Imputation of Counterfactual Outcomes when the Errors are Predictable

Benjamin O. Harrison

EMORY UNIVERSITY

23 February 2025

Econometrics/Reading Session

Synthetic Control: Recap

Time series dynamics matters.

- Artificial Counterfactual For High Dimensional Panel
- Framework
- Estimation

Part I

The Artificial Counterfactual Estimator

Synthetic Control: Recap

Recap

- The SCM construct the counterfactual outcomes as a weighted average of units from a donor pool using pre-intervention periods such that:

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \quad \sum_{j=2}^{J+1} w_j^* Y_{j2} = Y_{12},$$
$$\sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0}, \quad \text{and} \quad \sum_{j=2}^{J+1} w_j^* \mathbf{Z}_j = \mathbf{Z}_1$$

- We obtain the treatment effect for t (with $t > T_0$) as:

$$\widehat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

- The credibility of the SC depends on the extent to which it can fit the trajectory of Y_{1t} for an extended pre-intervention period
- There is no ex-ante guarantee that SC will be a good fit.
- The risk of overfitting in high dimensional settings. Ben-Michael et al., 2021

Common Challenges

- The SCM leverages on time series behavior of the treated and untreated unit prior the intervention.
- Implicit in this setup is the assumption that, the relationship between the treated unit and control unit remains stable after the treatment.
- There is no unobserved shocks or factors that distraught the time series behaviour post treatment.
- Example: unobserved economic shock or recession hits California after the proposition 99.
- The problem of Non-stationarity in variable of interest.
- The problem of dependencies of the errors across units and time.

Time series dynamics matters.

Classification of Counterfactual Imputation

A Optimization based approach: Classical Synthetic Control methods.

B Statistical Based Framework: Imposing model(parametric) on the counterfactual outcomes:

- Factor Models: Special example is the FarmTreat estimator by Fan et al., 2022. Capture unobserved shocks or factors.

$$Y_{i,t}(0) = \gamma_i' \boldsymbol{W}_{i,t} + \lambda_i' \boldsymbol{F}_t + e_{i,t}$$

- Autoregressive or Moving Average: Model time series dynamics. Ben-Michael et al., 2021 imposes an AR process with Regularization (Ridge, Lasso) to deal with over-fitting.
- State space model with Bayesian framework by Brodersen et al., 2015.
- ArCo Estimator: Deals with High Dimensional setting.

Artificial Counterfactual Estimator

- The goal here is to estimate the multivariate average treatment effect of an intervention on a treated unit.
- This extension is more applicable to macroeconomics models: examine the effect tax policy on {GDP, inflation and Unemployment.}
- This is a high dimensional setting as we have large dimensional set of variables from pool of donors.
- This estimator developed by Carvalho et al., 2018 is based on a simple two step procedure.
    1. Estimate the counterfactuals based on large dimensional set of variables from donor pool using LASSO operator.
    2. In stage 2, Compute the average intervention effect using the post intervention periods. Show that this estimator is asymptotically normal and consistent.

Framework

EMORY UNIVERSITY

- Suppose we have n units, indexed as $i = 1, \ldots, n$.
- For each unit at each time $t = 1 \cdots T$, we observe a realization of a set of variables $\boldsymbol{z}_{it} = \left(z_{it}^{1}, \ldots, z_{it}^{q_i}\right)' \in \mathbb{R}^{q_i}, q_i \geq 1$.
- Given the outcome of interest is $\boldsymbol{y}_t$, there exist a $\boldsymbol{h} : \mathbb{R}^{q_1} \mapsto \mathbb{R}^{q}$ is a measurable function such that $\boldsymbol{y}_t \equiv \boldsymbol{h}\left(\boldsymbol{z}_{1t}\right)$
- Unit $i = 1$ receives the intervention $\mathcal{D}_t$ at time T_0
- The observable variables of unit 1 as $\boldsymbol{y}_{1t} = \mathcal{D}_t \boldsymbol{y}_{1t}^{(1)} + (1 - \mathcal{D}_t)\, \boldsymbol{y}_{1t}^{(0)}$, where $\mathcal{D}_t = I\left(t \geq T_0\right)$

$$\boldsymbol{y}_t^{(1)} = \begin{cases} \boldsymbol{y}_t^{(0)}, & t = 1, \ldots, T_0 - 1 \\ \delta_t + \boldsymbol{y}_t^{(0)}, & t = T_0 \ldots, T \end{cases}$$

where $\{\delta_t\}_{t=T_0}^{T}$ is a deterministic sequence representing the treatment effect of unit 1.

Framework

EMORY UNIVERSITY

- The parameter of interest:

$$\boldsymbol{\Delta}_T = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^{T} \delta_t$$

- This is similar to the traditional average treatment effect on the treated (ATET).
- Construct let $\boldsymbol{y}_t^{(0)}$ using $\boldsymbol{Z}_{0t} = \left(\boldsymbol{z}'_{0t}, \ldots, \boldsymbol{z}'_{0t-p}\right)'$, where $\boldsymbol{z}_{0t} = (\boldsymbol{z}'_{2t}, \ldots, \boldsymbol{z}'_{nt})'$.
- The dimension f $\boldsymbol{Z}_{0t}$ depend upon the number of peers $(n-1)$, the number of variables per peer, $q_i, i = 2, \ldots, n$, and the choice of p.
- Consider an approximating model

$$\boldsymbol{y}_t^{(0)} = \mathcal{M}\left(\boldsymbol{Z}_{0t}, \boldsymbol{\theta}_0\right) + \boldsymbol{\nu}_t, t = 1, \ldots, T$$

where $\mathcal{M} : \mathcal{Z}_0 \times \Theta \to \mathcal{Y}$, $\mathbb{E}\left(\boldsymbol{\nu}_t\right) = 0$

ArCo Estimator

EMORY UNIVERSITY

Definition

The artificial counterfactual (ArCo) estimator is

$$\widehat{\boldsymbol{\Delta}}_T = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^{T} \widehat{\boldsymbol{\delta}}_t$$

where $\widehat{\boldsymbol{\delta}}_t \equiv \boldsymbol{y}_t - \mathcal{M}\left(\boldsymbol{Z}_{0t}, \widehat{\boldsymbol{\theta}}_{T_1}\right)$, for $t = T_0, \ldots, T$ and $\widehat{\boldsymbol{\theta}}_{T_1}$ is a consistent estimator for $\boldsymbol{\theta}_0$ using only the first T_1 observations of the data (pre-intervention).

- The core advantage of the ArCo estimator is that it disentangle the intervention effect from other macroeconomic or aggregate shocks that occur simultaneously.
- However this holds if the shock affects both the treated and control units, then it cancels out.

Identifying Assumptions

Assumption

The ArCo estimator is identified under the assumption that:

1. T_0 *is known i.e. No anticipation effects.*
2. $\boldsymbol{z}_{0t} \perp \mathcal{D}_s \quad \forall \quad t, s$ *i.e. stronger assumption version of No interference.*

- The first stage of the ArCo method requires a choice of $\mathcal{M}$.
- The choice of $\mathcal{M}$ is just an approximation of the true model $\boldsymbol{m}(\boldsymbol{Z}_{\boldsymbol{0t}}) \equiv \mathbb{E}(\boldsymbol{y}_{\boldsymbol{t}}{}^{(0)} \mid \boldsymbol{Z}_{0t})$
- Due to the dimension of $\boldsymbol{Z}_{0t}$, propose $\mathcal{M}(\boldsymbol{Z}_{0t}, \boldsymbol{\theta}_0) = \left(\boldsymbol{\theta}'_{0,1}\boldsymbol{x}_{1,t}, \ldots, \boldsymbol{\theta}'_{0,q}\boldsymbol{x}_{q,t}\right)'$, where $\boldsymbol{x}_t = \boldsymbol{h}_x(\boldsymbol{Z}_{0t})$ and both $\boldsymbol{x}_{j,t}$ and $\boldsymbol{\theta}_{0,j}$ are d_j- dimensional vectors for $j = 1, \ldots, q$.
- Set $\boldsymbol{r}_t \equiv \boldsymbol{m}(\boldsymbol{Z}_{0t}) - \mathcal{M}(\boldsymbol{Z}_{0t}, \boldsymbol{\theta}_0)$ as the approximation error and $\varepsilon_t \equiv \boldsymbol{y}_t^{(0)} - \boldsymbol{m}(\boldsymbol{Z}_{0t})$ as the projection error.

- We obtain the counterfactual model as:

$$y_{jt}^{(0)} = \boldsymbol{\theta}_{0,j}' \boldsymbol{x}_{j,t} + \nu_{jt}, \quad j = 1, \ldots, q$$

- The dimensions d_j of $\boldsymbol{x}_{j,t}$ can be very large, even larger than the sample size T, whenever the number of peers and/or the number of variables per peer is large.
- Therefore, we estimate $\boldsymbol{\theta}_0$ via

$$\widehat{\boldsymbol{\theta}} = \arg\min \left\{ \frac{1}{T_1} \sum_{t<T_0} \left(y_t - \boldsymbol{x}_t' \boldsymbol{\theta}\right)^2 + \varsigma \|\boldsymbol{\theta}\|_1 \right\}$$

where $\varsigma > 0$ is a penalty term and $\|\cdot\|_1$ denotes the ℓ_1 norm.

Concluding Remarks

EMORY UNIVERSITY

- The ArCo estimator under reasonable set of assumptions is asymptotically normal which makes inference in this setting less problematic.
- The use of LASSO, ArCo automatically selects the most informative donor units, reducing bias from weakly or noisy control units.
- If donor units are highly correlated with each other, LASSO will regularize the coefficients, avoiding overfitting.
- Hence the estimated θ_1 captures the cross-sectional dependence between treated unit and donor pool.
- Overlooks the predictability of errors that can distort conditional inference.

Part II

Imputation of Counterfactual Outcomes when the Errors are Predictable

Motivation

Motivation

Both ArCo and Traditional SCM, has two sources of uncertainty.

1. In-sample Uncertainty (Pre-Treatment):
   - Uncertainty that arises due finite sampling uncertainty.
   - Error obtained when estimating the model parameters using pretreatment data.
   - This uncertainty diminishes as pretreatment sample size T_0 increases.
2. Out-of-Sample Prediction Uncertainty:
   - Once the treatment occurs, we never observe the counterfactual outcome.
   - counterfactual is extrapolated using the fitted model from the pre-treatment data.
   - The error from extrapolation is not due to estimation but rather variation not explained by the model.
   - This uncertainty comes from misspecification. As such this error does not vanish asymptotically.

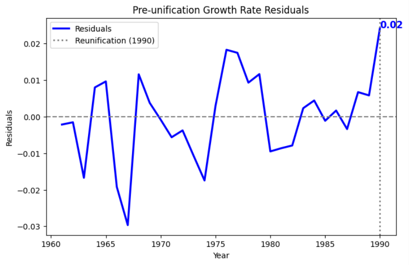
Motivation

- However, in many applications or models, the residuals are modeled as exogenous i.e. a random noise.
- If the model is truly exogenous, then the only source of prediction error is sampling uncertainty.
- However,in many cases residuals are predictable (serial correlation, cross-sectional dependence, or model misspecification),then the counterfactual model can be improved by adjusting for these patterns rather than treating them as pure noise.
- Gonçalves and Ng, 2024 highlights that many counterfactual estimation methods assume residuals are purely stochastic, but in practice, residuals often contain systematic variations that can be leveraged to enhance prediction accuracy.

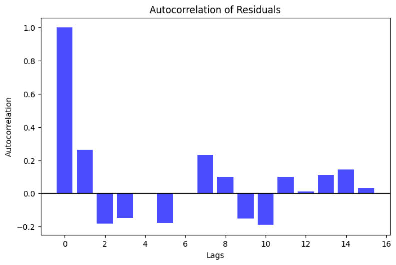
Motivating Example

- Consider estimating the economic impact of the 1990 German reunification on West Germany.
- let $Y_{1t} = \log$ GDP. Because the GDP data is non-stationary, use GDP growth.
- Estimate the counterfactual with a factor model using 16 OECD countries GDP.
- Compute the the (in-sample) residuals $\widehat{e}_{1,1:1T_0} = \Delta Y_{1,1:T_0}(0) - \widehat{\Delta Y}_{1,1:T_0}(0)$.
- The residuals are persistent. See Figure

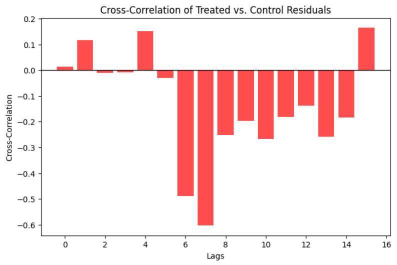
Residuals



Pre-unification Growth Rate Residuals
Residuals
Reunification (1990)
0.02
Residuals
Year
0.02
0.01
0.00
−0.01
−0.02
−0.03
1960
1965
1970
1975
1980
1985
1990



Autocorrelation of Residuals
Autocorrelation
Lags
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0
2
4
6
8
10
12
14
16



Cross-Correlation of Treated vs. Control Residuals
Cross-Correlation
Lags
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
0
2
4
6
8
10
12
14
16

Nonspherical Errors

- Gonçalves and Ng, 2024 examines how predictable errors (due to serial correlation or model misspecification) affect model-based imputation, making standard counterfactual estimates biased.
- propose a refined predictor of potential outcomes that accounts for correlated errors.
- They Propose the Practical Linear Unbiased Estimator (PLUP), based on the BLUP estimator by Goldberger, 1962.
- This proposed pre dictor is flexible and not limited to linear models (PUP).

Econometric Framework

Framework

- Let N_1 exposed units before the $N_0 = N - N_1$ unexposed units such that

$$Y_{it} = \begin{cases} Y_{it}(0), & i = 1, \ldots, N, & t = 1, \ldots, T_0 \\ Y_{it}(0), & i = N_1 + 1, \ldots, N, & t = T_0 + 1, \ldots, T \\ Y_{it}(1) & i = 1, \ldots, N_1, & t = T_0 + 1, \ldots, T \end{cases}$$

$$\delta_{i,T_0+h} = \underbrace{Y_{i,T_0+h}(1)} - \underbrace{Y_{i,T_0+h}(0)}$$

- $Y_{it}(0)$ has a pseudo true conditional mean (or mean-unbiased proxy) $m_{it} = \mathcal{M}(\beta; \mathcal{H})$ such that for each $i = 1, \ldots, N_1$;

$$\begin{aligned} Y_{it}(0) &= m_{it} + e_{it}, & t &= 1, \ldots, T \\ Y_{it}(1) &= m_{it} + \delta_{it} + e_{it}, & t &> T_0. \end{aligned}$$

- Let $\widehat{m}_{it}$ be a consistent estimate of m_{it}. Then $Y_{it}(0) = \widehat{m}_{it} + \widehat{e}_{it}$.

Prediction Error

- The treatment effect of unit i at a given time $t = T_0 + h$ is estimated as:

$$\hat{\delta}_{i,T_0+h} = Y_{i,T_0+h}(1) - \hat{Y}_{i,T_0+h}(0).$$
$$= Y_{i,T_0+h}(1) - Y_{i,T_0+h}(0) + \left(Y_{i,T_0+h}(0) - \hat{Y}_{i,T_0+h}(0) \right)$$
$$= \delta_{i,T_0+h} + m_{i,T_0+h} + e_{i,T_0+h} - \hat{m}_{i,T_0+h}$$

- The pointwise prediction error is:

$$\hat{\delta}_{i,T_0+h} - \delta_{i,T_0+h} = e_{i,T_0+h} + \left(m_{i,T0h} - \hat{m}_{i,T0+h} \right)$$

- This error has two sources of variation:
  - in-sample estimation of m_{it}
  - out-of-sample error e_{i,T_0+h} which does not vanish with sample size
- Therefore total prediction error variance is minimized asymptotically if m_{it} is chosen such that e_{it} does not contain predictable information.

Frame Title

EMORY UNIVERSITY

- To clarify the differences in the errors structure.
- suppose that unit 1 is being treated and the model is linear so that $m_{1t} = x_t'\beta$.
- The imputation error is given as: $\widehat{\delta}_{1,T_0+1} - \delta_{1,T_0+1} = -x_{T_0+1}'(\widehat{\beta} - \beta) + e_{1,T_0+1}$
- The variance of this prediction error:

$$\operatorname{var}\left(\widehat{\delta}_{1,T_0+1} - \delta_{1,T_0+1}\right) = \sigma_e^2 + x_{T_0+1}' \operatorname{var}(\widehat{\beta}) x_{T_0+1}$$

- Provided that $E\left[x_t e_{1t}\right] = 0, \implies \widehat{\beta}$ is consistent then $\operatorname{var}(\widehat{\beta}) \to 0$ as $T_0 \to \infty$
- Thus, the variance of imputation error is dominated by the out-of-sample error variance $\sigma_e^2 \equiv \operatorname{var}\left(e_{1,T_0+1}\right)$ asymptotically.
- This variance is minimized when e_{1,T_0+1} is uncorrelated.
- The residual correlation (serially or mutually) occurs from model misspecification.

Practical Unbiased Predictor

Practical Unbiased Predictor

- Recall the Best Linear Unbiased Predictor (BLUP) estimator Goldberger, 1962, which improves prediction by adjusting for the covariance structure of errors i.e.

$$y_t = X_t'\beta + e_t$$
$$y_t^{BLUP} = X'\beta_{GLS} + \Omega^{-1} e_{GLS}$$

- The Practical Unbiased Predictor (PUP) for serial correlation correction builds on Goldberger's Best Linear Unbiased Predictor (BLUP).
- Define the baseline predictor or model: $\widehat{Y}^{\text{baseline}}_{1,T_0+h} := \widehat{\mathcal{M}}_{1,T_0+h}\left(\boldsymbol{X}_{T_0+h};\boldsymbol{\theta}\right)$
- The PUP for serial correlation correction only is defined as

$$\widehat{Y}^{\text{PUP}}_{1,T_0+h} = \widehat{Y}^{\text{baseline}}_{1,T_0+h} + \widehat{\rho}_1^h \widehat{e}_{1,T_0}$$

- where $\hat{e}_{1,T_0} = Y_{1,T_0} - \widehat{Y}^{\text{baseline}}_{1,T_0+h}$ and $\widehat{\rho}_1 = \frac{\sum_{t=2}^{T_0} \widehat{e}_{1,t}\widehat{e}_{1,t-1}}{\sum_{t=2}^{T_0} \widehat{e}^2_{1,t-1}}$
- This is a simple correction, that takes into account the autorrelation structure.

EMORY UNIVERSITY

- On the other hand, the PUP for cross-correlation correction only is defined as

$$\widehat{Y}^{\mathrm{pup}}_{1,T_0+h} = \widehat{Y}^{\mathrm{baseline}}_{1,T_0+h} + \widehat{\boldsymbol{\pi}}_1 \widehat{\boldsymbol{e}}_{-1,T_0+h}$$

- where π_1 is the projection coefficient of $e_{1,t}$ on $\boldsymbol{e}_{-1,t} = (e_{2,t}, \ldots, e_{n,t})'$, where $e_{i,t} = Y_{i,t}(0) - \mathcal{M}_i\left(\boldsymbol{X}_t(0); \boldsymbol{\theta}\right)$.
- To account for both cross correlation serial correlation structure;

$$\widehat{Y}^{\mathrm{PUP}}_{1,T_0+h} = \widehat{Y}^{\mathrm{baseline}}_{1,T_0+h} + \sum_{s=0}^{p} \widehat{\rho}_{1s}\widehat{e}_{1,T_0-s} + \sum_{j=2}^{n}\sum_{s=-p}^{p} \widehat{\rho}_{j,s}\widehat{e}_{j,T_0+h-s}$$

  where $e_{i,t} = Y_{i,t}(0) - \mathcal{M}_i\left(\boldsymbol{X}_t(0); \boldsymbol{\theta}\right)$
- Note that we need the errors for all units now.it is a high-dimensional problem with $p + 1 + (n-1) \times (2p+1)$ parameters to be estimated.

Comment on Counterfactual Imputation with predictable Errors

PUP vs FarmTreat

- Fan et al., 2022 proposes FarmTreat which accounts accounts for high dimensional and nonstationarity settings.
- This method is a combination of latent factors and LASSO.
- The outcomes are modeled as:

$$
\begin{aligned}
Y_{i,t}(0) &= \gamma_i' \boldsymbol{W}_{i,t} + \lambda_i' \boldsymbol{F}_t + e_{i,t} \\
&= \gamma_1' \boldsymbol{W}_{1,t} + \boldsymbol{\lambda}_1' \boldsymbol{F}_t + \boldsymbol{\pi}_1' \boldsymbol{e}_{-1t} + v_t,
\end{aligned}
$$

  where $i, \boldsymbol{F}_t$ is a r-dimensional vector of common factors and λ_i its respective vector of loads for unit i; and $e_{i,t}$ is a zero mean idiosyncratic shock.
- π_1 is the coefficient of the linear regression of $e_{1,t}$ onto $\boldsymbol{e}_{-1t}$ and v_t the respective projection error.
- The authors also assume that $\boldsymbol{W}_{i,t}, \boldsymbol{F}_t$, and $e_{i,t}$ are mutually uncorrelated.

PUP vs FarmTreat

- Note that FarmTreat controls for cross-section dependence, and the parameter π_1 is estimated by LASSO regression of $\widehat{e}_{1,t}$ on $\widehat{\boldsymbol{e}}_{-1,t}$ on the pre-intervention sample.
- Therefore, FarmTreat is a PUP for cross-section correlation, with the baseline model being:

$$\widehat{Y}^{\text{baseline}}_{1,T_0+h} = \widehat{\gamma}_1' \boldsymbol{W}_{1,T_0+h} + \widehat{\boldsymbol{\lambda}}_1' \widehat{\boldsymbol{F}}_{T_0+h}.$$

- Medeiros, 2024 shows that FarmTreat can be easily modified into a PUP for both serial and cross-correlation correction:

$$\widehat{Y}_{1,T_0+h}(0) = \widehat{\gamma}_1' \boldsymbol{W}_{1,T_0+h} + \widehat{\boldsymbol{\lambda}}_1' \widehat{\boldsymbol{F}}_{T_0+h} + \widehat{\boldsymbol{\pi}}_1' \widehat{\boldsymbol{e}}_{-1,T_0+h} + \widehat{\phi} \widehat{e}_{1,T_0}$$

Simulation Design: Dependence Correction

- **Outcome equation:**

$$Y_{i,t} = \delta_{i,t} + \gamma_i' W_t + R_{i,t}$$

  - W_t includes a constant, trend, and Gaussian noise. - $R_{i,t}$ captures unobserved **factor dependence**.

- **Unobserved component:**

$$R_{i,t} = \lambda_i' F_t + e_{i,t}, \quad F_t = 0.8 F_{t-1} + \epsilon_t, \quad \lambda_i \sim \mathcal{N}(2, 1)$$

- **Serial and Cross-Sectional Correlation in Errors:**

$$e_{i,t} = \begin{cases} 0.8 e_{i,t-1} + \beta' e_{-1,t} + \nu_{i,t}, & \text{if } i = 1, \\ \nu_{i,t}, & \text{otherwise.} \end{cases}$$

  - Treated unit has **both autoregressive errors and cross-sectional dependence**.

- Consider the case $T = [75, 150]$ and $n = \{T, 2T, 3T\}$

Results

T	$n = T$	$n = 2 \times T$	$n = 3 \times T$
75	2.7285	2.1833	2.2728
150	2.2512	2.2536	2.1592

		Farmetreat	
	$n = T$	$n = 2 \times T$	$n = 3 \times T$
75	2.5448	2.0809	2.1877
150	2.0792	2.07513	2.0521

		Farmetreat + AR	
	$n = T$	$n = 2 \times T$	$n = 3 \times T$
75	1.2862	0.7940	0.9239
150	0.7670	0.5744	0.6215

PUP vs ArCo

PUP vs ArCo

- Recall that the ArCo model accounts for cross-section dependence given:

$$\widehat{Y}_{1,T_0+h}(0) = \widehat{\boldsymbol{\theta}}_1 \boldsymbol{Y}_{-1,T_0+h}(0)$$

- $\boldsymbol{\theta}_1$ is estimated by LASSO regression of $Y_{1,t}$ on $\boldsymbol{Y}_{-1,t}$ on the pre-intervention sample.

$$Y_{1t}^{(0)} = \theta_0' Y_{-1,t}^{0} + v_t$$

- The ArCo can be extended to incorporate serial correlation correction i.e.

$$\widehat{Y}_{1,T_0+h}(0) = \widehat{\boldsymbol{\theta}}_1 \boldsymbol{Y}_{-1,T_0+h}(0) + \widehat{\theta}_L Y_{1,T_0}(0)$$

- where both $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_L$ can be estimated by LASSO.
- Therefore the ArCo can be extended as PUP estimator

Potential Extension

PuP vs VAR

- One Potential Extension is to model the time dynamics as a VAR i.e.

$$\begin{bmatrix} 1 & \theta' \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} Y_{1,t}(0) \\ \boldsymbol{Y}_{0t}(0) \end{bmatrix}}_{:=\boldsymbol{Y}_t} = \boldsymbol{A}_0 + \boldsymbol{A}_1 \boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{A}_p \boldsymbol{Y}_{t-p} + \boldsymbol{e}_t$$

- This is a potentially high-dimensional VAR model, and some sort of regularization is needed.
- The VAR framework allows for potential feedback from the treated to the controls and could be a nice extension to capture spillover effects.
- This issue has been mostly neglected in the synthetic control literature.
- Another potential extension of model is to include factors in the model as in the FAVAR.

Conclusion

Thanks!

benjamin.harrison@emory.edu

benhars.com

@benhars

References

- Ben-Michael, E., Feller, A., & Rothstein, J. (2021).**The augmented synthetic control method.** *Journal of the American Statistical Association*, *116*(536), 1789–1803.
- Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015).**Inferring causal impact using bayesian structural time-series models.**
- Carvalho, C., Masini, R., & Medeiros, M. C. (2018).**Arco: An artificial counterfactual approach for high-dimensional panel time-series data.** *Journal of econometrics*, *207*(2), 352–380.
- Fan, J., Masini, R., & Medeiros, M. C. (2022).**Do we exploit all information for counterfactual analysis? benefits of factor models and idiosyncratic correction.** *Journal of the American Statistical Association*, *117*(538), 574–590.
- Goldberger, A. S. (1962).**Best linear unbiased prediction in the generalized linear regression model.** *Journal of the American Statistical Association*, *57*(298), 369–375.

Gonçalves, S., & Ng, S. (2024). **Imputation of counterfactual outcomes when the errors are predicatable.** *Journal of Business & Economic Statistics*, (just-accepted), 1–33.

Medeiros, M. C. (2024). **Counterfactual imputation: Comments on "imputation of counterfactual outcomes when the errors are predictable" by silvia gonçalves and serena ng.** *Journal of Business & Economic Statistics*, *42*(4), 1128–1132.