

tfe

Manoel Galdino

2024-05-20

## Introdução

Cerca de 20 anos atrás a literatura metodológica de ciência política sobre dados de Série de Tempo com Corte Transversal (TSCS na sigla em inglês, de Time Series Cross Section) estava devotada integralmente para as dificuldades de inferência estatística, isto é, como computar corretamente o erro padrão nesse tipo de regressão.

Após a revolução da credibilidade na economia e que se espalhou pela ciência política, os estudos mudaram para questões de identificação causal, deixando as questões de inferência estatística em segundo plano. O objetivo do presente trabalho é fazer um survey dessa literatura mais recente, feita nos últimos dez anos e ainda em franco desenvolvimento, de modo a atualizar o cientista político que fez seu treino metodológico sem contato com essa literatura.

## Literatura de Dados de Paine

A partir dos anos 80, emergiu uma literatura metodológica com ênfase em dados de Paine na economia (Arellano (2003), cf. Arkhangelsky and Imbens (2023)). Uma das questões primordiais dessa literatura dizia respeito a como modelar efeitos idiossincráticos das unidades que não variavam no tempo, isto é, quando usar modelos de efeitos fixos ou aleatórios (cf. Bell and Jones (2015) para uma discussão mais recente). Modelos de efeitos fixos seriam mais adequados quando houvesse correlação com covariáveis, enquanto efeitos aleatórios quando a correlação fosse inexistente. A partir disso, uma grande parte da literatura foi desenvolvida para garantir a estimação com T pequeno.

Motivado por questões substantivas, a dinâmica sempre foi algo importante na literatura de ciência política, como refletida pelos títulos de artigos hoje clássicos de Katz & Becker sobre levar a dimensão temporal seriamente, em particular a questão de variável dependente defasada e como modelar a dinâmica resultante.

## Notas Sol

Um aspecto negligenciado na literatura metodológica de ciência política e mencionado por Arkhangelsky and Imbens (2023) é a distinção entre dois tipos de dados: dados de TSCS clássicos, em que temos respostas  $y_{it}$  e tratamento  $D_{it}$  e o que eles chamam de row-column exchangeable data (RCED). O exemplo que eles dão são de consumidores  $i$  comprando itens  $j$ , ou no caso de corridas de aplicativo, consumidores  $i$  e motoristas  $j$ . Aqui, temos respostas  $y_{ij}$  e tratamento  $D_{ij}$ , isto é, dados duplamente indexados. A principal diferença entre dados de TSCS e RCED é que a ordem não importa nesse último, ou seja, podemos considerar não apenas as unidades  $i$  como permutáveis, mas também o segundo índice  $j$ . No caso de dados de TSCS, o índice temporal não é permutável. E no entanto, modelos de regressão com dois efeitos fixos e vários métodos de controle sintético podem ser usados igualmente para esse tipo de dado, o que significa que tais métodos não aproveitam a informação presente na ordem temporal dos dados para estimação de parâmetros. Onde tem havido algum tipo de atenção metodológica em ciência política é na análise de dados diádicos, típicos de relações internacionais. Não iremos abordar esse aspecto em nosso survey, mas recomendamos Bonhomme (2020) em econometria e Poast (2016).

## Within e Between

Considere que estamos interessados em estimar o efeito de instituições legadas pelo colonialismo sobre o desenvolvimento econômico dos países, como em Acemoglu, Johnson, and Robinson (2001). Uma formulação útil (Bell and Jones (2015)) para nossos propósitos é pensar no efeito das instituições do país  $i$  no período  $t$ ,  $D_{it}$  como sendo composta de dois efeitos: um efeito dentro do país “within” e um efeito entre países “between”:  $D_{it} = D_i^B + D_{it}^W$ . Em nosso exemplo, pode haver variação no efeito de instituições entre países ( $D_i^B$ ) devido à fatores usualmente omitidos das regressões, como questões geográficas, cultura política exportada pelos colonizadores etc. E pode haver variação no tempo para cada país, pois as instituições podem ter dependência de caminho, levando-as a se reproduzirem ao longo do tempo e impactando a renda per capita dos países.

A figura de Bartels (2008) é ilustrativa a esse respeito:

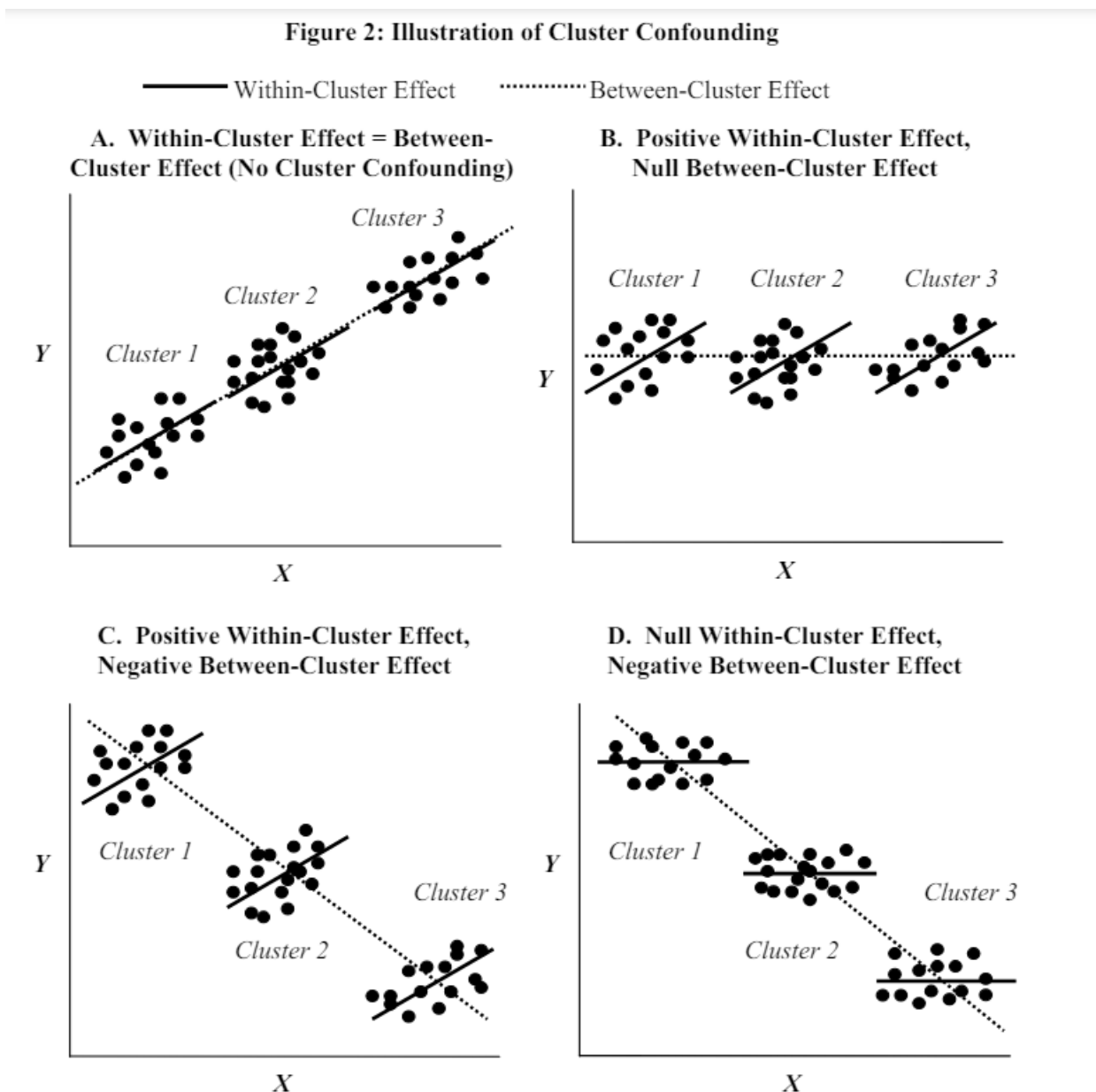
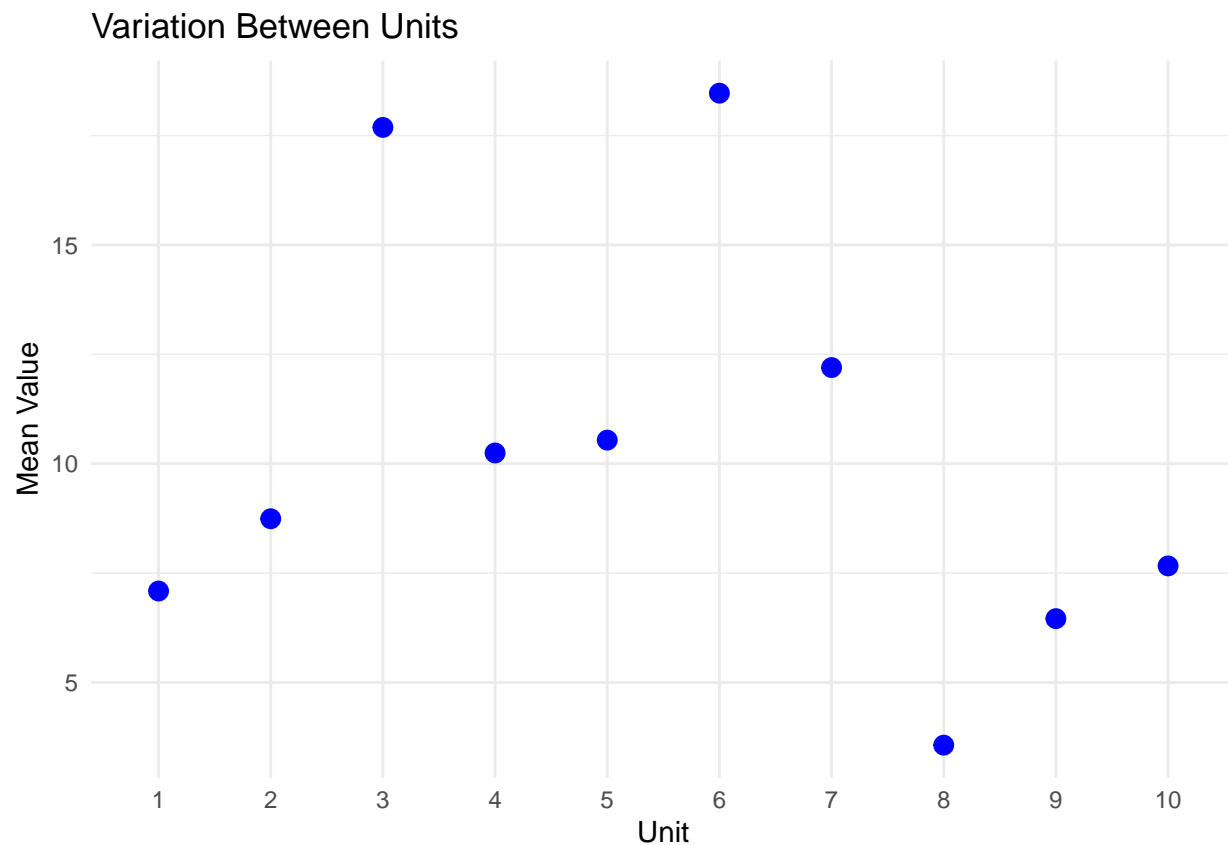
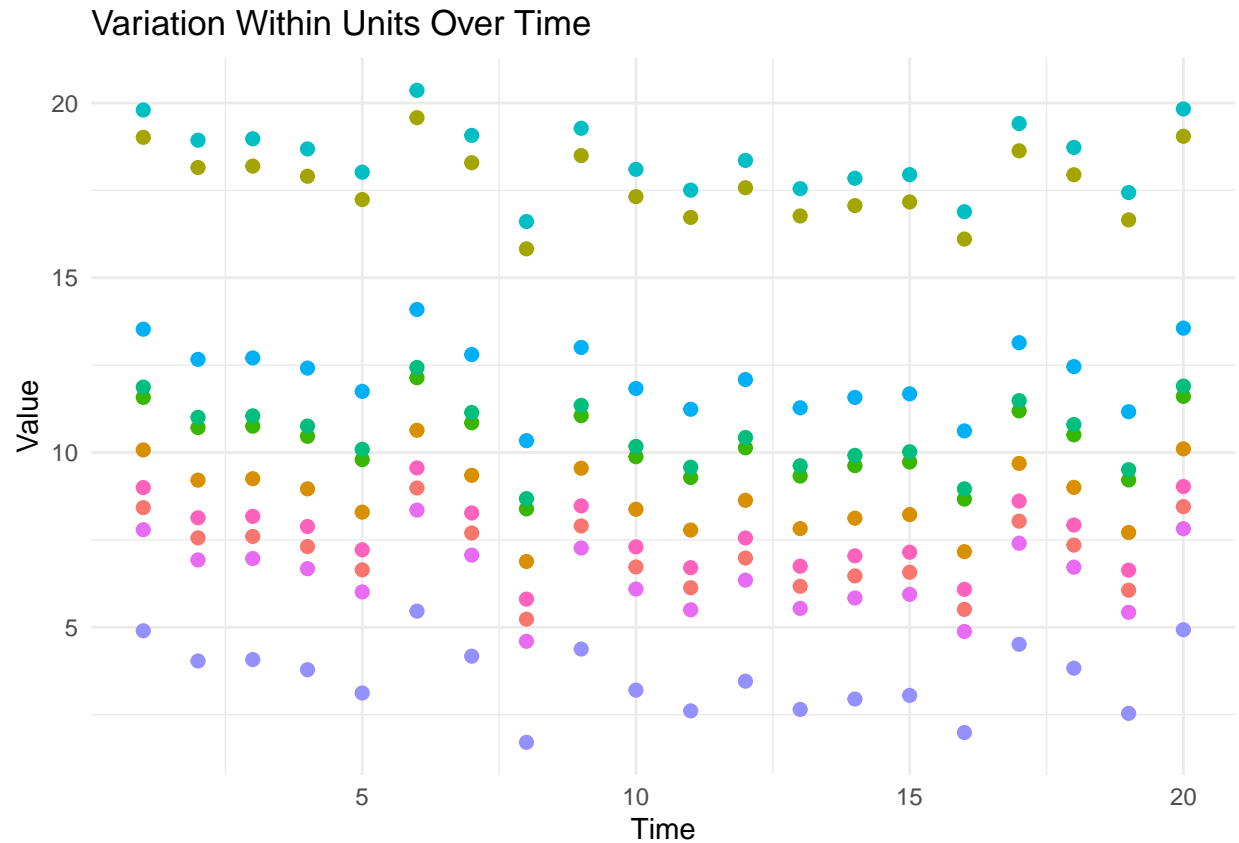


Figure 1: Descriptive text

Se a gente considerar que o painel D apresenta as unidades (indivíduos, países, estados, municípios, escolas etc.) como clusters, temos o caso arquetípico das variações não-observadas entre unidades capturadas pelos efeitos fixos das unidades. Já se os clusters representarem a dimensão temporal, isto é, cada cluster uma unidade de tempo (ano, mês, trimestre, dia etc.), então os efeitos fixos de tempo capturam a variação não-observada temporal.





## DAGs

O trabalho de Myint (2024) apresenta DAGs que codificam as suposições principais de identificação de TSCS com sequencial ignorability e paralell trends.

![Descriptive teDAGs tscs.png.png)

Texto sobre DAGs # References

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.
- Arellano, Manuel. 2003. *Panel Data Econometrics*. OXFORD.
- Arkhangelsky, Dmitry, and Guido Imbens. 2023. "Causal Models for Longitudinal and Panel Data: A Survey." National Bureau of Economic Research.
- Bartels, Brandon. 2008. "Beyond" Fixed Versus Random Effects": A Framework for Improving Substantive and Statistical Analysis of Panel, Time-Series Cross-Sectional, and Multilevel Data." *The Society for Political Methodology* 9: 1–43.
- Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3 (1): 133–53.
- Bonhomme, Stéphane. 2020. "Econometric Analysis of Bipartite Networks." In *The Econometric Analysis of Network Data*, 83–121. Elsevier.
- Myint, Leslie. 2024. "Controlling Time-Varying Confounding in Difference-in-Differences Studies Using the Time-Varying Treatments Framework." *Health Services and Outcomes Research Methodology* 24 (1): 95–111. <https://doi.org/10.1007/s10742-023-00305-2>.
- Poast, Paul. 2016. "Dyads Are Dead, Long Live Dyads! The Limits of Dyadic Designs in International Relations Research." *International Studies Quarterly* 60 (2): 369–74.