Nonparametric Network Causal Inference for Continuous Exposures in Mobile Source Air Pollution

Salvador Balkus, Rachel Nethery, Scott Delaney, Nima Hejazi

Harvard T.H. Chan School of Public Health

Motivation

Observational studies with data $O = (W, A, Y)$ in environmental health often seek to estimate a causal effect in a setting where...

1. The exposure A is **continuous-valued** + 2. The treatment of one unit affects the outcome of others (**interference**)

How can we perform machine-learning-based causal inference with continuous exposures and arbitrary network interference?

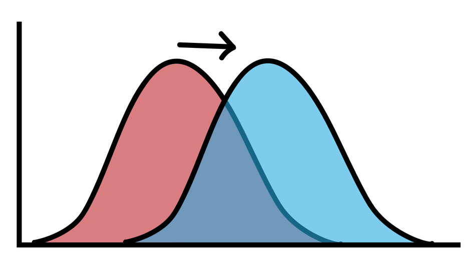
Modified Treatment Policy

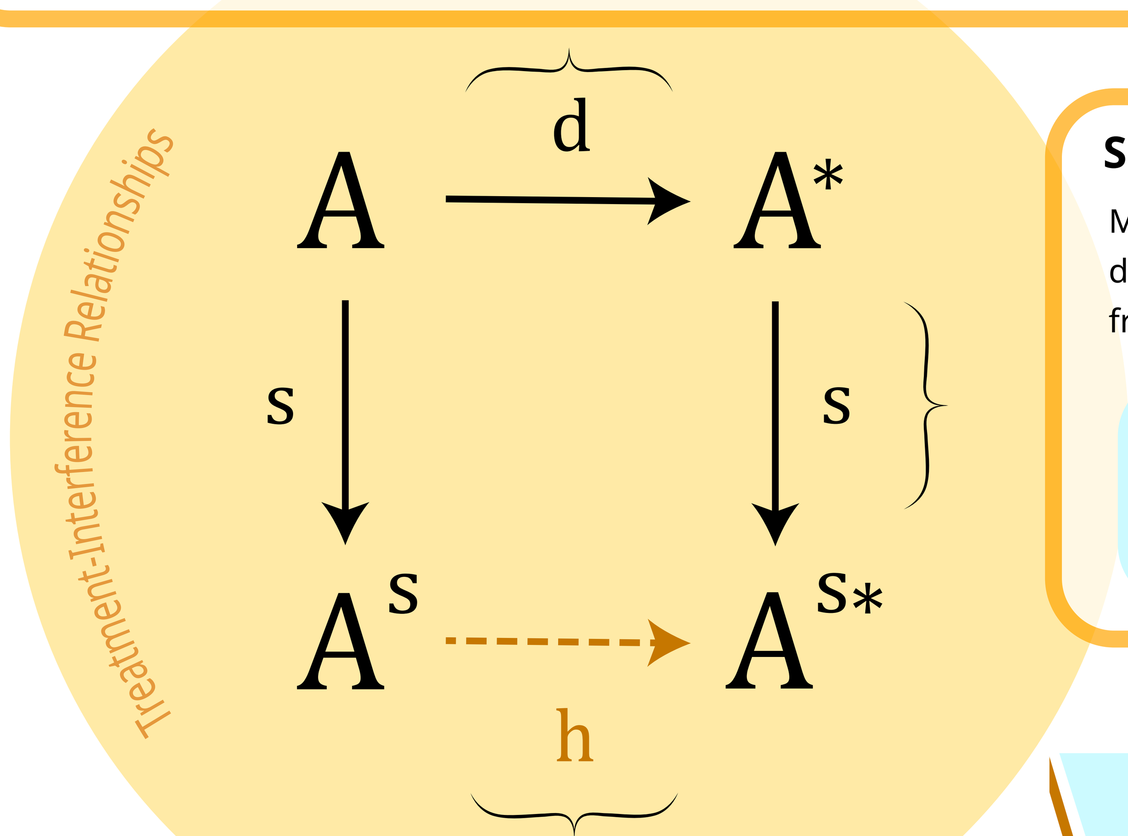
(Haneuse and Rotnizky, 2013; Diaz et al. 2018, 2021)

A causal estimand that addresses continuous-valued treatments by representing the counterfactual where some function d is applied to the natural value of exposure across all units.

Example (Additive Shift): $d(a_i, w_i; \delta) = a_i + \delta(w_i)$

Example (Multiplicative Shift): $d(a_i, w_i; \delta) = \delta(w_i) \cdot a_i$



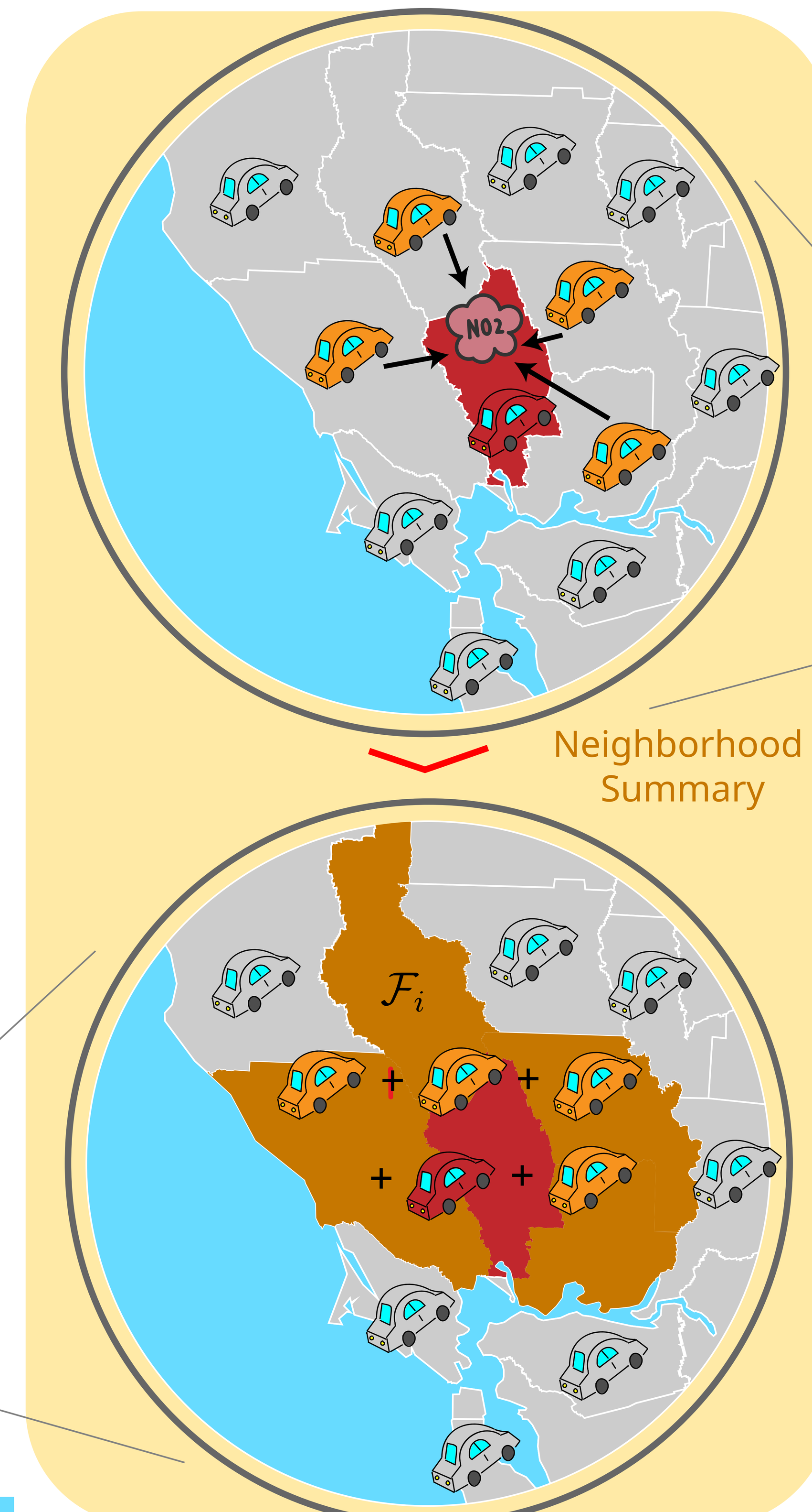


Summary Function:

Measures interference; a symmetric function describing the total exposure *received* by unit i from the set of its neighbors $\mathcal{F}_i$

Example (Weighted Sum of Neighbors):

$$s(i, \mathbf{a}, \mathbf{w}) = \sum_{j \in \mathcal{F}_i} w_j a_j$$



Induced Modified Treatment Policy:

A function h satisfying $(h \circ s)(a_i, w_i; \delta) = (s \circ d)(a_i, w_i; \delta)$ that is piecewise invertible (required to form a valid MTP)

Q: When does such a function exist?

Identifying Assumptions

(1) **Consistency:** $A^s = a^s \implies Y^{a^s} = Y$

(2) **Positivity:** $(a^s, w^s) \in \text{supp}(A^s, W^s) \implies (a^{s*}, w^s) \in \text{supp}(A^s, W^s)$

(3) **Conditional exchangeability*:** $Y^{h(a^s, \mathbf{w}; \delta)} \perp\!\!\!\perp A^s | W^s$ (*may be relaxed)

(4) **Piecewise smooth invertibility:** h has a piecewise differentiable inverse

Under (1)-(4), the Induced MTP is identified

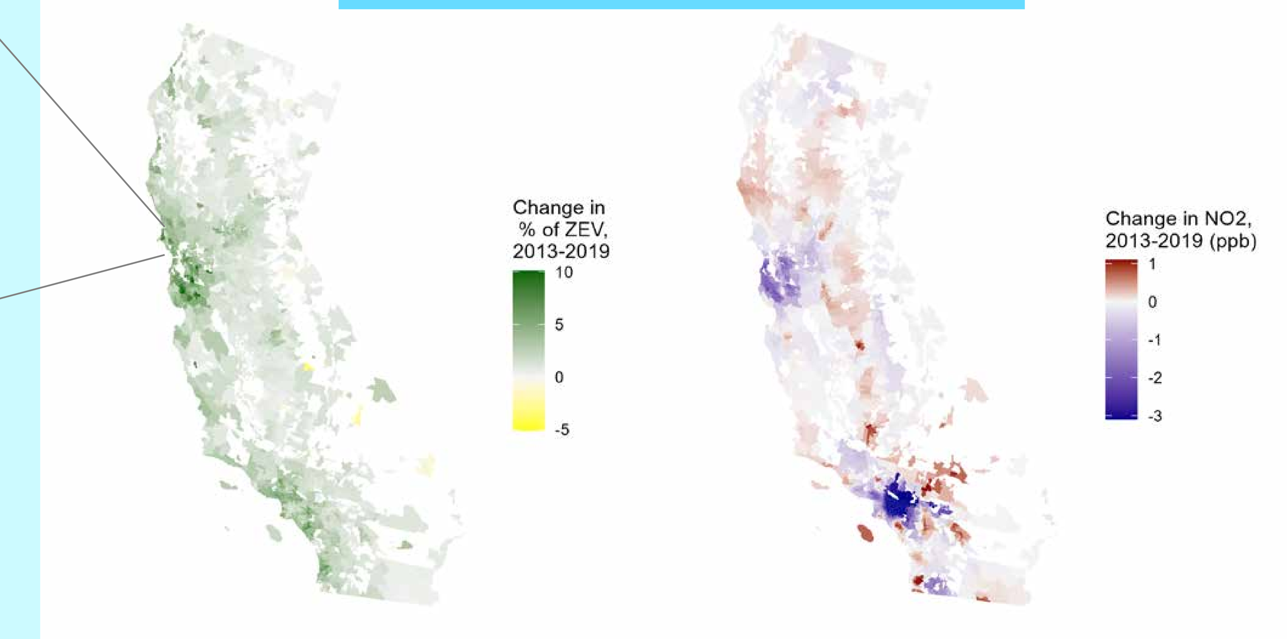
$$\psi = E(Y^{h(a^s, \mathbf{w}; \delta)}) = \int E(Y | A^s = h(a^s, \mathbf{w}; \delta), W^s) dP(A^s, W^s)$$

and can be efficiently estimated nonparametrically.

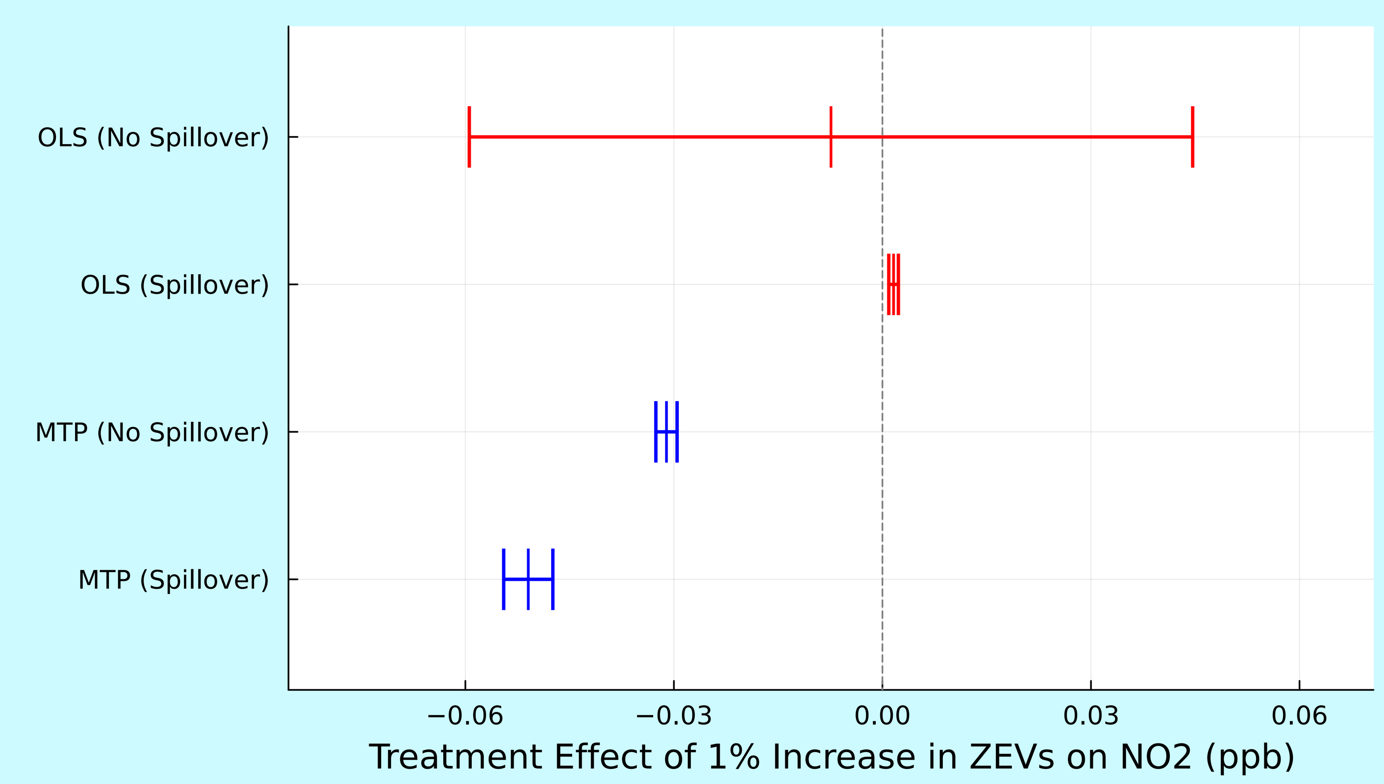
Data Application

Evaluating the effect of "zero-emission vehicle" (ZEV) uptake on NO2 pollution in California (unit of analysis: ZIP Code Tabulation Area)

Change in ZEV and NO2 across California



- Controlled for socioeconomic factors, road density, walkability, industrial zoning, and transportation use variables.
- MTP Estimation: Super Learning (*van der Laan et al. 2007*) with ElasticNet, Random Forest, and LGBM for outcome regression, KLIEP (*Sugiyama et al. 2007*) for density ratio.
- Covariate-conditional variance reported.



Theoretical Results

Definition (Selector): s is a selector if $s(i, \mathbf{a}, \mathbf{w}) = f(a_j, w_j)$ for some $a_j \in \mathcal{F}_i$ and f is piecewise smooth invertible.

Example (Maximum): $s(i, \mathbf{a}, \mathbf{w}) = \max\{a_j \in \mathcal{F}_i\}$

Definition (Aggregator): If $(\mathcal{A}, +, \cdot)$ is a field, and f, g field homomorphisms, s is an aggregator if $s(i, \mathbf{a}, \mathbf{w}) = g\Big(\sum_{j \in \mathcal{F}_i} f(a_j, w_j)\Big)$

Example (Weighted Mean): $s(i, \mathbf{a}, \mathbf{w}) = \frac{1}{|\mathcal{F}_i|} \sum_{j \in \mathcal{F}_i} w_j a_j$

Theorem 1

An induced modified treatment policy is piecewise smooth invertible (satisfying assumption 4) if either:

1. s is a selector **OR**
2. s is an aggregator (or a composition of them) with differentiable operations and the unit-level MTP takes the form

$$d(a_i, w_i; \delta) = \sum_j \Big(\delta \cdot a_i + q(w_i, \delta)\Big) \cdot 1(a_i \in \mathcal{A}_j)$$

Estimation

Nuisance Parameters

Outcome Regression: $\bar{Q}_i(a_i^s, w_i^s) = E(Y | A^s = a_i^s, W^s = w_i^s)$

Density Ratio: $H_i = \frac{p(h^{-1}(a_i^s, \mathbf{w}; \delta) | w_i^s)}{p(a_i^s | w_i^s)} \cdot \frac{d}{da^s} h^{-1}(a^s, \mathbf{w}; \delta)\Big|_{a^s = a_i^s}$

Estimators

Plug-In: $\hat{\psi}_n = \frac{1}{n} \sum_{i=1}^n \bar{Q}_i(a_i^s, w_i^s)$

IPW: $\hat{\psi}_n = \frac{1}{n} \sum_{i=1}^n H_i Y_i$

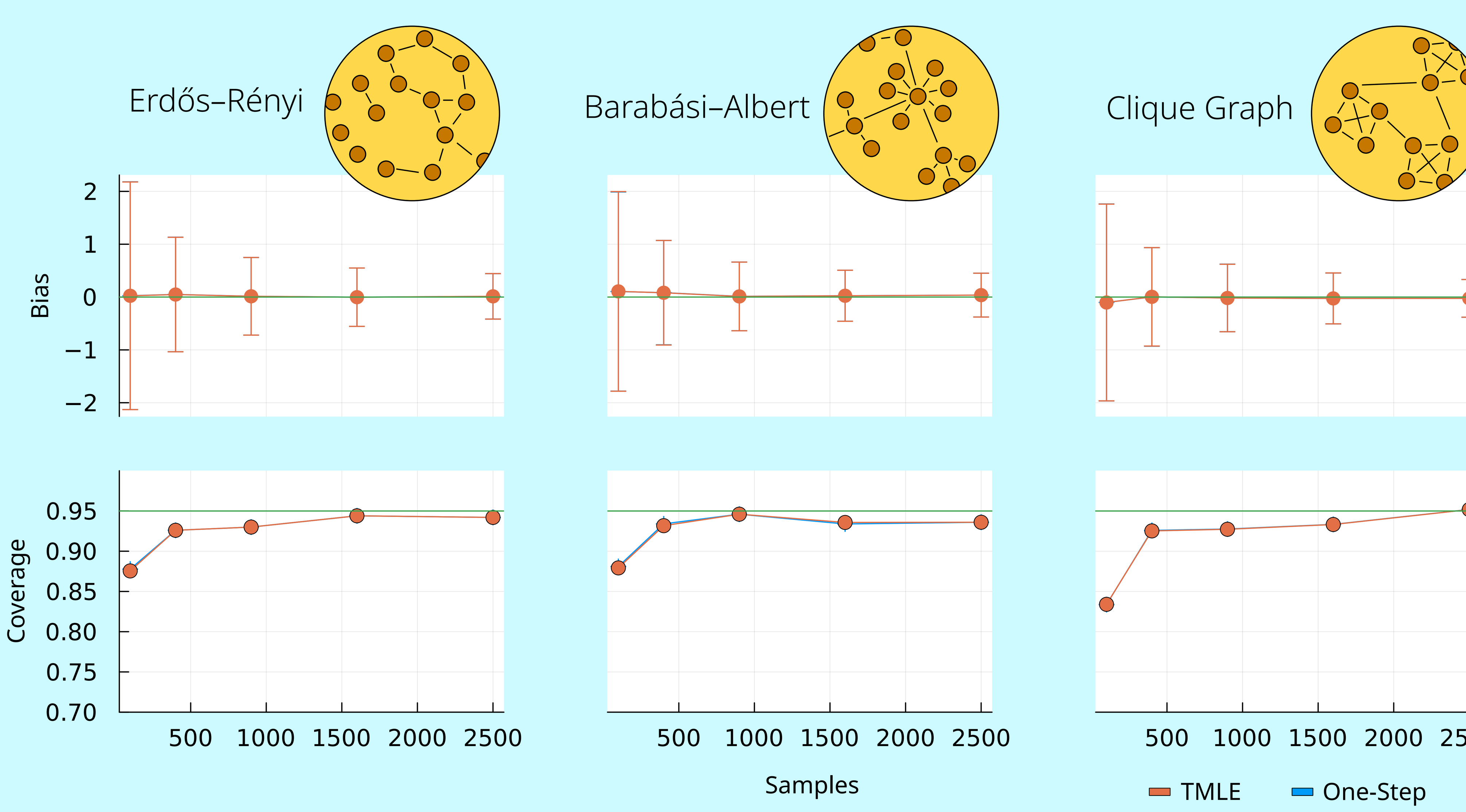
One-Step: $\hat{\psi}_n = \frac{1}{n} \sum_{i=1}^n H_i(Y_i - \bar{Q}_i(a_i^s, w_i^s)) + \bar{Q}_i(h(a_i^s; w_i^s, \delta))$

TMLE: $\hat{\psi}_n = \frac{1}{n} \sum_{i=1}^n \bar{Q}_i^\epsilon$ where $\bar{Q}_i^\epsilon$ fit via submodel fluctuation

Semiparametric Efficient & Doubly Robust Consistent

Simulations

- Evaluated Sum and Additive Shift with outcome a combination of nonlinear functions of 8 covariates + sine function treatment.
- Estimators: TMLE and One-Step with Super Learning for conditional mean, KLIEP for density ratio.
- 500 iterations for each sample size; EIF-based variance estimator for correlated treatment from Ogburn et al. (2022).



This work supported by NIH/NIEHS (T32ES007142) and NSF (DGE 2140743)