UF | UNIVERSITY of FLORIDA

Bipartite causal inference with interference, time series data, and a random network

Zhaoyan Song

with Georgia Papadogeorgou

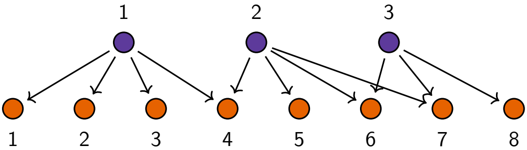
Spatiotemporal Causality Reading Group – May 14, 2025

Causal inference in observational settings with bipartite interference

- Causal inference provides the framework for studying the effect of an intervention on a population of interest
- Randomization is considered the gold standard for drawing causal inferences
  $\rightsquigarrow$ It can be unethical or impractical to implement
  $\rightsquigarrow$ Need to develop methodologies for observational data
- Methods can suffer due to units' causal interconnectedness
  $\rightsquigarrow$ Interference: a unit's potential outcomes depend on their own treatment, but also on the treatment of others
  $\rightsquigarrow$ Complicates how causal effects are defined and estimated

Causal inference in observational settings with bipartite interference

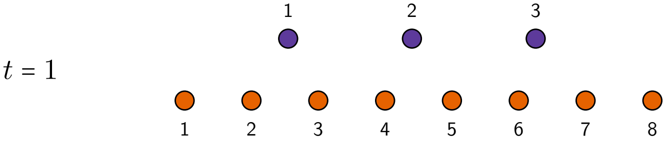
- In bipartite causal inference there are two distinct groups of units:
  - **interventional units** which are assigned treatment / control, and
  - **outcome units** on which we measure the outcome
- Units are connected through a bipartite network



Bipartite interference network with 3 interventional units and 8 outcome units.

Our contribution

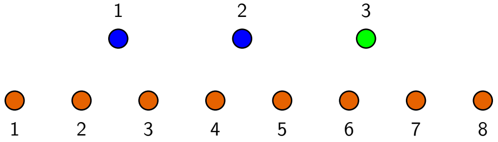
- We develop a causal inference framework for bipartite interference with time series observational data and a random network



Our contribution

- We develop a causal inference framework for bipartite interference with time series observational data and a random network

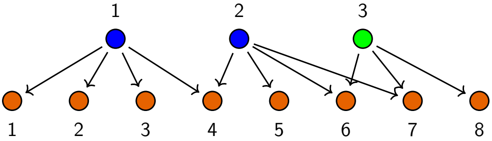
$t = 1$



Our contribution

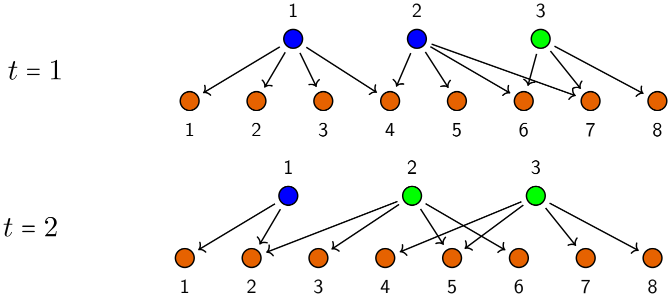
- We develop a causal inference framework for bipartite interference with time series observational data and a random network

$t = 1$



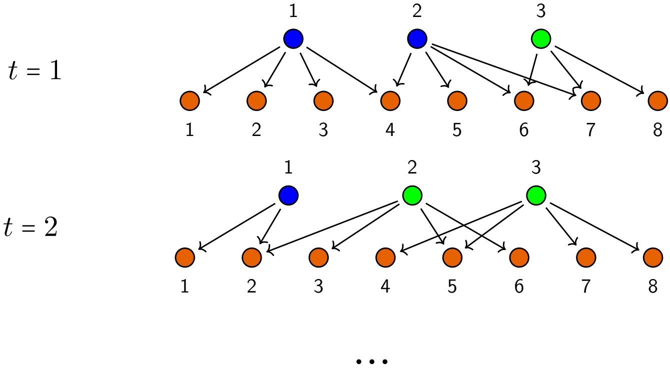
Our contribution

- We develop a causal inference framework for bipartite interference with time series observational data and a random network



Our contribution

- We develop a causal inference framework for bipartite interference with time series observational data and a random network



Existing literature

- Most work on causal inference assumes no interference
- Causal inference with interference is most often cross-sectional and considers a fixed and known network
  $\rightsquigarrow$ Bipartite interference: *e.g.*, Zigler and Papadogeorgou (2021); Harshaw et al. (2023); Pouget-Abadie et al. (2019)
- Some recent work studies uncertainty of the unknown network in a cross-sectional study
  $\rightsquigarrow$ Wikle and Zigler (2023)
- Causal inference with panel data generally assumes no interference
  $\rightsquigarrow$ Exceptions: *e.g.*, Grossi et al. (2020); Menchetti and Bojinov (2020)

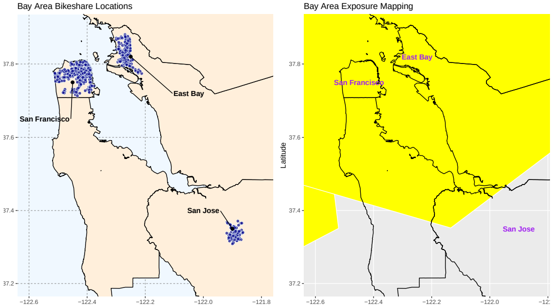
Our motivating context

The effect of wildfire smoke on transportation by bicycle

- The units correspond to geographical areas
  - **Interventional units** are forested areas where a wildfire might take place
  - **Outcome units** are urban areas where bicycle usage is measured
- The outcome units experience smoke from wildfires originating from interventional units according to a random network driven by geographic, atmospheric and weather conditions
- What is the effect of smoke from wildfires on total bikeshare time?

Our motivating context

The effect of wildfire smoke on transportation by bicycle



- Three outcome units in SF Bay area (SF, East Bay, San Jose)

  $\rightsquigarrow$ Outcome: Total riding time using Lyft's bikeshare program
- NOAA's Hazard Mapping System combines information on wildfires and smoke transport to deduct smoke exposure

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti}
- Time-invariant covariates $\boldsymbol{X}_i^*$
- Time-varying covariates $\boldsymbol{X}_{ti}$

- $\boldsymbol{A}_t : (A_{t1} \ A_{t2} \ \dots \ A_{tN})$
- $\boldsymbol{X}^* : (\boldsymbol{X}_1^{*\top} \ \boldsymbol{X}_2^{*\top} \ \dots \ \boldsymbol{X}_N^{*\top})^\top$
- $\boldsymbol{X}_t : (\boldsymbol{X}_{t1}^\top \ \boldsymbol{X}_{t2}^\top \ \dots \ \boldsymbol{X}_{tN}^\top)^\top$

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti} · presence/absence of wildfire
- Time-invariant covariates $\boldsymbol{X}_i^*$ · e.g. type of vegetation
- Time-varying covariates $\boldsymbol{X}_{ti}$ · e.g. humidity

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti} — presence/absence of wildfire
- Time-invariant covariates $\boldsymbol{X}_i^*$ — e.g. type of vegetation
- Time-varying covariates $\boldsymbol{X}_{ti}$ — e.g. humidity

- Outcome units $\mathcal{M} = \{m_1, m_2, \cdots, m_M\}$.

m_j

- Outcome Y_{tj}
- Time-invariant covariates $\boldsymbol{W}_j^*$ — $\boldsymbol{W}^* : (\boldsymbol{W}_1^{*\top}\ \boldsymbol{W}_2^{*\top}\ \ldots\ \boldsymbol{W}_N^{*\top})^\top$
- Time-varying covariates $\boldsymbol{W}_{tj}$ — $\boldsymbol{W}_t : (\boldsymbol{W}_{t1}^\top\ \boldsymbol{W}_{t2}^\top\ \ldots\ \boldsymbol{W}_{tN}^\top)^\top$

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti} · presence/absence of wildfire
- Time-invariant covariates $\boldsymbol{X}_i^*$ · e.g. type of vegetation
- Time-varying covariates $\boldsymbol{X}_{ti}$ · e.g. humidity

- Outcome units $\mathcal{M} = \{m_1, m_2, \cdots, m_M\}$.

m_j

- Outcome Y_{tj}
- Time-invariant covariates $\boldsymbol{W}_j^*$ · e.g. demographic information
- Time-varying covariates $\boldsymbol{W}_{tj}$ · e.g. temperature

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti} · presence/absence of wildfire
- Time-invariant covariates $\boldsymbol{X}_i^*$ · e.g. type of vegetation
- Time-varying covariates $\boldsymbol{X}_{ti}$ · e.g. humidity

- Outcome units $\mathcal{M} = \{m_1, m_2, \cdots, m_M\}$.

m_j

- Outcome Y_{tj}
- Time-invariant covariates $\boldsymbol{W}_j^*$ · e.g. demographic information
- Time-varying covariates $\boldsymbol{W}_{tj}$ · e.g. temperature

- Network
  - $G_{tij} \in \{0\ 1\}$: $n_i \overset{?}{\rightarrow} m_j$ · $\boldsymbol{G}_t$ interference network
  - Interaction time-invariant covariates · $\boldsymbol{P}^* : \{P_{ijs}^*\}$
  - Interaction time-varying covariates · $\boldsymbol{P}_t : \{P_{tijs}\}$

Notation

- Time: $t \in \{1, 2, \cdots, T\}$. Smooth temporal trend: $f(t)$
- Interventional units $\mathcal{N} = \{n_1, n_2, \cdots, n_N\}$.

n_i

- Treatment A_{ti} · presence/absence of wildfire
- Time-invariant covariates $\boldsymbol{X}_i^*$ · e.g. type of vegetation
- Time-varying covariates $\boldsymbol{X}_{ti}$ · e.g. humidity

- Outcome units $\mathcal{M} = \{m_1, m_2, \cdots, m_M\}$.

m_j

- Outcome Y_{tj}
- Time-invariant covariates $\boldsymbol{W}_j^*$ · e.g. demographic information
- Time-varying covariates $\boldsymbol{W}_{tj}$ · e.g. temperature

- Network
  - $G_{tij} \in \{0\ 1\}$: $n_i \overset{?}{\rightarrow} m_j$ · based on e.g. wind patterns
  - Interaction time-invariant covariates · e.g. relative distance of areas
  - Interaction time-varying covariates · e.g. traffic in given day

Estimands

- Potential outcomes: $Y_{tj}(\boldsymbol{a}_t, \boldsymbol{g}_{t.j})$

- Assumption (Exposure mapping) There exists function $h_{tj} : \mathcal{A} \times \mathcal{G}_{.j} \to \mathcal{E}$ for which if $h_{tj}(\boldsymbol{a}_t, \boldsymbol{g}_{t.j}) = h_{tj}(\boldsymbol{a}'_t, \boldsymbol{g}'_{t.j})$, then $Y_{tj}(\boldsymbol{a}_t, \boldsymbol{g}_{t.j}) = Y_{tj}(\boldsymbol{a}'_t, \boldsymbol{g}'_{t.j})$

  $\rightsquigarrow$ Potential outcomes can be denoted as $Y_{tj}(e_{tj})$

- The estimand is specific to each outcome unit:

$$\widetilde{\tau}_j(e, e') = \frac{1}{\sum_t I(E_{tj} = e)} \sum_t \left(Y_{tj}(e) - Y_{tj}(e')\right) I(E_{tj} = e)$$

  $\rightsquigarrow$ Advantages over estimands that average over units (we will return to this later)

Unconfoundedness

- *Assumption*
  - Unconfoundedness of the interventional units' treatment assignment

$$\boldsymbol{A}_t \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j}.$$

  - Unconfoundedness of the bipartite network

$$\boldsymbol{G}_{t.j} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j}.$$

Unconfoundedness

- *Assumption*
  - Unconfoundedness of the interventional units' treatment assignment

$$\boldsymbol{A}_t \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j}.$$

  - Unconfoundedness of the bipartite network

$$\boldsymbol{G}_{t.j} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j}.$$

Proposition 1

An outcome unit's exposure assignment is conditionally independent of potential outcomes

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j}.$$

Unconfoundedness

- Essentially all causal inference with interference that uses exposure mappings *assumes* exposure unconfoundedness
- This formalization of exposure unconfoundedness
  - acknowledges the bipartite context
  - places assumptions on the processes that give rise to a unit's exposure
  - provides practical insights for identifying confounders
  - renders confounding adjustment more tangible and actionable in the bipartite setting

Estimation based on matching time periods

Focusing on binary exposures

- One estimand for each outcome unit that averages across time
- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- Time-invariant covariates are implicitly "matched" across time
  $\rightsquigarrow$ Nothing to do for potentially high-dimensional $\boldsymbol{X}^*, \boldsymbol{W}^*, \boldsymbol{P}^*$

Estimation based on matching time periods

Focusing on binary exposures

- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- We propose matching algorithms that match time periods for **one** outcome unit, as follows

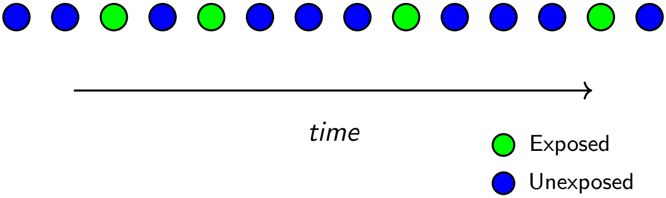
Estimation based on matching time periods

Focusing on binary exposures

- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- We propose matching algorithms that match time periods for **one** outcome unit, as follows



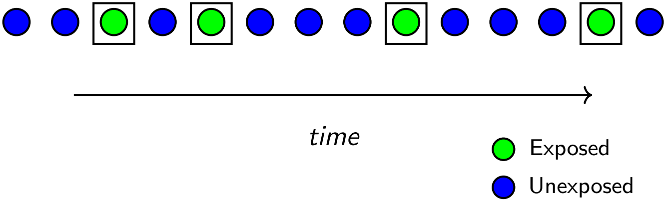
Estimation based on matching time periods

Focusing on binary exposures

- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- We propose matching algorithms that match time periods for **one** outcome unit, as follows



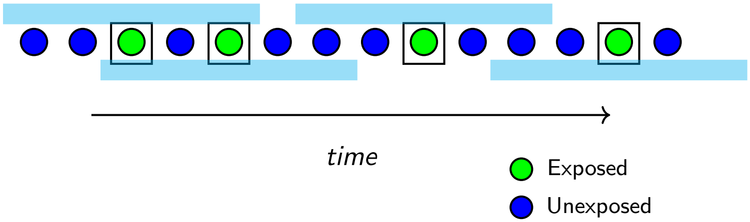
Estimation based on matching time periods

Focusing on binary exposures

- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- We propose matching algorithms that match time periods for **one** outcome unit, as follows



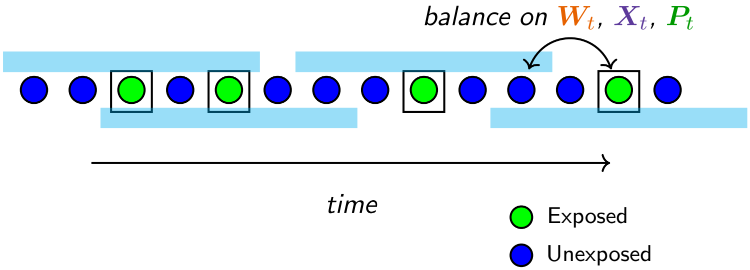
Estimation based on matching time periods

Focusing on binary exposures

- Estimation under the exposure unconfoundedness

$$E_{tj} \perp\!\!\!\perp \mathcal{Y}_{tj}(\cdot) \mid f(t), \boldsymbol{X}^*, \boldsymbol{X}_{t.}, \boldsymbol{W}^*, \boldsymbol{W}_{tj}, \boldsymbol{P}^*, \boldsymbol{P}_{t.j.}$$

- We propose matching algorithms that match time periods for **one** outcome unit, as follows



The matching algorithms

- Maximize number of matched exposed time periods
- Satisfy balance constraints on
  - ⇝ time to control temporal trends $f(t)$
  - ⇝ time-varying covariates of the outcome unit $\boldsymbol{W}_t$
  - ⇝ summaries of time-varying covariates of interventional units $\boldsymbol{X}_t$
  - ⇝ summaries of interaction covariates $\boldsymbol{P}_t$
- Match an exposed time periods to 1, 2 or (1 or 2) unexposed time periods

---

Connections with matching and balancing methods in the literature (e.g., Zubizarreta, 2012; Keele et al., 2014; Hainmueller, 2012)

Matching example: Matching 1-1 approach

(A) $$\max_{\boldsymbol{a}} \sum_{t_e, t_u} a_{t_e t_u},$$

(A.1) $$\sum_{t_u} a_{t_e t_u} \leq 1, \quad \forall t_e \in \mathcal{T}_e, \quad \text{and} \quad \sum_{t_e} a_{t_e t_u} \leq 1, \quad \forall t_u \in \mathcal{T}_u.$$

(A.2) $$\left| \sum_{t_e, t_u} a_{t_e t_u} (t_e - t_u) \right| \leq \delta \sum_{t_e, t_u} a_{t_e t_u}.$$

(A.3) $$|a_{t_e t_u} (t_e - t_u)| \leq \epsilon, \quad \forall t_e \in \mathcal{T}_e, \forall t_u \in \mathcal{T}_u.$$

(A.4) $$\left| \sum_{t_e, t_u} a_{t_e t_u} (\boldsymbol{W}_{t_e} - \boldsymbol{W}_{t_u}) \right| \leq \mathbf{1}_{p_W} \cdot \delta' \sum_{t_e, t_u} a_{t_e t_u},$$

(A.5) $$\left| \sum_{t_e, t_u} a_{t_e t_u} (\widetilde{\boldsymbol{X}}_{t_e} - \widetilde{\boldsymbol{X}}_{t_u}) \right| \leq \mathbf{1}_{p_X} \cdot \delta' \sum_{t_e, t_u} a_{t_e t_u}, \quad \text{and}$$

$$\left| \sum_{t_e, t_u} a_{t_e t_u} (\widetilde{\boldsymbol{P}}_{t_e} - \widetilde{\boldsymbol{P}}_{t_u}) \right| \leq \mathbf{1}_{p_P} \cdot \delta' \sum_{t_e, t_u} a_{t_e t_u}.$$

The matching estimator

- For exposed time period t matched to the unexposed t', impute its missing potential outcome as

$$\widehat{Y}_t(0) = Y_{t'}$$

- Estimate the causal effect on the exposed using

$$\widehat{\tau} = \frac{1}{\sum_t I(E_t = 1)} \sum_t \left(Y_t - \widehat{Y}_t(0)\right) I(E_{tj} = 1)$$

The matching estimator

The estimator's bias is bounded, and can be made arbitrarily small

Theorem 1

Suppose
$Y_t(e) = \theta + \beta e + h_0(t) + \sum_{s=1}^{p_W} h_s(W_{ts}) + \sum_{s=1}^{p_X} h_{p_W+s}(\widetilde{X}_{ts}) + \sum_{s=1}^{p_P} h_{p_W+p_X+s}(\widetilde{P}_{ts}) + \epsilon_t(e)$,
with $\mathrm{E}(\epsilon_t(e)|\cdot) = 0$ *and functions* $h_0, h_1, \ldots, h_{p_W+p_X+p_P}$ *that are* K*-times differentiable with* $|h_s^{(k)}(x)| \leq c$*, then*

$$|\mathrm{E}(\hat{\tau} - \tau)| \leq C_T\delta + C_{WXP}\delta' + C_{TWXP}\ell^{K-1},$$

where C_T, C_{WXP} *and* C_{TWXP} *are constants and* δ, δ', l *are controlled by the researcher.*

The advantages of temporally-average estimands for each outcome unit

- We are more familiar with estimands in cross-sectional studies that average across units
- Measuring and adjusting for temporal covariate information might be easier and lower-dimensional than unit covariate information
- Number of units might be small (or even 1)
- Estimand informs us of effect heterogeneity across units, which can be policy-relevant

The impact of wildfire smoke exposure on bikeshare hours

- 1,003 days, with ~ 140 exposed ones
- Balance on daily temperature, humidity, precipitation, wind speed, and wind direction as potential time-varying confounders
- Based on our framework, we conjecture that no additional covariates are needed for confounding adjustment!
  - factors influencing wildfire occurrence and smoke dispersion are unlikely to impact biking activity in distant locations
  - economic indices fluctuating over time that might affect biking activity are likely unrelated to wildfire occurrence

The impact of wildfire smoke exposure on bikeshare hours

	San Francisco			East Bay			San Jose		
Naïve-t	0.973	(1.000)		0.110	(1.000)		0.022	(0.810)	
Maching 1-1	**-0.601**	**(0.014)**	101	-0.043	(0.190)	103	-0.002	(0.235)	100
Matching 1-1/2	**-0.521**	**(0.021)**	101	-0.031	(0.265)	103	-0.002	(0.196)	100
Matching 1-2	-0.100	(0.356)	56	-0.001	(0.490)	59	-0.001	(0.279)	59

The columns correspond to estimate, p-value, and number of matched exposed time periods.

Summary

We provide a comprehensive framework for bipartite interference with time series observational data and a random network

- Define estimands as temporal averages of time-specific effect for each outcome unit

  $\rightsquigarrow$ Important advantages over estimands that average across units

- Provide unconfoundedness assumptions on the treatment assignment of interventional units and the network process to establish the outcome unit's exposure unconfoundedness

  $\rightsquigarrow$ tangible, actionable confounding adjustment

- Focusing on binary exposures, we develop three matching estimators

  $\rightsquigarrow$ guarantee small estimation bias

References

Giulio Grossi, Patrizia Lattarulo, Marco Mariani, Alessandra Mattei, and O Oner. Synthetic control group methods in the presence of interference: The direct and spillover effects of light rail on neighborhood retail activity. *arXiv preprint arXiv:2004.05027*, 2020.

Jens Hainmueller. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20(1):25–46, 2012.

Christopher Harshaw, Fredrik Sävje, David Eisenstat, Vahab Mirrokni, and Jean Pouget-Abadie. Design and analysis of bipartite experiments under a linear exposure-response model. *Electronic Journal of Statistics*, 17(1):464–518, 2023.

Luke Keele, Rocío Titiunik, and José R Zubizarreta. Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout. *Journal of the Royal Statistical Society, Series A*, 178:223–239, 2014.

Fiammetta Menchetti and Iavor Bojinov. Estimating causal effects in the presence of partial interference using multivariate bayesian structural time series models. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (21-048), 2020.

Jean Pouget-Abadie, Kevin Aydin, Warren Schudy, Kay Brodersen, and Vahab Mirrokni. Variance reduction in bipartite experiments through correlation clustering. *Advances in Neural Information Processing Systems*, 32, 2019.

Nathan B Wikle and Corwin M Zigler. Causal health impacts of power plant emission controls under modeled and uncertain physical process interference. *arXiv preprint arXiv:2306.05665*, 2023.

Corwin M Zigler and Georgia Papadogeorgou. Bipartite causal inference with interference. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 2021.

José R Zubizarreta. Using mixed integer programming for matching in an observational study of kidney failure after surgery. *Journal of the American Statistical Association*, 107(500): 1360–1371, 2012.

The manuscript is available on arXiv (*arXiv:2404.04775*)

This work was supported by the National Science Foundation under Grant No. 2124124.

Questions?