Who's Gaming the System?

A Causally-Motivated Approach for Detecting Strategic Adaptation

Trenton Chang[1], Lindsay Warrenburg[2], Sae-Hwan Park[2], Ravi B. Parikh[2,3], Maggie Makar[1], Jenna Wiens[1]

[1]Division of Computer Science & Engineering, University of Michigan

[2]Perelman School of Medicine, University of Pennsylvania [3]Winship Cancer Institute, Emory University

NeurIPS 2024

Let’s talk about American health insurance.

Medicare in the U.S.

U.S. permanent residents and citizens over 65 are eligible for government-subsidized healthcare.

63.8M

Total Enrolled*

27.5M

In *Medicare Advantage (MA)** *(private insurance, gov't reimbursement)*

**Most recent figures from Center for Medicare & Medicaid Services, 2021. Individuals under 65 with certain conditions may be eligible for Medicare, and those over 65 who have not paid taxes for sufficiently long may not qualify.*

How to game insurance

x: demographics, health history

y: reported diabetes diagnosis (yes/no)

Diabetes reimbursement rate: $1,000 / patient

CMS
CENTERS FOR MEDICARE & MEDICAID SERVICES

(U.S. government agency that administers public healthcare)

x_1	…	x_d	y
0	…	1	1
1	…	1	0
0	…	1	0
1	…	0	1
1	…	0	1
0	…	0	0
1	…	1	1

Diagnosis reports

Okay, look away for a second…

(insurance company)

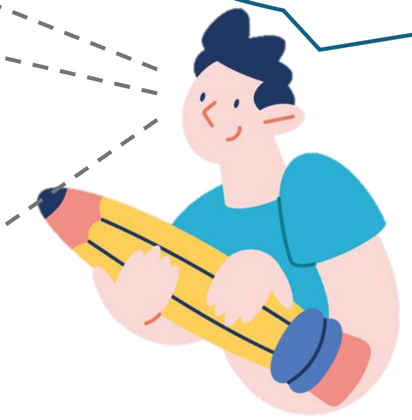
Motivating problem: insurance gaming

y: reported diabetes diagnosis

x_1	...	x_d	y
0	...	1	1
1	...	1	**1**
0	...	1	**1**
1	...	0	1
1	...	0	1
0	...	0	**1**
1	...	1	1

Diagnosis reports

More diagnoses,
more $$$!



(insurance company)

Gaming is expensive

Wall Street Journal, July 2024

EXCLUSIVE HEALTHCARE

Insurers Pocketed $50 Billion From Medicare for Diseases No Doctor Treated

Questionable diagnoses of HIV and other maladies triggered extra Medicare Advantage payments; 'It's anatomically impossible'

EXCLUSIVE HEALTHCARE

Medicare Paid Insurers Billions for Questionable Home Diagnoses, Watchdog Finds

Nurse visits said to be worth $1,869 each for Medicare Advantage companies

Wall Street Journal, Oct. 2024

			y?
			1
1	...	1	1
0	...	1	1
1	...	0	1
1	...	0	1
0	...	0	1
1	...	1	1

Goal: Identify the worst offenders

(insurance company)

Insurance plans maximize utility

$$\Delta(d_p^*; \lambda_p) = \arg\max_{\bar{d} \in [0,1]}$$

The observed diagnosis rate

Reward for increasing the diagnosis rate

Cost for deviation from ground truth

Insurance plans maximize utility

The gaming deterrence parameter:
costs *scale* differently by plan

$$\Delta(d_p^*; \lambda_p) = \arg\max_{\bar{d} \in [0,1]} R(\bar{d}) - \boxed{\lambda_p} c(\bar{d} - d_p^*)$$

Not all actions are equally
"costly" to all plans!

**however, note that the underlying cost c (and also R) are shared across plans.*

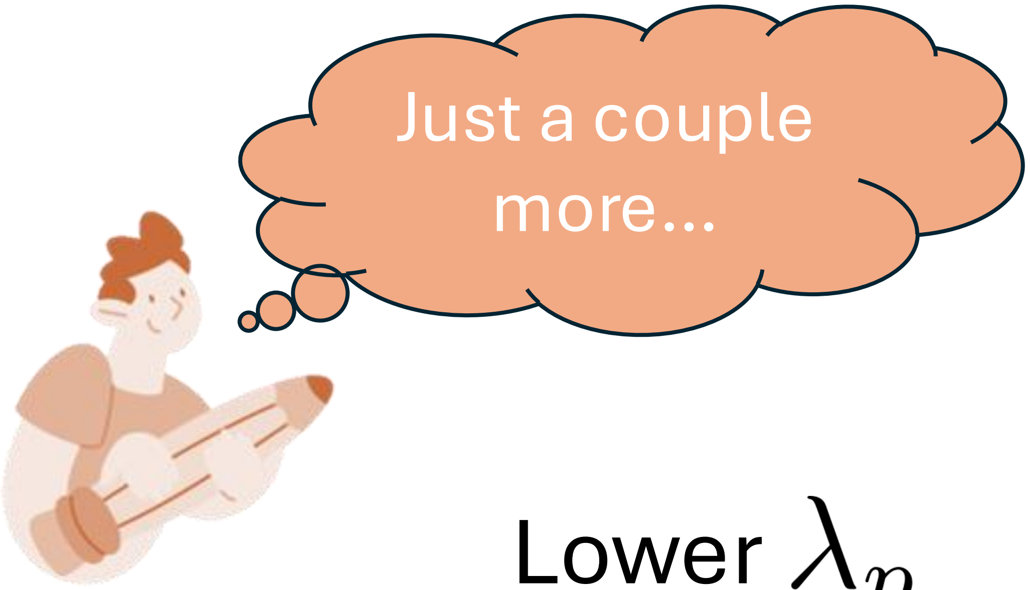
Gaming deterrence: intuition

$$\Delta(d_p^*; \lambda_p) = \arg\max_{\bar{d}\in[0,1]} R(\bar{d}) - \lambda_p c(\bar{d} - d_p^*)$$

Insurance Plan A

Higher λ_p

(more deterred from gaming)



Insurance Plan B

Lower λ_p

(less deterred from gaming)

Gaming deterrence: intuition

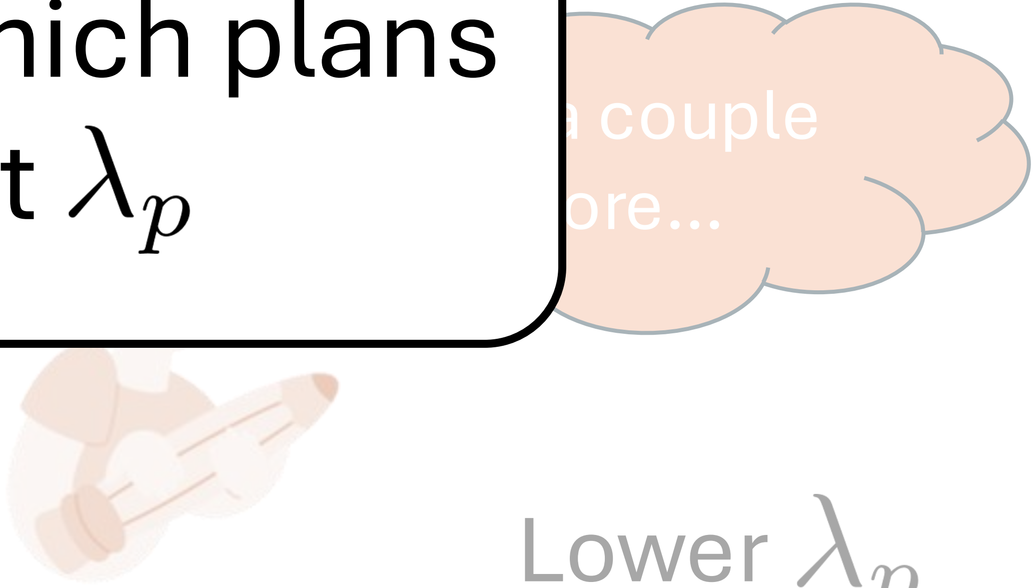
$$\Delta(d_p^*; \lambda_p) = \arg\max R(\bar{d}) - \lambda_p c(\bar{d} - d_p^*)$$

Goal: Find out which plans have the smallest λ_p

Insurance Plan A

Higher λ_p

(more deterred from gaming)



Insurance Plan B

Lower λ_p

(less deterred from gaming)

We can't tell definitively if a plan is gaming...

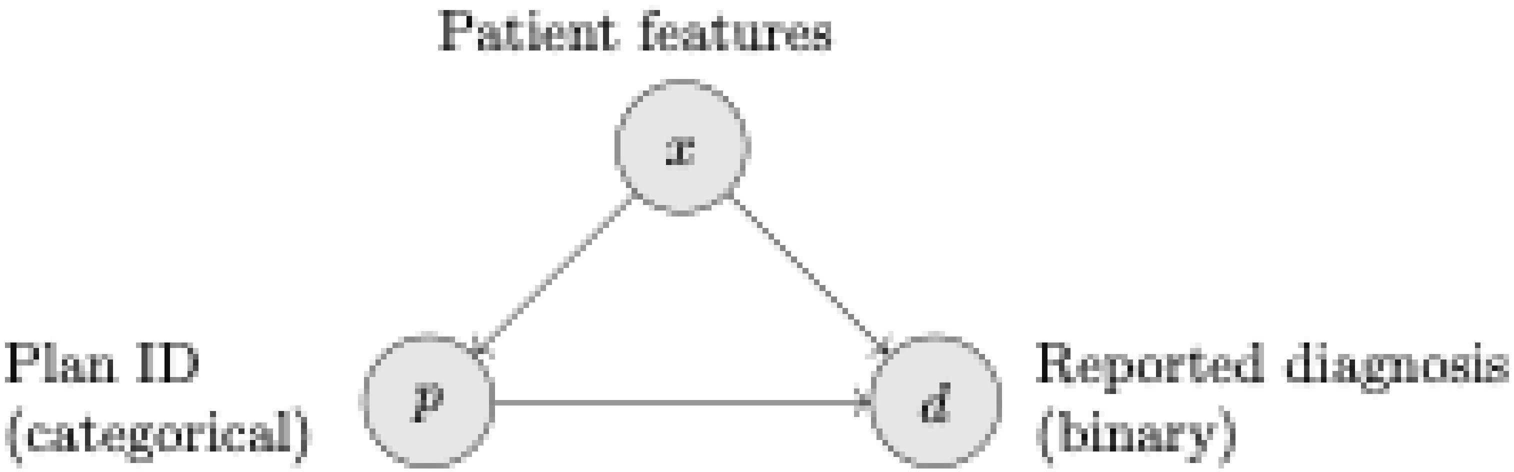
Given our assumptions about the utility function, for any agent p:

$$\lambda_p \in \left[\frac{R'(\Delta_p(d_p^*))}{c'(\Delta_p(d_p^*))}, \infty\right)$$

Every possible value of ground truth is consistent with a different value of the gaming deterrence parameter.

...but we can globally rank plans by gaming deterrence!

Given this causal graph:



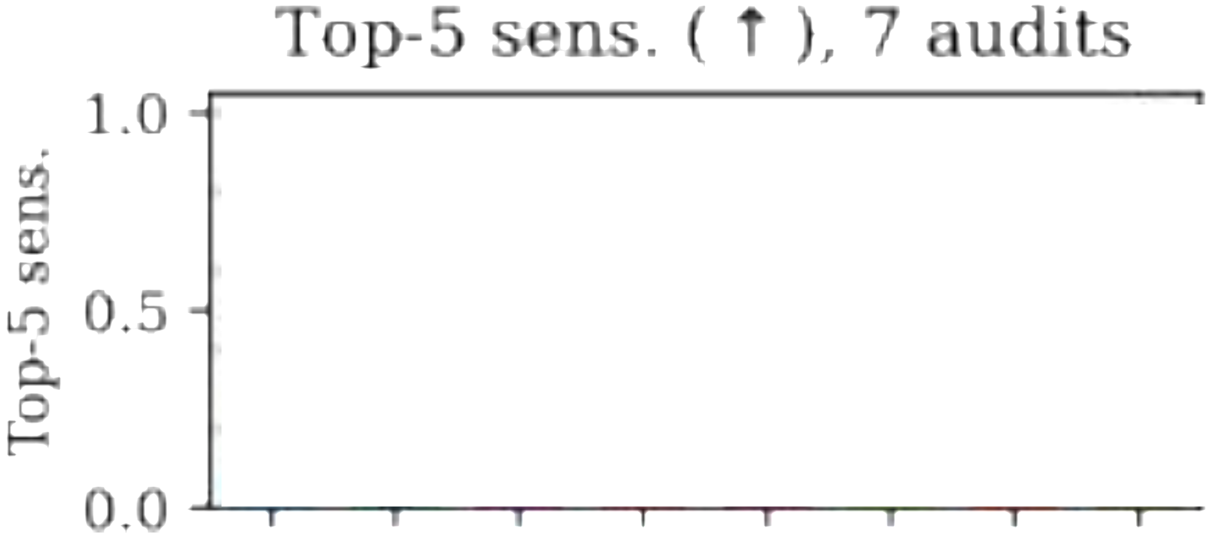
$$\lambda_{\color{blue}{PlanA}} > \lambda_{\color{orange}{PlanB}}$$

$$\Leftrightarrow$$

$$\mathbb{E}[d \mid x, p = \color{blue}{A}] - \mathbb{E}[d \mid x, p = \color{orange}{B}] < 0$$

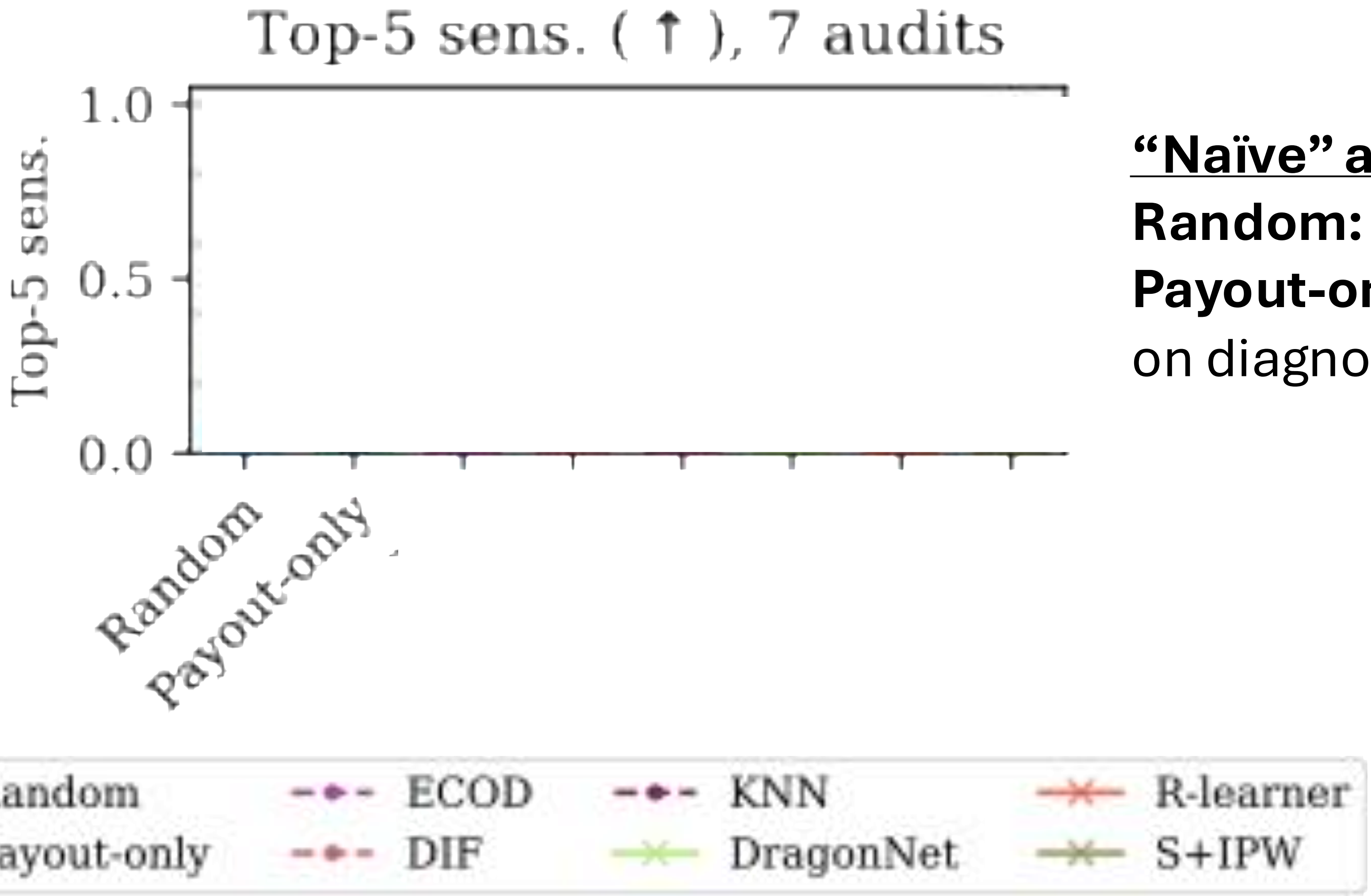
*Under the three standard assumptions for valid causal inference + assumptions used in our utility function

In simulation: our approach requires **fewer audits** to identify the **worst offenders.**



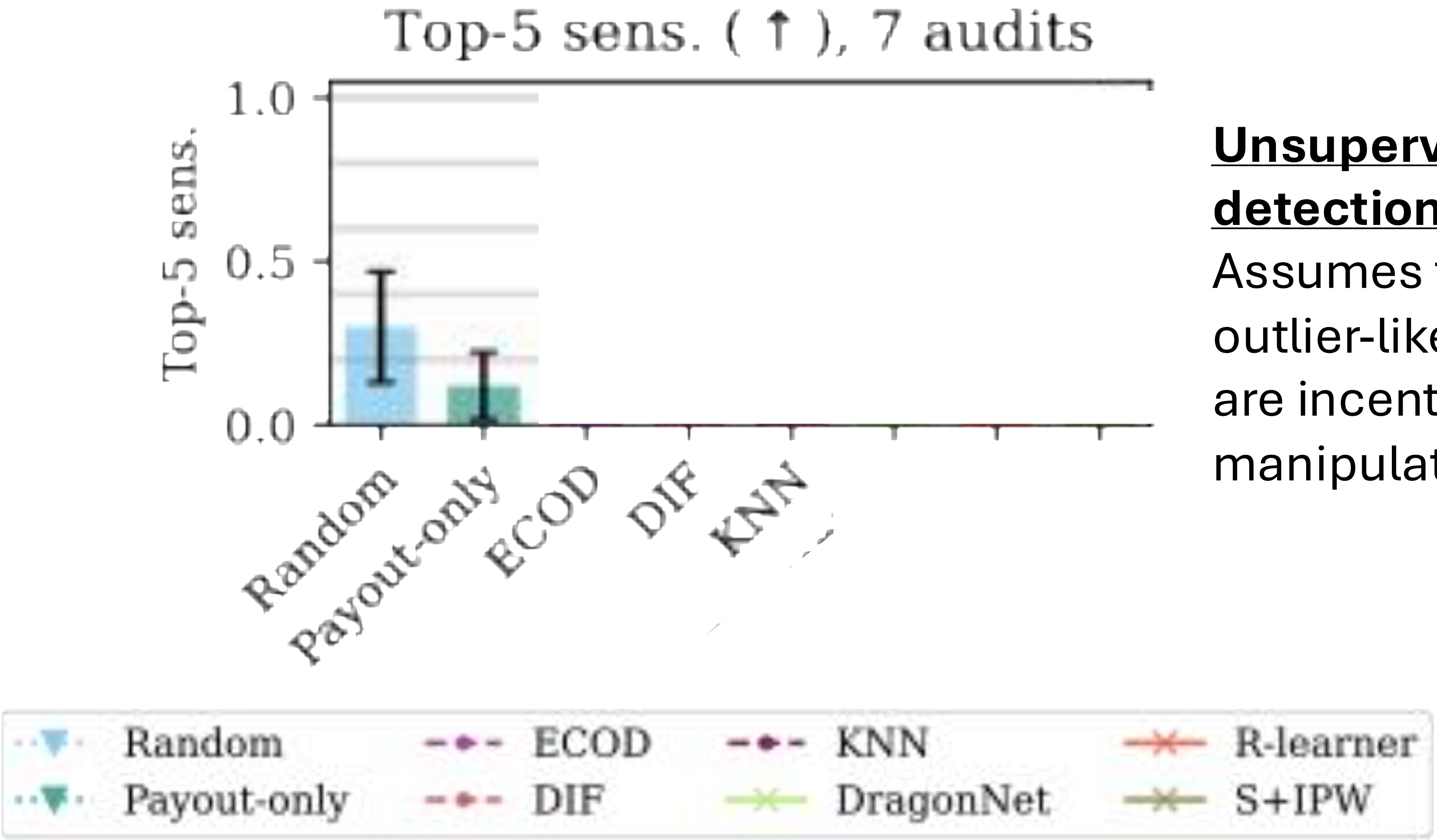
(methods on x-axis)

In simulation: our approach requires **fewer audits** to identify the **worst offenders.**



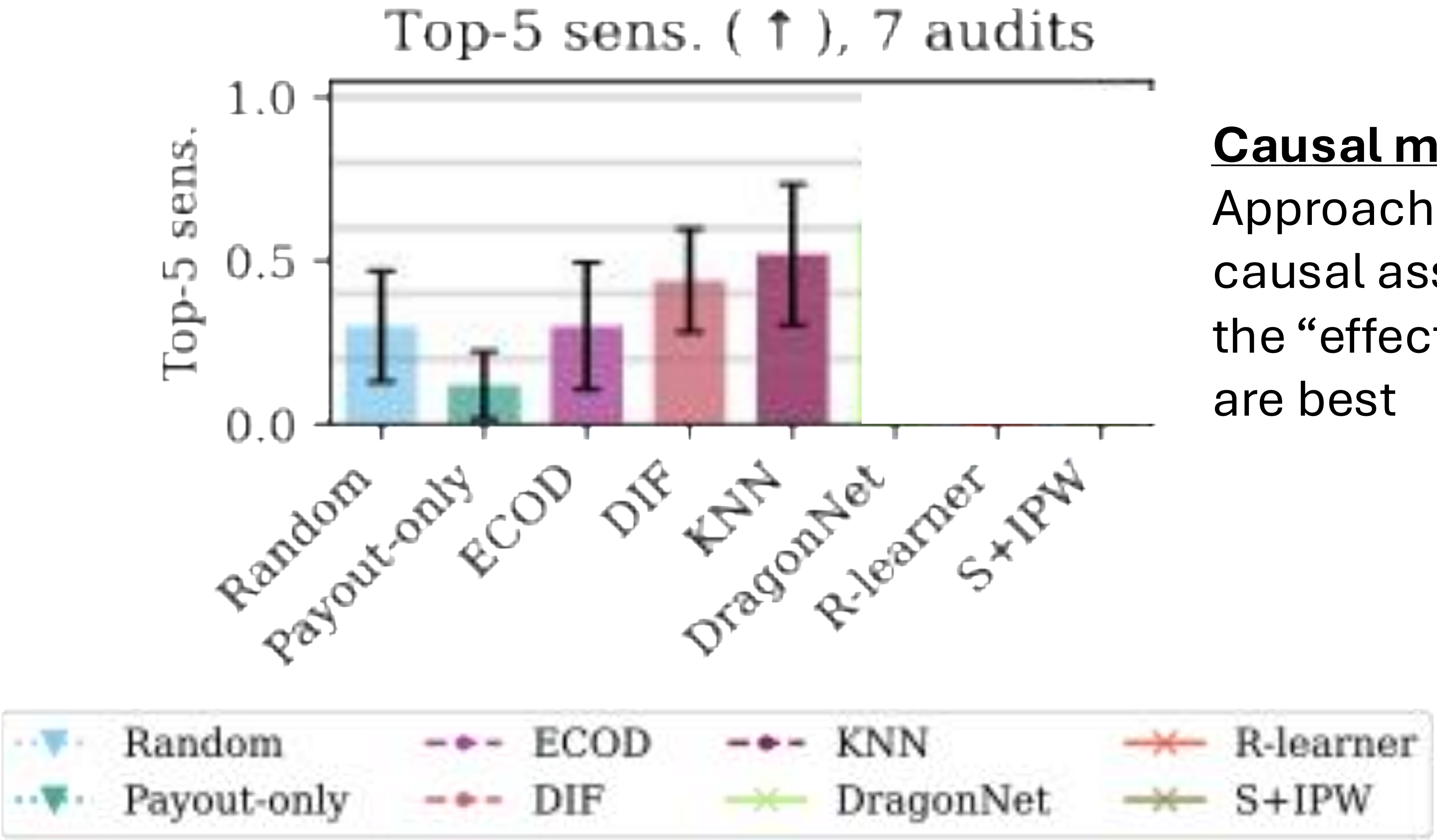
<u>“Naïve” approaches</u>
Random: audit at random
Payout-only: audit based on diagnosis rates

In simulation: our approach requires **fewer audits** to identify the **worst offenders.**



<u>Unsupervised anomaly detection</u>
Assumes that fraud is outlier-like — but *all* plans are incentivized to manipulate labels!

In simulation: our approach requires **fewer audits** to identify the **worst offenders.**



<u>Causal methods</u>
Approaches leveraging causal assumptions about the "effect" of the plan itself are best

Takeaways

- Systematic gaming be hard to detect
- Detecting gaming definitively is infeasible (without unrealistic assumptions)
- But a ranking of the gaming deterrence parameter is possible via causal effect estimation