Controlling for Unmeasured Confounding in Panel Data using Minimal Bridge Functions: From Two-Way Fixed Effects to Factor Models

Imbens, Kallus, Mao (2021)

Presented by Apoorva Lal

October 20, 2021

Introduction

- Panel data models offer a variety of ways to control for unmeasured confounding
- Synthetic control, factor models, and matrix-completion require $N_0, T_0 \to \infty$
- This paper: causal inference from panel data under a linear factor model with *fixed-T* asymptotics
  - use **bridge functions** that transform pre-treatment variables to control confounding
  - generically non-unique, so need regularised GMM

Setup

- N units, time $t \in \{-T_0, \ldots, -1, 0, 1, \ldots, T_1\}$
- Treatment : $A_i = 1$ (!), Potential outcomes: $Y_{i,t}(0)$ and $Y_{i,t}(1)$
  - $\mathcal{C} := \{i : A_i = 0\}$, $\mathcal{T} := \{i : A_i = 0\}$, sizes N_0, N_1 respectively
  - pre $:= \{t : t < 0\}$, post $:= \{t : t > 0\}$
- Covariates : $X_i \in \mathbb{R}^d$
- $(A_i, X_i, Y_{i,t}(0), Y_{i,t}(1), -T_0 \leq t \leq T_1)$ iid drows from common population

Estimand: ATT at $t = 0$

$$\text{ACT} = \mathbb{E}\left[Y_{\cdot,0}(1) | A = 1\right] - \underbrace{\mathbb{E}\left[Y_{\cdot,0}(0) | A = 1\right]}_{=: \gamma^*}$$

Assumptions

1. **A1 Linear factor model** for $Y_{i,t}(0) = \mathbf{V}_t^\top \mathbf{U}_i + \mathbf{b}_t^\top \mathbf{X}_i + \varepsilon_{i,t}$; $U_i, V_t \in \mathbb{R}^r$
   - $Y_0(0) \perp\!\!\!\perp A | X, U$
2. **A2 Positivity**: $\mathbf{Pr}\left(A = a | X, U\right) > 0$ almost surely.

TWFE

Special-case of factor model with $Y_{i,t}(0) = U_i + b_t + \varepsilon_{i,t}$. Implies the following estimator for the counterfactual mean:

$$\widehat{\gamma}_{\text{DID}} = \underbrace{\frac{1}{T_0 N_1} \sum_{t \in \text{pre}} \sum_{i \in \mathcal{T}} Y_{i,t}}_{\mathbb{E}\left[\frac{1}{T_0}\mathbf{1}_{T_0}^\top Y_{\text{pre}} | A=1\right]} - \underbrace{\frac{1}{T_0 N_0} \sum_{t \in \text{pre}} \sum_{i \in \mathcal{C}} Y_{i,t}}_{\mathbb{E}\left[\frac{1}{T_0}\mathbf{1}_{T_0}^\top Y_{\text{pre}} | A=0\right]} + \underbrace{\frac{1}{N_0} \sum_{i \in \mathcal{C}} Y_{i,0}}_{\mathbb{E}[Y_0 | A=0]}$$

Bridge Functions I : DiD

- DiD estimand can be written as $\gamma^*_{\text{DID}} = \mathbb{E}\left[h(Y_{\text{pre}}; \theta^*_{\text{DID}})|A=1\right]$ where
  - $h(Y_{\text{pre}}; \theta) = \theta_1' Y_{\text{pre}} + \theta_2$
  - $\theta_{\text{DID},1} = \mathbf{1}_{T_0}/T_0$
  - $\theta_{\text{DID},2} = \mathbb{E}\left[-\frac{1}{T_0}\sum_{t\in\text{pre}} Y_t + Y_0\right] = b_0 - b_{\text{pre}}^\top \theta^*_1$
- This transformation is learned based on control units' outcomes $Y_{i,t} \forall\, i \in \mathcal{C}, t \in \text{pre} \cup \{0\}$
- One of many valid transformations

$$\gamma^* = \mathbb{E}\left[h(Y_{\text{pre}}; \theta^*|A=1)\right] \quad \forall \theta \in \Theta^*_{FE} := \left\{\theta^* : \mathbf{1}_{T_0}^\top \theta^*_1 = 1, \theta^*_2 = b_0 - b_{\text{pre}}^\top \theta^*_1\right\}$$

- which satisfy the condition

$$\mathbb{E}\left[Y_0(0) - h(Y_{\text{pre}}; \theta^*|U, A=0)\right] = 0 \;\forall\; \theta^* \in \Theta^*_{\text{FE}}$$

Definition: Bridge Functions

A function $h(Y_{\text{pre}}, X)$ is called a bridge function if

$$\mathbb{E}\left[Y_0(0) - h(Y_{\text{pre}}, X; \theta^*) | U, A = 0, X\right] = 0 \tag{1.1}$$

In words, Bridge functions give some transformations of the pre-treatment outcomes and covariates such that the **unmeasured confounding effects on this transformation exactly reproduce those on the counterfactual outcome**.
Conceptual link to proxy-controls / proxy-outcomes literature.
Under A1 and A2,

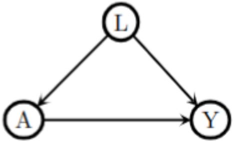
$$\gamma^* = \mathbb{E}\left[Y_0(0) | A = 1\right] = \mathbb{E}\left[h(Y_{\text{pre}}, X) | A = 1\right]$$

Equivalent moment equation:

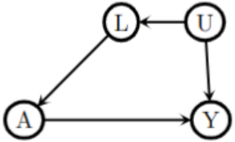
$$\mathbb{E}\left[A(h(Y_{\text{pre}}, X; \theta) - \gamma^*)\right] = 0 \tag{1.2}$$

Detour - Proxy Controls and Proxy Outcomes

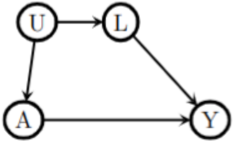
- Tchetgen-Tchetgen et al review paper
- 3 types of proxies
  - **1: confounds**: variables that are common causes of treatment and outcome variables
  - **2, 3: treatment (outcome) inducing confounding proxies**: potential cause of *outcome (treatment)* which is related with the *treatment (outcome)* through an unmeasured common cause for which the variable is a proxy
    - **proxy outcomes :** W not affected by $A - Y_{\text{pre}}$
    - **proxy treatments:** Z does not affect W or $Y - Y_{\text{post}}$



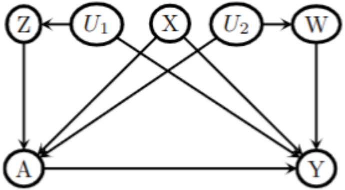
(a) Type (a) proxy.



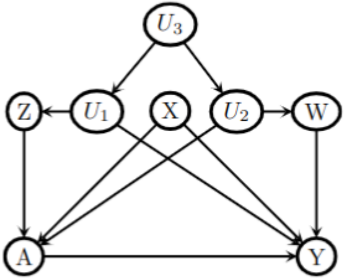
(b) Type (b) proxy.



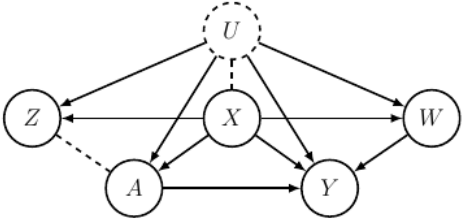
(c) Type (c) proxy.



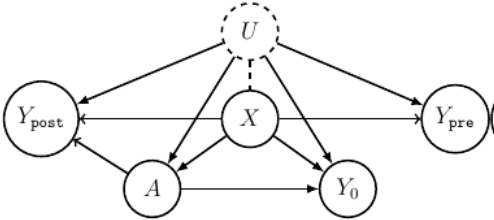
(d) Coexistence of type (a)(b)(c) proxies when NUC holds.



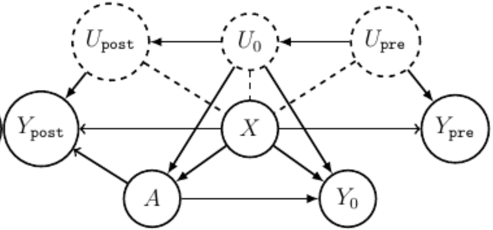
(e) Coexistence of type (a)(b)(c) proxies when NUC fails.



(a) Negative controls.



(b) Panels with time-invariant counfounders.



(c) Panels with time-varying confounders.

Bridge functions for Factor Models

Existence of bridge functions requires

- **A3 V-rank :** $\mathbf{V}_{\text{pre}} \in \mathbb{R}^{T_0 \times r}$ has full column rank (r)
- $\exists \theta_1^* \in \mathbb{R}^{T_0}$ to solve

$$\mathbf{V}_{\text{pre}}^{\top} \theta_1^* = V_0 \tag{1.3}$$

- $\implies U^{\top} \mathbf{V}_{\text{pre}}^{\top} \theta_1^* = U^{\top} V_0$

Bridge functions for LFMs

$$h(Y_{\text{pre}}, X, \theta^*) = {\theta_1^*}^{\top} Y_{\text{pre}} + \theta_2^* X$$
$$\forall\ \theta^* \in \Theta^* = \left\{ \theta^* : \mathbf{V}_{\text{pre}}^* \theta_1^* = V_0, \theta_2^* = b_0 - {\mathbf{B}_{\text{pre}}^*}^{\top} \theta_1^* \right\}$$

Bridge functions: Identification

- **A4: $\varepsilon-$Independence** Assume $\varepsilon_{\text{post}} \perp\!\!\!\perp (\varepsilon_{\text{pre}}, \varepsilon_0)$ so that $Y_{\text{post}} \perp\!\!\!\perp (Y_{\text{pre}}, Y_0) | X, U, A$; How plausible is this? How to falsify?
- **A5: More rank**: following two matrices have full rank $\widetilde{r} = r + d$

$$\mathbb{E}\left[\begin{bmatrix} U & X \end{bmatrix} \begin{bmatrix} U^\top \\ X^\top \end{bmatrix} \mid A = 0 \right] \in \mathbb{R}^{\tilde{r}\times\tilde{r}}, \quad \begin{bmatrix} \mathbf{V}_{\text{post}} & \mathbf{B}_{\text{post}} \\ 0_{d\times r} & I_{d\times d} \end{bmatrix} \in \mathbb{R}^{(T_1+d)\times\tilde{r}}$$

  - Second-order moment matrix invertible
  - post outcomes need to be informative about unmeasured confounders

$$\mathbb{E}\left[m(\mathcal{O}, \theta)\right] = \mathbb{E}\left[(1-A)\left(Y_0 - h\left(Y_{\text{pre}}, X; \theta^*\right)\right) \begin{bmatrix} Y_{\text{post}} \\ X \end{bmatrix}\right] = \mathbf{0}_{(T_1+d)\times 1} \quad (1.4)$$

Solutions to 1.4 corresponds to a valid bridge function($\theta^* \in \Theta^*$), which implies γ^* is identifiable.

Regularized GMM for Bridge Non-Uniqueness

- whenever $T_0 > r$, there are infinitely many bridge functions
- vanilla GMM is a bad idea
- Instead, target **Minimal Bridge Function**

$$\theta^*_{\text{min}} := \text{argmin}\left\{\|\theta\|_2 : \mathbb{E}\left[m(\mathcal{O}, \theta)\right] = 0\right\}$$

Given p.d. weight matrix **A6: Weight matrix, vanishing tuning param** $\mathcal{W}_{m,N} \rightarrow \mathcal{W}_{m,\infty}, \lambda_N \rightarrow 0$

$$\widehat{\theta} = \underset{\theta}{\text{argmin}}\left((\widehat{\mathbb{E}}_N[m(O,\theta)])^\top \mathcal{W}_{m,N} \widehat{\mathbb{E}}_N[m(O,\theta)] \right) + \lambda_n \|\theta\|_2^2$$

$$\left\| \widehat{\theta} - \theta^*_{min} \right\|_2 = \mathrm{O}\left(\lambda_n + \frac{1}{N\lambda_N} + \frac{1}{\sqrt{N}} \right)$$

Then solve $\widehat{\mathbb{E}}_N[g(O, \widehat{\theta}, \widehat{\gamma}] = 0$ to get counterfactual mean $\widehat{\gamma}$.

Counterfactual Mean is RAL

- Semiparametric theory
  - Kennedy (2015) review article
  - Chernozhukov et al chapter on regularized GMM
- counterfactual mean can be shown to be distributed

$$\sqrt{N}(\widehat{\gamma} - \gamma^*) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \psi(O_i, \theta^*_{\text{min}}, \gamma^*, \mathcal{W}_{m,\infty}) + \mathrm{O}\left(\underbrace{\lambda_N \sqrt{N}}_{\to 0} + \underbrace{\frac{1}{\sqrt{\lambda_N N}}}_{\to \infty} \right)$$

- Consistent plug-in estimator for asymptotic variance $\widehat{\sigma}^2$ is the variance of the influence function

Summary

- Proposes a method to estimate causal effects in linear factor model with time-varying confounding using bridge functions
  - that transform pre-treatment variables to control for confounding
- doesn't need $T \to \infty$, just needs $T_0 \geq r$ (where r is the 'number' of unmeasured confounders) how do we know r?
- regularized GMM estimation for non-unique bridge functions

Discussion points

- **intuition for assumptions?**
  - strong independence assumption
  - rank condition
- Ideas for other causal problems that can be posed as GMM problems?