Estimating Continuous Treatment Effects with Recourse Data

Alessandro Marchese[1] Jeroen Berrevoets[2] Niels Martin[3] Sam Verboven[1]

[1]Vrije Universiteit Brussel [2]King's College London [3]Universiteit Hasselt alessandro.marchese@vub.be AlessandroMarchese/Recourse-CADR

NeurIPS page

NEURAL INFORMATION PROCESSING SYSTEMS

TL;DR

Recourse explanations — *what would have had to change for a different outcome* — reveal where individual outcomes flip. They **identify dose-response curves even at doses never observed**, and they plug into **any neural estimator** as an auxiliary loss term.

1 The positivity problem

Goal: the dose-response $\mu(\mathbf{x}, t) = \mathbb{E}[Y(t) \mid \mathbf{X} = \mathbf{x}]$ of a continuous treatment t on a binary outcome Y.

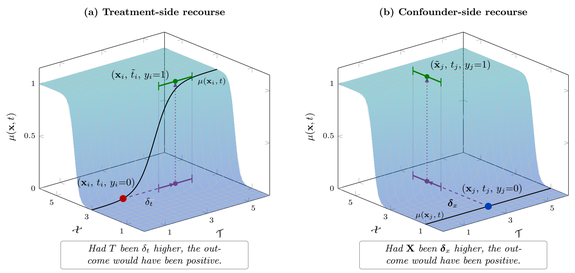
Obstacle: adjustment needs **positivity**, $p(t \mid \mathbf{x}) > 0$ at *every* dose — rarely met by routine treatment assignment in medicine or pricing.

2 Recourse data

Records often say *what would have changed the outcome*:

- *"A higher dose would have been effective."*
- *"The product had features I did not want."*
- *"The organ's cold ischemic time was too long."*

Such **counterfactual recourses** have an interventional reading: each one pins a nearby point of the response surface where the outcome flips.



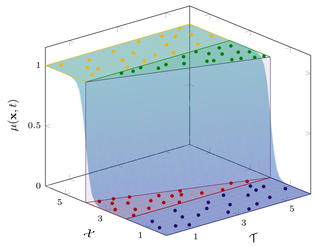
(a) A treatment-side recourse pins a point of the unit's own curve. **(b)** When no dose works, a confounder-side recourse certifies that no boundary exists within the dose range.

Formally, a recourse is the minimal-cost intervention that flips a negative outcome, under a treatment-preferential cost $\mathrm{cost}(\boldsymbol{\delta}_x, 0) \gg \mathrm{cost}(\mathbf{0}, \delta_t) > 0$:

$$\boldsymbol{\delta} = (\boldsymbol{\delta}_x, \delta_t) = \arg\min_{\boldsymbol{\delta}_x,\, \delta_t} \ \mathrm{cost}(\boldsymbol{\delta}_x, \delta_t) \quad \text{s.t.} \quad f_Y(\mathbf{x} + \boldsymbol{\delta}_x,\ t + \delta_t,\ u_Y) = 1$$

3 Recourses reveal boundaries

Threshold response $f_Y(\mathbf{x}, t, u_Y) = \mathbf{1}\{t \geq g(\mathbf{x}, u_Y)\}$ with boundary $B := g(\mathbf{X}, U_Y)$. A faithful treatment-side recourse reveals it: $b_i = t_i + \delta_{t,i}$. As the dose range is bounded, units fall into **three response strata**:



NR Never-responders
$B > T_{\max}$
$\mu^{\mathrm{NR}} \equiv 0$

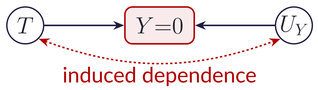
SR Sensitive responders
$T_{\min} < B \leq T_{\max}$
$\mu^{\mathrm{SR}} = F_{B|\mathbf{x},\mathrm{SR}}(t)$

AR Always-responders
$B \leq T_{\min}$
$\mu^{\mathrm{AR}} \equiv 1$

Treatment-side recourses mark SR, confounder-side ones mark NR, and AR is recovered as a residual.

4 Only failures explain

Only failures reveal boundaries, $y_i = 0 \Leftrightarrow b_i > t_i$, and conditioning on $Y = 0$ opens a collider:



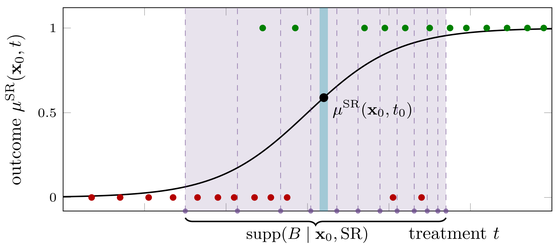
The observed boundaries are therefore **left-truncated**, favouring high thresholds:

$$p(b \mid \mathbf{x}_0, Y{=}0, \mathrm{SR}) = \frac{f_{B|\mathbf{x}_0,\mathrm{SR}}(b)\ F_{T|\mathbf{x}_0}(b^-)}{P(Y{=}0 \mid \mathbf{x}_0, \mathrm{SR})}$$

$F_{T|\mathbf{x}_0}(b^-)$ is the **Lynden–Bell selection kernel**; reweighting each revealed boundary by its inverse undoes the truncation, provided that

Recourse positivity

$F_{T|\mathbf{x}_0}(b^-) > 0$ for all $b \in \mathrm{supp}(B \mid \mathbf{x}_0, \mathrm{SR})$
treatment mass *somewhere below* each boundary — no density is needed *at* the evaluation dose itself.



No treatment is observed at t_0, yet the thresholds b_i revealed by recourses (purple) still identify $\mu^{\mathrm{SR}}(\mathbf{x}_0, t_0)$.

5 Lynden–Bell identification

Sensitive-responder CADR

Under the assumptions above, weighting each revealed boundary by $1/F_{T|\mathbf{x}_0}(b^-)$ gives

$$\mu^{\mathrm{SR}}(\mathbf{x}_0, t_0) = P(Y{=}0 \mid \mathbf{x}_0, \mathrm{SR}) \times \int_{T_{\min}}^{t_0} \frac{p(b \mid \mathbf{x}_0, Y{=}0, \mathrm{SR})}{F_{T|\mathbf{x}_0}(b^-)}\, db$$

even when $p(t_0 \mid \mathbf{x}_0) = 0$, i.e. at unobserved doses.

Adding back never-responders (from confounder-side recourses) and always-responders (the residual) gives:

Full-population CADR

$$\mu(\mathbf{x}_0, t_0) = 1 - \underbrace{P(\mathrm{CR}{=}1 \mid \mathbf{x}_0)}_{\text{never-responders}} - \underbrace{P(\mathrm{TR}{=}1 \mid \mathbf{x}_0) \int_{t_0}^{T_{\max}} \frac{p(b \mid \mathbf{x}_0, Y{=}0, \mathrm{SR})}{F_{T|\mathbf{x}_0}(b^-)}\, db}_{\text{sensitive units still failing at } t_0}$$

Every term is observable, so μ is identified even at doses with *no treatment mass at all*.

6 Auxiliary recourse loss

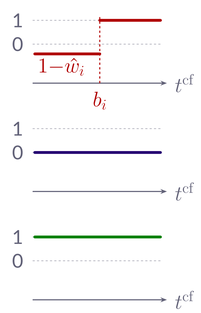
Pseudo-outcomes are evaluated at random doses, with a Lynden–Bell weight per treatment-side recourse:

$$t_j^{\mathrm{cf}} \sim \mathrm{Unif}(T_{\min}, T_{\max}) \qquad \hat{w}_i = 1/\hat{F}_{T|\mathbf{x}_i}(b_i^-)$$

TR **treatment-side**
$\tilde{y}_{ij} = 1 - \hat{w}_i\, \mathbf{1}\{t_j^{\mathrm{cf}} < b_i\}$

CR **confounder-side**
$\tilde{y}_{ij} = 0$

$y{=}1$ **positive outcome**
$\tilde{y}_{ij} = 1$



Only treatment-side recourses carry dose information; the other units anchor the curve at 0 or 1. The pseudo-outcome is unbiased, $\mathbb{E}[\tilde{y}(t) \mid \mathbf{x}] = \mu(\mathbf{x}, t)$, so the loss can be added to any neural learner:

$$\mathcal{L}_{\mathrm{Rec}}(\theta) = \frac{1}{nk} \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\hat{\mu}_\theta(\mathbf{x}_i, t_j^{\mathrm{cf}}) - \tilde{y}_{ij}\right)^2$$

$$\mathcal{L}(\theta) = \mathcal{L}_{\mathrm{Base}}(\theta) + \lambda\, \mathcal{L}_{\mathrm{Rec}}(\theta)$$

7 Experiments

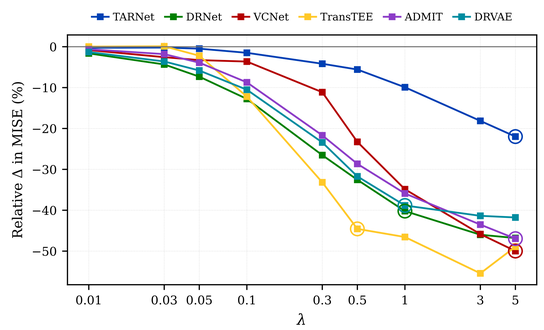
Synthetic: threshold-crossing data whose treatments cover only part of the dose range, so positivity fails while recourse positivity holds.

UNOS-PTR: organ-offer acceptance vs. cold ischemic time; refusals that cite this time are treatment-side recourses, all other refusals are confounder-side.

	Synthetic			UNOS-PTR		
MISE ↓	Base	+Rec	+LB	Base	+Rec	+LB
Recourse-only	–	.0900	**.0466**	–	.5662	**.5172**
TARNet	.4305	.1041	**.0866**	1.1628	1.1329	**1.1043**
DRNet	.4076	.0651	**.0428**	1.2510	.7436	**.4542**
VCNet	.0589	.0610	**.0312**	.8172	.8090	**.7820**
TransTEE	.0933	.0594	**.0379**	1.0486	.4853	**.3030**
ADMIT	.4015	.0735	**.0430**	1.2441	.8290	**.5609**
DRVAE	.4230	.0679	**.0466**	1.1829	.6000	**.4545**

Mean over 10 seeds. +Rec adds the recourse loss to each base learner; +LB additionally applies the Lynden–Bell weights.

8 Reweighting corrects the bias



Change in MISE from Lynden–Bell weighting across loss weights λ; the circled points mark the tuned λ of each model.

Takeaways

- Recourses are causal evidence: boundary samples and never-responder certificates.
- Identification under recourse positivity — no density needed at the evaluation dose.
- Lynden–Bell reweighting is best or tied for every model, metric and dataset.