

# Recent Developments in Methods for Causal Inference with Cross-Sectional Time-Series Data\*

## *Workshop Notes*

Roberto Valli<sup>†</sup>

December 22, 2022

## 1 Introduction

Political science is often characterized by research questions that cannot be studied experimentally because ethical or practical reasons prevent treatment manipulation. The impossibility of using true randomization in polisci research makes all the more important to find alternative ways to estimate causal effects of policies and other interventions. These notes discuss recent developments in a specific family of statistical methods that exploit within and across-unit variation over time to approximate estimates of causal effects.

The goal of these notes is to tie together a recent literature that highlights the limitations of standard difference-in-differences methods (DiD) when it comes to estimating causal effects in complex settings that violate its basic assumptions. The interesting thing about this literature is that not only it develops extremely fast, but it accompanies a harsh critique of common modelling choices to many solutions, practical advice, and R packages to implement more suitable models. However, the fast-moving and technical nature of the literature can be confusing. This is why I wrote these notes, to guide applied researchers to the right references and modelling choices without having to first read everything under the sun.

Note that this document is meant to be a practical guide for advanced researchers with prior knowledge of statistics, basic causal inference designs and programming in R. It is by no means a technical paper, nor an introduction to any of these subjects.

The notes proceed as follows. First, I give a brief overview of causal inference methods with cross-sectional time-series data (CSTS) and its limitations. I start from the standard DiD model and its estimators, and walk through its critique in the literature. Second, I present a heuristic decision model to select the right CSTS method based on one's data. Third, I provide a description of various methods, their assumptions and implementation in R.

---

\*Find the workshop's GitHub repo at [https://github.com/RobertoValli/panel\\_models\\_workshop](https://github.com/RobertoValli/panel_models_workshop).

<sup>†</sup>PhD candidate, International Conflict Research, ETH Zurich: [roberto.valli@icr.gess.ethz.ch](mailto:roberto.valli@icr.gess.ethz.ch)

## 2 Causality and Cross-Sectional Time-Series Data

The problem of estimating causal effects can be thought of as the challenge of predicting counterfactual outcomes. The impossibility of observing the outcome of a process if something (the treatment condition) had been different is often referred to as the *fundamental problem of causal inference*. The Neymar-Rubin counterfactual outcome framework formalizes one way of thinking about causality in terms of counterfactuals.

Consider an outcome of interest  $Y_{it}$  that is observed for individual  $i$  at time  $t$ . Also, consider a binary treatment  $D_{it} \in \{0, 1\}$ . The individual causal effect of  $D$  on  $Y$  can be defined as

$$\text{ICE} = Y_{it}(1) - Y_{it}(0)$$

However, since only one of the two terms is ever observed in reality, statistical methods of causal inference find ways to estimate *average causal effects* by predicting the counterfactual outcome under the unobserved treatment condition. In fact, while applied statistics are often divided between causal inference and prediction tasks, prediction is very much at the heart of causal inference methods, although with a strong focus on bias reduction.

Let's consider the classic two-period DiD setup and how it can be thought of as a prediction exercise. The outcome is assumed to be a continuous variable  $Y_{it}$  observed over two periods  $t = 1, 2$ , whereas the treatment is assumed to be a dichotomous variable  $D_{it} = 0, 1$ . The observed value of the outcome can be expressed as:

$$Y_{it} = Y_{it}(1)D_{it} + Y_{it}(0)D_{it}$$

The major contribution of the classic DiD literature (**Angrist2009**) is to demonstrate that under SUTVA<sup>1</sup> assumptions the average treatment effect on the treated (ATT) can be defined as

$$\tau = E(Y_{it}(1) - Y_{it}(0) | D_{it} = 1)$$

Under the additional assumption of so-called parallel trends, meaning that the individual *change* in outcome under control between  $t = 1, t = 2$  is independent of treatment status, i.e.

$$E(Y_{i2}(1) - Y_{i1}(0) | D_{it} = 1) = E(Y_{i2}(1) - Y_{i1}(0) | D_{it} = 0)$$

it is possible to statistically identify the ATT as

$$\tau = E(Y_{it}(1)) - E(Y_{it}(0))$$

Now, in the basic two-period setting there are multiple ways of estimating  $\tau$ , all of which produce equivalent estimates. One might use a non-parametric approach by estimating a difference in means as done in the classic **Card1990empty citation** paper. Alternatively, one might use a two-way fixed effects regression (TWFE) or a first-difference regression which both have the advantage of producing measures of statistical uncertainty (**Angrist2009**).

However, as soon as one's case of interest goes beyond this simple setup, the assumptions and common estimation methods for the ATT start to be violated. Next, I discuss some of the major problems identified by the literature and touch on the solutions out there.

---

<sup>1</sup>The *stable unit treatment value assumption* is a set of assumptions such that: i) units composition does not change over time, ii) no treatment heterogeneity exists, iii) no treatment spillovers take place.

### 3 When the Basic DiD Design Breaks Down

The simplest violation of the basic DiD setup emerges when units are observed for more than two periods  $t = 1, 2, \dots, N$ . In this case, **Angrist2009** already note how the first-difference model cannot be considered an unbiased estimator of  $\tau$ , since taking first-differences over multiple periods introduces autocorrelation of the outcome and violates the independence of the error term. Nevertheless, the same is not true for the two-way fixed effects model, which can be used in multiple-period settings.

The most significant result in the recent literature on CSTS estimators, and probably the one that contributed the most to its quick development, is the demonstration that the TWFE estimator is a biased estimator of  $\tau$  in the common setting of *staggered treatment adoption*. Whenever treated units are not exposed to treatment at the same point in time, but rather adopt treatment in several cohorts, the TWFE estimate of  $\hat{\tau}$  is biased towards zero (**GoodmanBacon2021; Callaway2021**). Researchers demonstrated that TWFE estimates are a weighted average of DiD estimates comparing all pairs of periods. The major issue is that some comparisons are undesirable, meaning that they use already-treated cohorts as counterfactual for other treated units. A second problem is that some of the cohort-effects' weights are negative, which creates problem with the interpretation of the aggregate estimate of  $\hat{\tau}$ .