Matrix Completion Estimation in Varying Panel Data Settings

as discussed with Prof. Martin Schumann

In Athey et al. (2021), the authors synthesize the literature on Causal Inference in Panel Data Econometrics. They characterize the common goal of prominent approaches in this field as estimating average (causal) treatment effects by imputing missing potential outcomes. They show that two popular frameworks for causal inference, the unconfoundedness and synthetic control approaches, can be viewed as matrix completion (MC) estimators. They then introduce a new MC estimator that they claim shows improved accuracy compared to existing methods in certain settings. MC estimation (MCE) is a technique that so far has been used extensively in the Computer Science and Statistics literatures, but has not been widely adopted in Econometrics. The empirical part of their paper is focused on comparing the accuracy of their new estimator to several other methods, including a Difference-in-Differences (DID) approach.

Since the initial publication of Athey et al. (2021), there has been a deluge of advancements in the DID literature. In the characterization of Roth et al. (2023), substantial parts of this frontier work focus on assessing and reducing estimation bias when canonical assumptions of DID are relaxed (cf. Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfoeuille, 2018; de Chaisemartin & D'Haultfœuille, 2020; Goodman-Bacon, 2021). There is currently no literature that incorporates these advances into the MCE framework.

The main goals of this proposed thesis are threefold:

1. To review and synthesize the contributions to Panel Data Econometrics made in Athey et al. (2021),

2. To examine whether recent advances in DID methods can be incorporated into the MCE framework, and

3. To compare the accuracy of Athey et al. (2021)'s new MC estimator with that of recent DID methods in data regimes in which canonical DID assumptions are not met.

References

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix Completion Methods for Causal Panel Data Models [Appendix at https://arxiv.org/pdf/1710.10251.pdf]. *Journal of the American Statistical Association*, *116*(536), 1716–1730. https://doi.org/10.1080/01621459.2021.1891924

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods [Publisher: Elsevier]. *Journal of Econometrics*, *225*(2), 200–230.

de Chaisemartin, C., & D'Haultfoeuille, X. (2018). Fuzzy Differences-in-Differences. *The Review of Economic Studies*, *85*(2), 999–1028. https://doi.org/10.1093/restud/rdx049

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature [arXiv:2201.01194 [econ, stat]]. https://doi.org/10.48550/arXiv.2201.01194