Space-Time-Meta Causal Inference: A Weighting Perspective

José R. Zubizarreta

02/02/2026

Advances in Quantitative Medical Care
Institute for Mathematical and Statistical Innovation

Outline

Transparency in health care research

- Rosenbaum:
  - “Transparency means making evidence evident.”

Transparency in health care research

- Rosenbaum:
  - “Transparency means making evidence evident.”
  - “With enough model, you don’t need data.”

Transparency in health care research

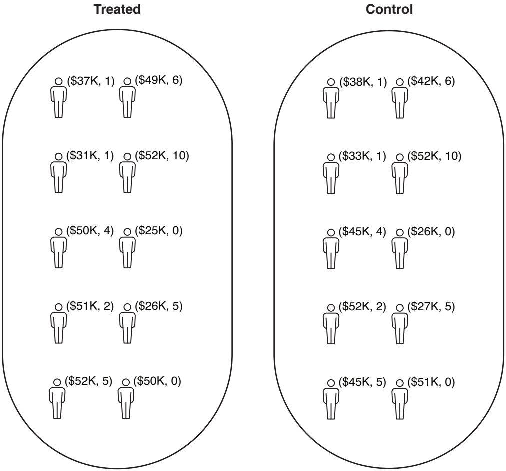
- Rosenbaum:
  - "Transparency means making evidence evident."
  - "With enough model, you don't need data."
- To me, this means being crystal clear how are we acting on the data to build a contrast.

Transparency in health care research

- Rosenbaum:
  - “Transparency means making evidence evident.”
  - “With enough model, you don’t need data.”
- To me, this means being crystal clear how are we acting on the data to build a contrast.
- And here, the benchmark of a randomized experiment is critical.

Randomized experiments

Chattopadhyay and Z. (2024)



Diagnostic Dashboard

Covariate Balance

	Income	Visits
Treated	$38,934	4.2
Control	$39,325	4.3
Difference:[1] $_{\text{Treated-Control}}$	0.04	0.03

Study Representativeness

	Income	Visits
Target	$38,934	4.2
Difference:[2] $_{\text{Treated-Target}}$	0.00	0.00
Difference:[2] $_{\text{Control-Target}}$	0.04	0.03

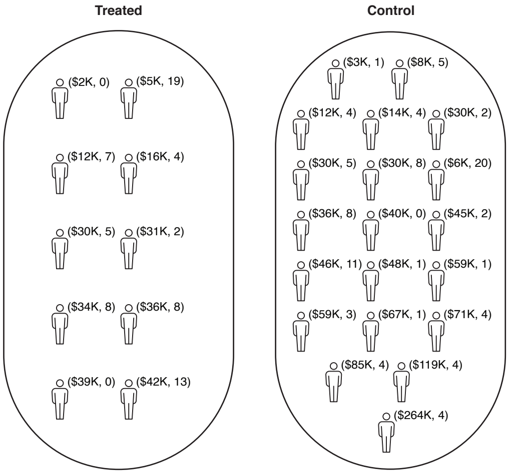
Sample Size

	Effective	Nominal
Treated	100	100
Control	100	100

Differential Weighting[3]

	Min	Q1	Med	Q3	Max
Treated	1	1	1	1	1
Control	1	1	1	1	1

Observational studies



Diagnostic Dashboard

Covariate Balance

	Income	Visits
Treated	$27,216	4.6
Control	$44,388	3.8
Difference:[1] $_{\text{Treated-Control}}$	0.83	0.16

Study Representativeness

	Income	Visits
Target	$27,216	4.6
Difference:[2] $_{\text{Treated-Target}}$	0.00	0.00
Difference:[2] $_{\text{Control-Target}}$	1.56	0.17

Sample Size

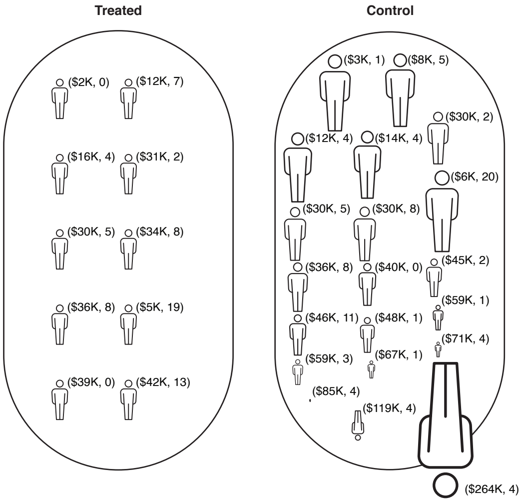
	Effective	Nominal
Treated	100	100
Control	200	200

Differential Weighting[3]

	Min	Q1	Med	Q3	Max
Treated	1	1	1	1	1
Control	1	1	1	1	1

Regression

Chattopadhyay and Z. (2023)



Diagnostic Dashboard

Covariate Balance

	Income	Visits
Treated	$28,533	4.5
Control	$28,533	4.5
Difference:[1] Treated-Control	0.00	0.00

Study Representativeness

	Income	Visits
Target	$27,216	4.6
Difference:[2] Treated-Target	0.12	0.03
Difference:[2] Control-Target	0.12	0.03

Sample Size

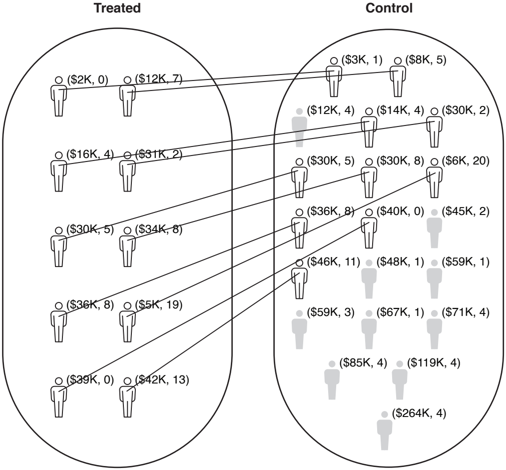
	Effective	Nominal
Treated	99	100
Control	164	200

Differential Weighting[3]

	Min	Q1	Med	Q3	Max
Treated	0.7	0.9	1.0	1.1	1.2
Control	-3.8	0.8	1.1	1.4	1.9

Matching

Z. (2013)



Diagnostic Dashboard

Covariate Balance

	Income	Visits
Treated	$27,216	4.6
Control	$27,829	4.5
Difference:[1] Treated-Control	0.03	0.02

Study Representativeness

	Income	Visits
Target	$27,216	4.6
Difference:[2] Treated-Target	0.00	0.00
Difference:[2] Control-Target	0.06	0.03

Sample Size

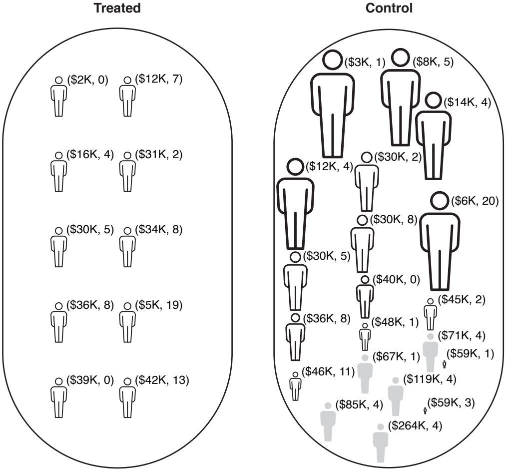
	Effective	Nominal
Treated	100	100
Control	100	200

Differential Weighting[3]

	Min	Q1	Med	Q3	Max
Treated	1.0	1.0	1.0	1.0	1.0
Control	0.0	0.0	1.0	1.0	1.0

Weighting

Z. (2015)



Diagnostic Dashboard

Covariate Balance

	Income	Visits
Treated	$27,216	4.6
Control	$27,243	4.6
Difference:[1] Treated-Control	0.001	0.00

Study Representativeness

	Income	Visits
Target	$27,216	4.6
Difference:[2] Treated-Target	0.00	0.00
Difference:[2] Control-Target	0.002	0.00

Sample Size

	Effective	Nominal
Treated	100	100
Control	125	200

Differential Weighting[3]

	Min	Q1	Med	Q3	Max
Treated	1.0	1.0	1.0	1.0	1.0
Control	0.0	0.4	1.0	1.6	2.7

Translation through optimization

- These three methods offer a similar weighting representation.
  - Explains how each method acts on each observation in the data.
  - Provides a more precise connection to the hypothetical experiment.

minimize *dispersion of the weights*

subject to

covariate balance adjustments

extrapolation/interpolation requirements

Just you weight

$$\hat{w} = \arg\min_{w} \{\mathcal{D}(w) : \; w \in \mathcal{B}_\delta \cap \mathcal{A}_+\}$$

Connection from the optics of optimization

Cohn and Z. (2022), Wang and Z. (2020), Chattopadhyay and Z. (2023)

Matching (PM1):

$$\underset{\boldsymbol{m}}{\text{maximize}} \sum_{i:Z_i=0} m_i$$

subject to

$$\left| \sum_{i:Z_i=0} m_i B_k(X_i) - B_k(\boldsymbol{X}^*) \right| \leq \delta_k, \quad k = 1, 2, \ldots, K$$

$$m_i \in \{0, 1\}, i : Z_i = 0$$

Weighting (SBW):

$$\underset{\boldsymbol{w}}{\text{minimize}} \sum_{i:Z_i=0} (w_i - \bar{w}_c)^2$$

subject to

$$\left| \sum_{i:Z_i=0} w_i B_k(X_i) - B_k(\boldsymbol{X}^*) \right| \leq \delta_k, \quad k = 1, 2, \ldots, K$$

$$\sum_{i:Z_i=0} w_i = 1$$

$$w_i \geq 0, i : Z_i = 0$$

Regression (W-MRI...):

$$\underset{\boldsymbol{w}}{\text{minimize}} \sum_{i:Z_i=0} (w_i - \tilde{w}_i^{\text{base}})^2 / w_i^{\text{scale}}$$

subject to

$$\left| \sum_{i:Z_i=0} w_i B_k(X_i) - B_k(\boldsymbol{X}^*) \right| \leq \delta_k, \quad k = 1, 2, \ldots, K$$

$$\sum_{i:Z_i=0} w_i = 1$$

Reminder

- These are generic methods for variable adjustment.
  - They can be used under various identification strategies.

Causal inference with spatio/temporal/meta data...

- **... a weighting perspective:**

Causal inference with spatio/temporal/meta data...

- **... a weighting perspective:**
- Standard approach: regression modeling.

Causal inference with spatio/temporal/meta data...

- **... a weighting perspective:**
- Standard approach: regression modeling.
- Common struggles: opaque aggregation, model dependence, complicated interpretation...

Causal inference with spatio/temporal/meta data...

- **... a weighting perspective:**
- Standard approach: regression modeling.
- Common struggles: opaque aggregation, model dependence, complicated interpretation...
- Our approach: we recast these methods through a weighting framework that subsumes and extends them.

Causal inference with spatio/temporal/meta data...

- **... a weighting perspective:**
- Standard approach: regression modeling.
- Common struggles: opaque aggregation, model dependence, complicated interpretation...
- Our approach: we recast these methods through a weighting framework that subsumes and extends them.
- Viewing these diverse problems through the weighting lens:
  - Clarifies aggregation,
  - Simplifies estimation,
  - Enables diagnostics.

Outline

Outline

Joint work

- Sophie Woodward, Francesca Dominici.

Spatial confounding and regression modeling

- Spatial regression models aim to adjust for unmeasured confounding by modeling spatially structured errors.
- Three widely used models:
  - Random effects (RE) models;
  - Conditional autoregressive (CAR) models;
  - Gaussian process (GP) models.
- Debate: how well do these methods adjust for **unmeasured spatial confounding**?

  (We provide a finite sample characterization.)

Basic spatial models

- General model: $Y = X\beta + \tau Z + \epsilon$.
- Spatial structure is in the error term ϵ:
  - **RE**: clusters (e.g., states).
  - **CAR**: neighbors.
  - **GP**: geographic distances.
- All try to capture hidden area-level confounding.

Three models

- **RE** model:

$$Y_i = \beta^\top X_i + \tau Z_i + \gamma_{C_i} + \xi_i$$

where $\gamma_{C_i} \sim N(0, \rho^2)$ is a random intercept for cluster C_i, and $\xi_i \sim N(0, \sigma^2)$.

- **CAR** model:

$$Y_i = \beta^\top X_i + \tau Z_i + \phi_i + \xi_i$$

where $\phi \sim N(0,\ \rho^2(D - A)^{-1})$, A is the adjacency matrix, $D_{ii} = \sum_j A_{ij}$, and $\xi_i \sim N(0, \sigma^2)$.

- **GP** model:

$$Y_i = \beta^\top X_i + \tau Z_i + \nu_i + \xi_i$$

where $\nu \sim \mathrm{GP}(0,\ \rho^2 K)$, K_{ij} is a kernel/covariance based on spatial distance, and $\xi_i \sim N(0, \sigma^2)$.

Weighting representation

Proposition

The GLS estimator of τ *can be expressed as*

$$\hat{\tau}_{GLS} = \sum_{i:Z_i=1} w_i Y_i - \sum_{i:Z_i=0} w_i Y_i$$

with weights of $\mathbf{w} = (w_1, \ldots, w_n)$ *that admit the following closed-form representation*

$$\mathbf{w} = \mathbf{M} \frac{(\mathbf{I}_n - \mathbf{\Sigma}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T)\mathbf{\Sigma}^{-1}\mathbf{Z}}{\mathbf{Z}^T\mathbf{\Sigma}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{\Sigma}^{-1})\mathbf{Z}}, \quad (1)$$

where $\mathbf{M}$ *is the diagonal matrix with* (i, i) *entry* $M_{ii} = 2Z_i - 1$.

Implicit adjustment

Proposition

Suppose that $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_n + \rho^2 \mathbf{S}$, *where* $\mathbf{S}$ *is an* $n \times n$ *positive semidefinite matrix. Let* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *be the eigenvectors of* $\mathbf{S}$ *with corresponding eigenvalues* $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$*. Consider the following quadratic programming problem:*

$$\min_{\mathbf{w} \in \mathbb{R}^n} \left\{ \sigma^2 \sum_{i=1}^{n} w_i^2 + \rho^2 \sum_{k=1}^{n} \lambda_k \Big(\sum_{i:Z_i=1} w_i v_{ki} - \sum_{i:Z_i=0} w_i v_{ki} \Big)^2 \right\} \tag{2}$$

$$\textit{subject to} \sum_{i:Z_i=1} w_i = \sum_{i:Z_i=0} w_i = 1 \textit{ and } \sum_{i:Z_i=1} w_i \mathbf{X}_i = \sum_{i:Z_i=0} w_i \mathbf{X}_i. \tag{3}$$

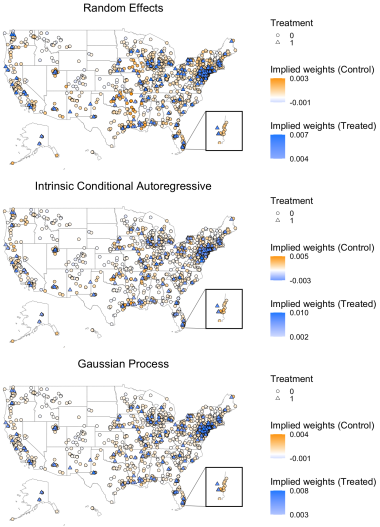
The solution to this problem are the implied weights of $\hat{\tau}_{GLS}$*,*

$$\mathbf{w} = \mathbf{M} \frac{(\mathbf{I}_n - \boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T)\boldsymbol{\Sigma}^{-1}\mathbf{Z}}{\mathbf{Z}^T\boldsymbol{\Sigma}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1})\mathbf{Z}}.$$

Key insights

- The spatially autocorrelated error term produces approximate balance on a hidden set of covariates, thereby adjusting for a specific class of unmeasured confounders.
- The error covariance structure can be equivalently expressed as regressors under a linear model.
- So: **from error structure to covariate structure** to spatial balance.

How spatial regression builds causal contrasts



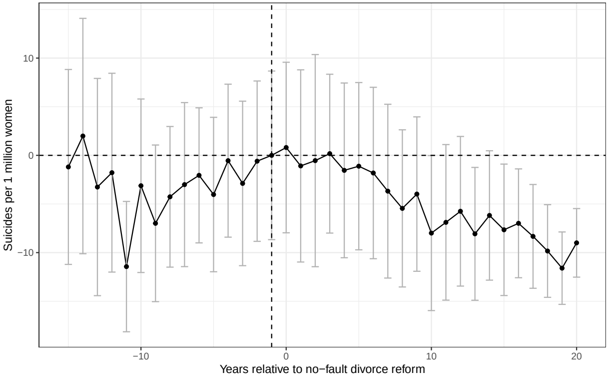
Outline

Joint work

- Ambarish Chattopadhyay, Yuzhou Lin, Zhu Shen.



The events study plot



- Running example: impact of divorce reforms on female suicide.
- **TWFE model**: $Y_{it} = \alpha_i + \beta_t + \boldsymbol{X}_i^\top \gamma + \sum_{\substack{t-F_i=a \\ t-F_i \neq -\infty}}^{b} \tau_l \mathbb{1}(t - F_i = l) + \epsilon_{it}, \forall i, t.$

Setup

- Panel dataset with:
  - N units followed for T periods.
    - Units, time points, **observations**.
  - Staggered interventions.
  - Baseline covariates X_i, times of treatment initiation F_i, treatment assignment indicators Z_{it}, observed outcomes Y_{it}^{obs}.

		Year					
		2000	2001	2002	2003	2004	2005
Unit	1	○	●	●	●	●	●
	2	○	○	●	●	●	●
	3	○	○	○	●	●	●
	4	○	○	○	○	●	●
	5	○	○	○	○	○	●
	6	○	○	○	○	○	○

Implied weights decomposition

- TWFE model: $Y_{it} = \alpha_i + \beta_t + \boldsymbol{X}_i^\top \gamma + \sum_{\substack{t-F_i=a \\ t-F_i \neq -\infty}}^{b} \tau_l \mathbb{1}(t - F_i = l) + \epsilon_{it}, \forall i, t.$
- **Equivalent** representation

$$\hat{\tau}_l^{\text{WLS}} = \underbrace{\sum_{i=1}^{n} \sum_{t=1}^{T} w_{it} Y_{it}^{\text{obs}} \mathbb{1}(t - F_i = l)}_{\text{treatment component}} - \underbrace{\sum_{i=1}^{n} \sum_{t=1}^{T} w_{it} Y_{it}^{\text{obs}} \mathbb{1}(t - F_i \neq l)}_{\text{control component}}$$

  for suitable weights w_{it}.
- We provide a similar formulation for Sun and Abraham (2021).

The TWFE experiment

- Target estimand: $\text{ATE}^{\mathcal{P}}_{2003}(2002, \infty)$.

		Year					
		2000	2001	2002	2003	2004	2005
Unit	1	○C	●C	●T	●C	●C	●C
	2	○C	○C	●C	●T	●C	●C
	3	○C	○C	○C	●C	●T	●C
	4	○C	○C	○C	○C	●C	●T
	5	○C	○C	○C	○C	○C	●C
	6	○C	○C	○C	○C	○C	○C

$$\hat{\tau}_l^{\text{WLS}} = \underbrace{\sum_{i=1}^{n}\sum_{t=1}^{T} w_{it} Y_{it}^{\text{obs}} \mathbb{1}(t - F_i = l)}_{\text{treatment component}} - \underbrace{\sum_{i=1}^{n}\sum_{t=1}^{T} w_{it} Y_{it}^{\text{obs}} \mathbb{1}(t - F_i \neq l)}_{\text{control component}}$$

Building robust contrasts: a weighting approach

- For any $r, s \in \mathcal{T}^+$, write $\mathcal{C}_{r,s} = \{(i,t) : F_i = r, t = s\}$ for the set of observations where treatment is initiated at time r and the outcome is measured at time s.

$$\mathbb{E}(\hat{\mu}) - \mu = \mathbb{E}\left\{ \sum_{(i,t)\in\mathcal{C}_{r,s}} w_{it} m_{r,s}(\boldsymbol{X}_i) - \frac{1}{n}\sum_{i=1}^{n} m_{r,s}(\boldsymbol{X}_i) \right\}.$$

- $\mathcal{C}_{r,s}$ is a credible set of observations to build the contrast, valid under **increasingly stronger** assumptions:

Building robust contrasts: a weighting approach

- For any $r, s \in \mathcal{T}^+$, write $\mathcal{C}_{r,s} = \{(i,t) : F_i = r, t = s\}$ for the set of observations where treatment is initiated at time r and the outcome is measured at time s.

$$\mathbb{E}(\hat{\mu}) - \mu = \mathbb{E}\left\{ \sum_{(i,t)\in\mathcal{C}_{r,s}} w_{it} m_{r,s}(\boldsymbol{X}_i) - \frac{1}{n}\sum_{i=1}^{n} m_{r,s}(\boldsymbol{X}_i) \right\}.$$

- $\mathcal{C}_{r,s}$ is a credible set of observations to build the contrast, valid under **increasingly stronger** assumptions:

- Minimum variance, population targeted, sample bounded estimator.

Building robust contrasts: a weighting approach

- For any $r, s \in \mathcal{T}^+$, write $\mathcal{C}_{r,s} = \{(i,t) : F_i = r, t = s\}$ for the set of observations where treatment is initiated at time r and the outcome is measured at time s.

$$\mathbb{E}(\hat{\mu}) - \mu = \mathbb{E}\left\{ \sum_{(i,t)\in\mathcal{C}_{r,s}} w_{it} m_{r,s}(\boldsymbol{X}_i) - \frac{1}{n}\sum_{i=1}^{n} m_{r,s}(\boldsymbol{X}_i) \right\}.$$

- $\mathcal{C}_{r,s}$ is a credible set of observations to build the contrast, valid under **increasingly stronger** assumptions:

- Minimum variance, population targeted, sample bounded estimator.
- The TWFE estimator is a particular case.

Statistical properties

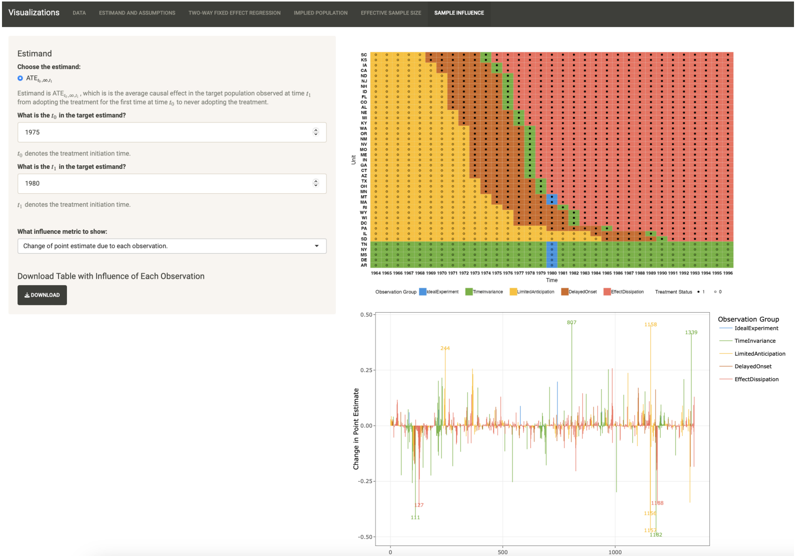
Theorem (consistency)

Under the previous assumptions, $\widehat{ATE}^{\mathbb{P}}_{t_0,\infty,t_1}$ is consistent for $ATE^{\mathbb{P}}_{t_0,\infty,t_1}$ if any of the following conditions are satisfied: (a) The treatment initiation models are correctly specified for both time t_0 and ∞, (b) the potential outcome models are correctly specified for both time t_0 and ∞, (c) the treatment initiation model for time t_0 and the potential outcome model for time ∞ are both correctly specified, (d) the treatment initiation model for time ∞ and the potential outcome model for time t_0 are both correctly specified.

Theorem (asymptotic normality and semiparametric efficiency)

Under the previous assumptions, as $n \to \infty$, $n^{1/2}(\widehat{ATE}^{\mathbb{P}}_{t_0,\infty,t_1} - ATE^{\mathbb{P}}_{t_0,\infty,t_1})$ converges in distribution to a Normal random variable with mean 0 and variance V_τ, where V_τ equates to the semiparametric efficiency bound for $ATE^{\mathbb{P}}_{t_0,\infty,t_1}$.

Interactive software



Outline

Wenqi Shi

Towards personalized evidence integration and synthesis

Setup and notation

- **There are** n units nested in m studies, with sizes $n_i \ni \sum_{i=1}^{m} n_i = n$.
- Index the j-th patient in study i by ij.
- Write:
  - Binary treatment assignment as $Z_{ij} \in \{0, 1\}$.
  - Baseline covariates as $x_{ij} = (x_{ij,1}, \ldots, x_{ij,p})' \in \mathbb{R}^p$.
  - Potential outcomes by $\{Y_{ij}(0), Y_{ij}(1)\}$.
  - Observed outcome as $Y_{ij} = Z_{ij} Y_{ij}(1) + (1 - Z_{ij}) Y_{ij}(0)$.

Example: one-stage ID meta-analysis

- One-stage individual-patient data (ID) meta-analysis model:

$$y_{ij} = \phi_i + \tau_i z_{ij} + \beta_i^\top x_{ij} + \gamma_i^\top x_{ij} z_{ij} + \epsilon_{ij}.$$

Example: one-stage ID meta-analysis

- One-stage individual-patient data (ID) meta-analysis model:

$$y_{ij} = \phi_i + \tau_i z_{ij} + \beta_i^\top x_{ij} + \gamma_i^\top x_{ij} z_{ij} + \epsilon_{ij}.$$

- Common instances:
  - Fixed effects: different constants are used for different studies.
  - Common effect: e.g., $\tau_i = \tau$ for all i.
  - Random effects: e.g., $\tau_i \sim \mathcal{N}(\tau, \sigma_\tau^2)$.

Example: one-stage ID meta-analysis

- One-stage individual-patient data (ID) meta-analysis model:

$$y_{ij} = \phi_i + \tau_i z_{ij} + \beta_i^\top x_{ij} + \gamma_i^\top x_{ij} z_{ij} + \epsilon_{ij}.$$

- Common instances:
  - Fixed effects: different constants are used for different studies.
  - Common effect: e.g., $\tau_i = \tau$ for all i.
  - Random effects: e.g., $\tau_i \sim \mathcal{N}(\tau, \sigma_\tau^2)$.
- Key questions:
  - What defines the target population and the corresponding estimand?
  - How is each fundamental **"atom"** of information used in the analysis?

Estimand and populations

- We have a sample of n patients from the study population $\mathcal{P}$.
- $\mathcal{T}$ represents the target population.
- $\mathbb{P}$ and $\mathbb{T}$ are probability measures for the study and target populations with densities $p(\cdot)$ and $t(\cdot)$.
- The treated and control groups in $\mathcal{P}$ are described by $\mathbb{P}_1$ and $\mathbb{P}_0$, with densities $p_1(\cdot)$ and $p_0(\cdot)$.
- We wish to learn the (target) average treatment effect for $\mathcal{T}$:

$$\tau := \mathbb{E}_{\mathbb{T}}[Y(1) - Y(0)].$$

A *w*-formula

- The conditions in A1 jointly allow us to **identify** the target average treatment effect (τ) using the formula:

$$\tau = \mathbb{E}_{\mathbb{P}_1}\left[\frac{t(X)}{p(X)e(X)}Y\right] - \mathbb{E}_{\mathbb{P}_0}\left[\frac{t(X)}{p(X)(1-e(X))}Y\right].$$

Properties of the weighting factor

- The **weighting factor** adjusts for biases from observed covariate imbalances between the populations and treatment groups.
- This factor is always non-negative and, when $f(x) = 1$, the weights sum to one in expectation.
- In general:

$$\mathbb{E}_{\mathbb{P}_1}\left[\frac{t(X)}{p(X)e(X)}f(X)\right] = \mathbb{E}_{\mathbb{P}_0}\left[\frac{t(X)}{p(X)\{1-e(X)\}}f(X)\right] = \mathbb{E}_{\mathbb{T}}\left[f(X)\right]$$

for any function f.

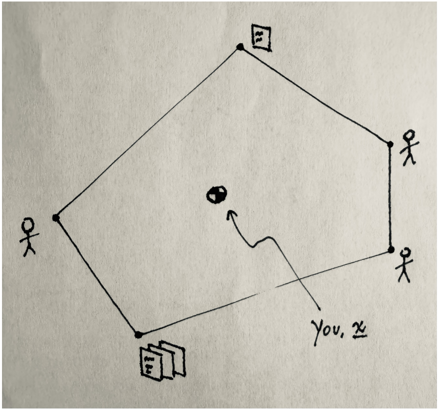
A linear estimator

- This motivates using a linear estimator

$$\hat{\tau}^{\circ} := \sum_{i=1}^{m} \sum_{j:Z_{ij}=1} \hat{w}_{1,ij} Y_{ij} - \sum_{i=1}^{m} \sum_{j:Z_{ij}=0} \hat{w}_{0,ij} Y_{ij},$$

with weights $\hat{w}_z$ that approximate the previous conditions.

Patient profile ⊕



Minimal weights

$$\arg\min_{w_z} \Big\{ \sum_{i=1}^{m} \sum_{j:Z_{ij}=z} \psi(w_{ij}) : \Big| \sum_{i=1}^{m} \sum_{j:Z_{ij}=z} w_{ij} B_k(x_{ij}) - B_k^{\circledast} \Big| \leq \delta_k, \text{for } k \in \mathcal{A},$$

$$\Big| \sum_{j:Z_{ij}=z} w_{ij} B_k(x_{ij}) - \big(\sum_{j:Z_{ij}=z} w_{ij} \big) B_k^{\circledast} \Big| \leq \delta_k, \text{for } k \in \mathcal{W} \text{ and } i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} \sum_{j:Z_{ij}=z} w_{ij} = 1, w_{ij} \geq 0 \Big\}.$$

- ψ is a dispersion measure; e.g., the variance of the weights.
- $\mathcal{A}$ and $\mathcal{W}$ index across and within studies balancing conditions.
- $B_k(x_{ij})$ are basis functions; $B_k^{\circledast}$ are their corresponding target values.
- Build **sample bounded** and unbounded estimators: $\hat{\tau}^{\circledast+}$ and $\hat{\tau}^{\circledast-}$.

Detecting relevant studies/units

Theorem

Under assumptions 1 and 3 and conditions S.1, the following probability limits hold as the sample size $n \to \infty$:

$$\Pr_{\mathbb{P}}\{n\hat{w}(X) > 0 \mid X \in V\} \to 1, \quad \Pr_{\mathbb{P}}\{n\hat{w}(X) = 0 \mid X \in \mathrm{Supp}(\mathbb{P}) \setminus V\} \to 1.$$

- A means to **detect** deviant studies and trim the sample using the original data (e.g., without estimating the propensity score).

Takeaways

- Personalized meta-analyses:
  - Target the covariate profiles of specific patients.
  - Provide estimates that are sample-bounded.
- General weighting framework:
  - Handle both individual-patient and aggregate-level data (ID/AD).
  - Encompass common regression approaches to meta-analyses.

Takeaways

- Personalized meta-analyses:
  - Target the covariate profiles of specific patients.
  - Provide estimates that are sample-bounded.
- General weighting framework:
  - Handle both individual-patient and aggregate-level data (ID/AD).
  - Encompass common regression approaches to meta-analyses.
- Properties, diagnostics:
  - Establish consistency and asymptotic normality.
  - **Detect** relevant study/unit donors for the target profile.

Takeaways

- Personalized meta-analyses:
  - Target the covariate profiles of specific patients.
  - Provide estimates that are sample-bounded.
- General weighting framework:
  - Handle both individual-patient and aggregate-level data (ID/AD).
  - Encompass common regression approaches to meta-analyses.
- Properties, diagnostics:
  - Establish consistency and asymptotic normality.
  - **Detect** relevant study/unit donors for the target profile.

    (Result of independent interest for regression analysis.)

Outline

Concluding remarks

- A unified, weighting-based framework that encompasses and extends regression methods.

Concluding remarks

- A unified, weighting-based framework that encompasses and extends regression methods.
- For spatial data, modeling errors equals imposing specific covariate structures via regularization.
- For panel data, we clarify how two-way fixed effects aggregate information for causal contrasts.
- For meta data, we propose a diagnostic to detect measurements whose covariates deviate from the target, guiding potential exclusion.

Concluding remarks

- A unified, weighting-based framework that encompasses and extends regression methods.
- For spatial data, modeling errors equals imposing specific covariate structures via regularization.
- For panel data, we clarify how two-way fixed effects aggregate information for causal contrasts.
- For meta data, we propose a diagnostic to detect measurements whose covariates deviate from the target, guiding potential exclusion.
- This **weighting perspective** streamlines estimation and offers diagnostics for causal inference in complex data.

Acknowledgments

- Work supported by PCORI.

References

- Shen, Z., Chattopadhyay, A., Lin, Y., and Zubizarreta, J. R., "An Anatomy of Event Studies: Hypothetical Experiments, Exact Decomposition, and Weighting Diagnostics," *arXiv:2410.17399*.
- Shi, W., and Zubizarreta, J. R., "On the Use of Weighting for Personalized and Transparent Evidence Synthesis," *arXiv:2509.00228*.
- Woodward, S. M., Dominici, F., and Zubizarreta, J. R., "Understanding Spatial Regression Models from a Weighting Perspective in an Observational Study of Superfund Remediation," *arXiv:2509.00228*.

References

- Chattopadhyay, A., and Zubizarreta, J. R. (2023), "On the Implied Weights of Linear Regression for Causal Inference," *Biometrika*, 110, 615–629.
- Chattopadhyay, A., and Zubizarreta, J. R. (2024), "Causation, Comparison, and Regression," *Harvard Data Science Review*, 6.1.
- Wang, Y., and Zubizarreta, J. R. (2020), "Minimal Dispersion Approximate Balancing Weights: Asymptotic Properties and Practical Considerations," *Biometrika*, 107, 93–105.

Space-Time-Meta Causal Inference: A Weighting Perspective

José R. Zubizarreta

02/02/2026

Advances in Quantitative Medical Care
Institute for Mathematical and Statistical Innovation

Regularity assumptions

Assumption (treatment initiation)

The conditional probability for treatment initiation at time t satisfies $\Pr(F_i = t|\boldsymbol{X}_i = \boldsymbol{x}) = [n\rho'\{g_t^*(\boldsymbol{x})\}]^{-1}$ *for all* $\boldsymbol{x} \in Supp(\boldsymbol{X}_i)$*, where* $\rho(\cdot)$ *is a smooth function obtained from the dual of the weighting problem, and* $g_t^*(\cdot)$ *is a smooth function that satisfies* $\sup_{\boldsymbol{x} \in Supp(\boldsymbol{X}_i)} |g_t^*(\boldsymbol{x}) - \boldsymbol{B}(\boldsymbol{x})^\top \boldsymbol{\lambda}_{1t}^*| = O(K^{-r_t})$ *for some* $\boldsymbol{\lambda}_{1t}^* \in \mathbb{R}^K$ *and* $r_t > 1$*.*

Assumption (potential outcomes)

The conditional mean function of the potential outcome at t_1 *under treatment initiation at* t *satisfies* $\sup_{\boldsymbol{x} \in Supp(\boldsymbol{X}_i)} |m_{t,t_1}(\boldsymbol{x}) - \boldsymbol{B}(\boldsymbol{x})^\top \boldsymbol{\lambda}_{2t}^*| = O(K^{-s_t})$ *for some* $\boldsymbol{\lambda}_{2t}^* \in \mathbb{R}^K$ *and* $s_t > 3/4$*, with* $||\boldsymbol{\lambda}_{2t}^*||_2||\boldsymbol{\delta}||_2 = o(1)$*.*