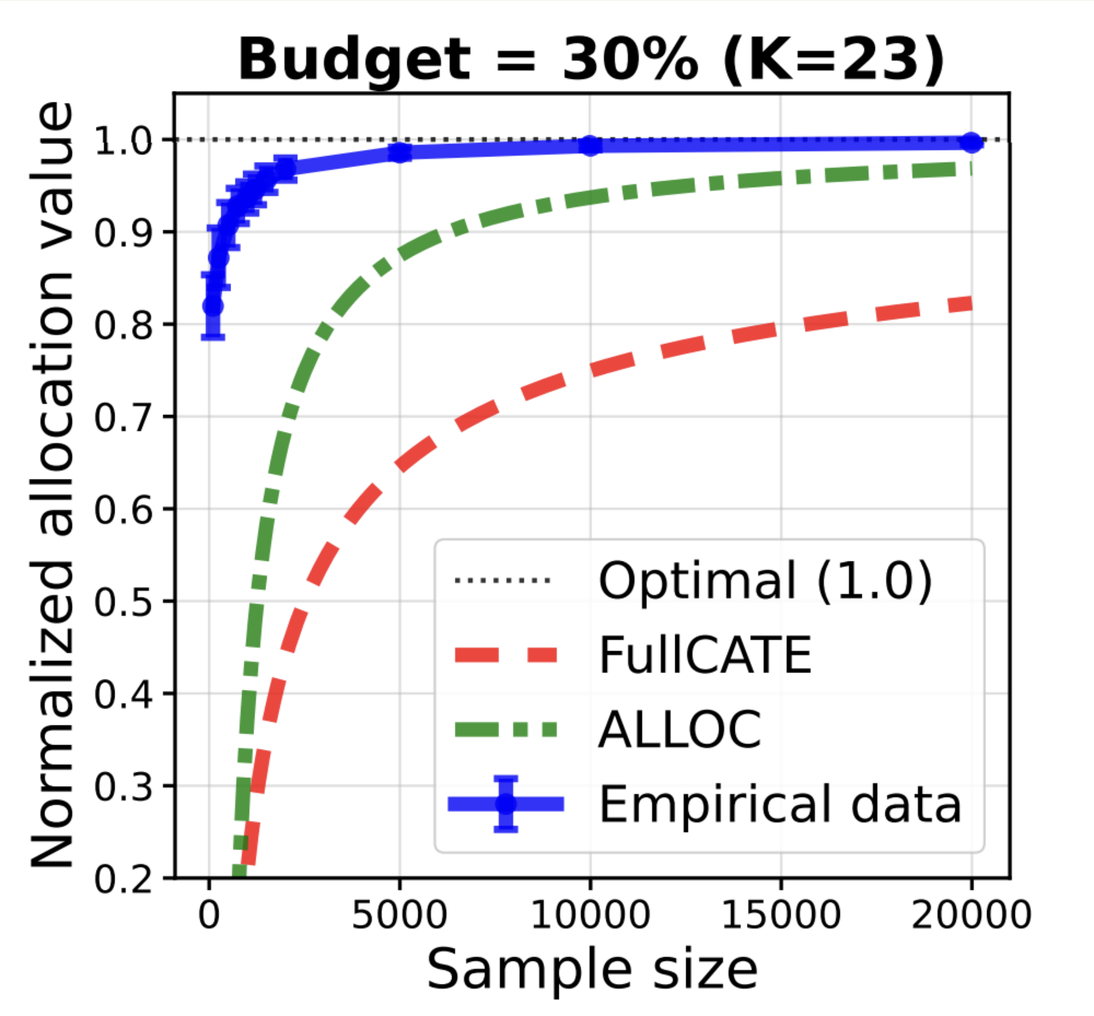
(d) Acupuncture, age.

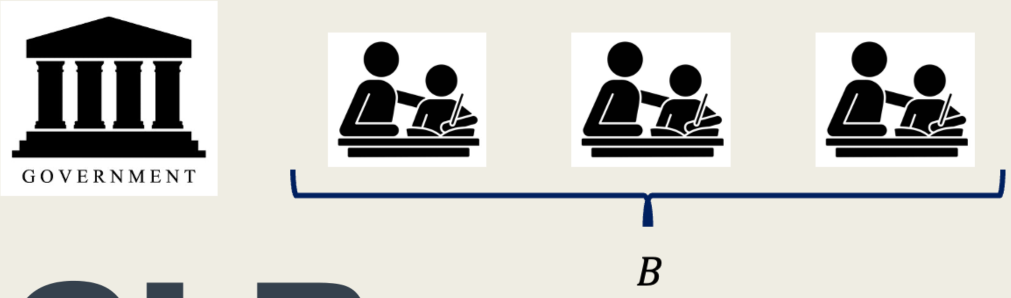


(e) Post-op, BMI.

1. STAR (*education*)
2. TUP (*economic development*)
3. NSW (*labor economics*)
4. Acupuncture (*healthcare*)
5. Post-op pain (*healthcare*)

RCT experiments





THANK YOU!

See you in Rio!