9^{th} International Conference on Computational Social Science IC^{2}S^2
July 17-20, 2023, Copenhagen, Denmark

Ameliorating the Problem of Temporal Validity When Studying Fast-Moving Systems

Keywords: Metascience; Temporal Validity; External Validity; Quantitative Description; Digital Media

Extended Abstract

The "credibility revolution" has forced computational social scientists to confront the limits of our methods for creating general knowledge. As a result, many practitioners aim to generate valid but local knowledge and then synthesize and apply that knowledge to predict what will happen in a target context. Positivist social science has until recently been hamstrung with other, more immediate threats to validity and inference, but we argue that recent advances in statistical approaches to the problem of external validity reveal limits of the current paradigm. This article and the term "temporal validity" illustrate the intrinsic limits of agnostic (that is, assumption-free) external validity when the target setting is in the future.

These limits, we argue, suggest a re-orientation of social science methodology. We should acknowledge that no research design, no empirical knowledge, is perfectible; instead, we should explicitly aim to increase the amount and quality of knowledge we produce. It is useful to characterize this perspective as "Meta-Science," an emerging social/intellectual movement within the social sciences. "Temporal validity" and the implied "knowledge decay" thus represent a meta-scientific intervention aimed at increasing the usefulness of the knowledge we produce.

Indeed, we use the term "temporal validity" to make our critique legible to social scientists who are used to thinking in terms of "validity," but our statistical formulation reveals the limitations of this conceptual frame. "Validity" implies a binary; we cannot really talk about a finding or research design being "somewhat valid" or "somewhat invalid." Either of these terms would immediately be read as "not valid."

So we translate the problem of generalizability into the equally familiar but more continuous consideration of statistical power, drawing in particular on the idea of the effective sample size as used in (for example) [Aronow and Samii, 2016].

Following the logic of agnostic statistics, we develop a statistical approach to bounding the "knowledge decay" resulting from rapidly changing objects of inquiry—a first order concern for many computational social scientists. This approach is intrinsically metascientific. Traditional social science methodology is oriented towards "validity"—that is, towards crafting more perfect papers. The problem of the imperfectibility of social science becomes obvious if we want to apply the knowledge we have created and stored in an academic article.

We conceptualize this problem by modelling a larger "research program" composed of a community of scholars and their resources (time and money). Following the paradigm of "solution-oriented social science" [Watts, 2017] or "the economist as plumber" [Duflo, 2017] we posit the optimand of this research program as the contribution that the knowledge they produce can make to the solution of a problem; that is, the "lift" contributed by that knowledge to a prediction about the outcome of some action or policy change.

Crucially, by the definition of time, all of the knowledge we produce is about the past, and all action is in the future.

We assume that the research program thus has a fixed resource endowment and model the tradeoff faced between conducting more causal research (ideally, RCTs) and conducting more quantitative description. The latter is necessary for bounding the rate of knowledge decay caused by changing covariate space and changing treatment effects.

The following proposition is the heart of our statistical model; Figure 1 is a schematic representation of the relevant parameters. The specific decision calculus is of course contingent on the parameter values in a given context. In particular, the relative cost of descriptive and causal knowledge, researcher intuitions about the rate of change of the system under study, and the degree of treatment effect heterogeneity. Our framework gives social scientists practical ways to reason about these questions and even to estimate relevant parameter values.

Proposition 1 Weakening standard assumptions for external validity [c.f. Egami and Hartman, 2022], assume that the right covariates are measured for ignorability of setting, but positivity does not necessarily hold. Also assume that treatment effects conditional on these covariates do not change by more than L in one unit of time, nor does the conditional treatment effect function change by more than M for a one unit change in covariate space.

Then the temporal validity error from period t_0 to period t is upper bounded by:

$$L(t-t_0)+M\,W_1(\mathbb{P}(\mathbf{X};t),\mathbb{P}(\mathbf{X};t_0))$$

This bound is tight without making strong exclusion restrictions or parametric assumptions. Examples of such assumptions would be that treatment effects are constant ($L=M=0$), that treatment effects do not change over time ($L=0$) or that there is change in covariate values over time $W_1(\mathbb{P}(\mathbf{X};t),\mathbb{P}(\mathbf{X};t_0))=0$. In other settings for causal inference, such assumptions are viewed as implausibly strong, such as in the growing literature on difference–in–difference designs [e.g. Callaway and Sant'Anna, 2021]. The latter can be directly estimated using off-the-shelf computational methods in polynomial time, but requires regularly collecting relevant descriptive data.

References

Peter M Aronow and Cyrus Samii. Does regression produce representative estimates of causal effects? American Journal of Political Science, 60(1):250–267, 2016.

Brantly Callaway and Pedro HC Sant'Anna. Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2):200–230, 2021.

Esther Duflo. The economist as plumber. American Economic Review, 107(5):1–26, 2017.

Naoki Egami and Erin Hartman. Elements of external validity: Framework, design, and analysis. American Political Science Review, page 1–19, 2022. doi: 10.1017/S0003055422000880.

Duncan J Watts. Should social science be more solution-oriented? Nature Human Behaviour, 1(1):0015, 2017.

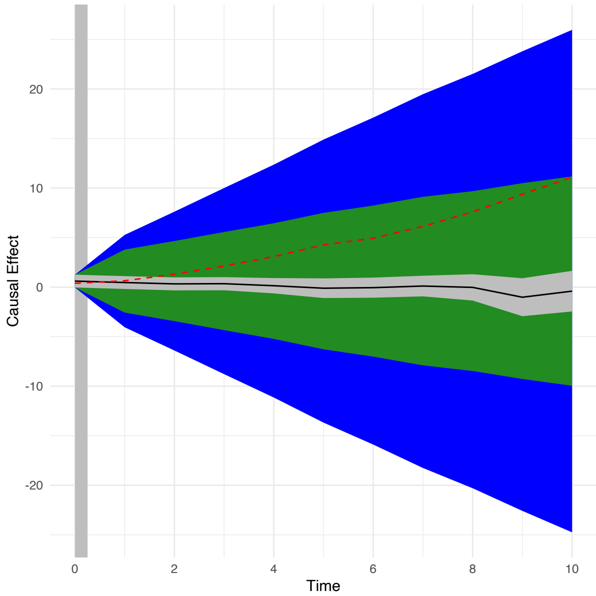

Figure 1: Example non-parametric bounds on a causal effect as a function of time. The blue region represents the impact of potentially changing treatment effects, while the green region represents the growth in uncertainty due to covariate shift. The grey region denotes the naïve confidence intervals of a reweighting estimator which grossly underestimate the epistemic uncertainty.