When do common time series estimands have nonparametric causal meaning?*

Ashesh Rambachan[†] Neil Shephard[‡]

October 12, 2021

Abstract

In this paper, we introduce the nonparametric, direct potential outcome system as a foundational framework for analyzing dynamic causal effects of assignments on outcomes in observational time series settings. Using this framework, we provide conditions under which common predictive time series estimands, such as the impulse response function, generalize impulse response function, local projection, and local projection instrument variables, have a nonparametric causal interpretation in terms of such dynamic causal effects.

Keywords: causality, instrumental variable, potential outcome, prediction, shock, time series.

1 Introduction

In this paper, we introduce the nonparametric, direct *potential outcome system* as a foundational framework for analyzing dynamic causal effects of assignments on outcomes in observational time series settings. We consider settings in which there is a single unit (e.g., macroeconomy or market) observed over time. At each time period $t \geq 1$, the unit receives a vector of assignments W_t, and an associated vector of outcomes Y_t are generated. The outcomes are causally related to the assignments through a *potential outcome process*, which is a stochastic process that describes what would be observed along counterfactual assignment paths. A dynamic causal effect is generically defined as the comparison of the potential outcome process along different assignment paths at a fixed point in time.

*This is a revised version of "A nonparametric dynamic causal model for macroeconometrics." We thank Iavor Bojinov, Gary Chamberlain, Fabrizia Mealli, James M. Robins, Frank Schorfheide and James H. Stock for conversations that have helped developed our thinking on causality. We are also grateful to many others, including conference participants, for their comments on earlier versions of this paper. All remaining errors are our own.

[†]Harvard University, Department of Economics: asheshr@g.harvard.edu

[‡]Harvard University, Department of Economics and Department of Statistics: shephard@fas.harvard.edu

Importantly, we place no functional form restrictions on the potential outcome process, no restrictions on the extent to which past assignments may causally affect outcomes, nor common time series assumptions such as "invertibility" or "recoverability" on the system of potential outcomes and assignments. Most leading econometric models used to study dynamic causal effects in time series settings, such as the structural vector moving average model and (both linear and non-linear) structural vector autoregressions, can be cast as special cases of the direct potential outcome system by introducing these additional restrictions on the potential outcome process or the full system. In this sense, the direct potential outcome system provides a flexible, nonparametric foundation upon which to analyze dynamic causal effects in time series settings.

We then analyze conditions under which predictive time series estimands, as the impulse response function (Sims, 1980), generalized impulse response function (Koop et al., 1996), local projection (Jordá, 2005) and local projection instrumental variables (Jordá et al. (2015)), have a nonparametric causal interpretation in terms of dynamic causal effects of assignments on outcomes. That is, under what conditions do these common time series estimands have a nonparametric causal interpretation as measuring how movements in the outcomes Y_{t+h} for some $h \geq 0$ are dynamically caused by changes in the assignments W_t? In exploring this question, we focus on four data environments, which place alternative assumptions on what output is observed by the researcher.

First, we analyze a benchmark case in which the researcher directly observes both the outcomes Y_{t+h} and the assignments W_t generated by the potential outcome system through time. We show that impulse response functions, local projections, and generalized impulse response functions of the outcome Y_{t+h} on the assignment W_t identify a dynamic average treatment effect, a weighted average of marginal average treatment effects, and a filtered average treatment effect respectively if the assignments W_t are randomly assigned. Random assignment requires that the assignment must be independent of the potential outcome process (which is familiar from cross-sectional causal inference) *and* the assignments must be independent over time. These results provide an interpretation of a macroeconomic "shock" purely in terms of assumptions on the assignment process. Furthermore, these results provide a new perspective on a rapidly growing empirical literature in macroeconomics that attempts to construct measures of underlying economic shocks, and then use these constructed measures to estimate dynamic causal effects using reduced form methods. Nakamura and Steinsson (2018b) refers to this empirical strategy as "direct causal inference." Our first set of results therefore provide conditions that these constructed shocks must satisfy in order for the resulting reduced form estimates to be causally interpretable.

Second, we provide a special case of the direct potential outcome system to incorporate instrument variables Z_t that causally affect the assignment W_t but not the outcome Y_{t+h}. Provided the researcher directly observes the instrument, the assignment, and outcome, it is natural to consider the causal interpretation of the time-series analogue of common instrumental variable estimands. We focus attention on dynamic instrument variable estimands that take the ratio of an impulse response function of the outcome Y_{t+h} on the instrument Z_t (a "reduced-form") relative to an impulse response function of the assignment W_{t+h} on the instrument Z_t (a "first stage"). We show that such dynamic instrumental variable estimands identify an appropriately defined dynamic "local average treatment effect" in the spirit of Imbens and Angrist (1994), provided the instrument is randomly assigned and satisfies a monotonicity condition that is familiar from cross-sectional causal inference. The dynamic local average treatment effects that we characterize measure the average dynamic treatment effect of the assignment on the h-period ahead outcome conditional on the event that the instrument causally affects the treatment.

We further analyze the case in which the researcher only observes the instrument Z_t and outcome Y_{t+h} but not the treatment W_t itself. This is an important case. Empirical researchers in macroeconomics increasingly use "external instruments" to identify the dynamic causal effects of *unobservable* economic shocks on macroeconomic outcomes (e.g., see Jordá et al., 2015; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018a; Ramey and Zubairy, 2018; Stock and Watson, 2018; Plagborg-Møller and Wolf, 2020; Jordá et al., 2020). In this research, it is common for empirical researchers to analyze estimands that involve two distinct elements of the outcome vector $Y_{j,t+h}, Y_{k,t}$ and the instrument. For example, given an external instrument Z_t for the unobserved monetary policy shock W_t (e.g., Kuttner, 2001; Cochrane and Piazessi, 2002; Gertler and Karadi, 2015), it is common to measure the dynamic causal effect of the monetary policy shock on unemployment $Y_{j,t+h}$ by (1) estimating a reduced-form impulse response function of unemployment on the external instrument, (2) estimating a "first stage" impulse response function of the Federal Funds rate $Y_{k,t}$ on the external instrument, (3) report the ratio of these impulse responses (e.g., Jordá et al., 2015; Stock and Watson, 2018; Jordá et al., 2020).

We show that dynamic IV estimands constructed in this manner are causally interpretable, and nonparametrically identify a *relative*, dynamic local average treatment effect that measures the causal response of the h-step ahead outcome $Y_{j,t+h}$ to a change in the treatment $W_{k,t}$ that raises the contemporaneous outcome $Y_{k,t}$ by one unit among compliers (i.e., in the monetary policy example, the dynamic causal effect of unemployment to a monetary policy shock that raises the Federal funds rate by one unit). This result therefore provides a motivation for the recent surge in interest in external instruments in empirical macroeconomics — provided one exists, an external instrument can be used to identify causally interpretable estimands without resorting to functional form restrictions on the outcomes and without even directly observing the treatment itself.

Finally, we conclude by briefly discussing the most challenging data environment in which the researcher only observes the outcomes Y_{t+h}, but not the treatments W_t nor any external instruments Z_t. This is the leading setting considered by much of the foundational and influential research on model-based approaches to analyzing dynamic causal effects in macroeconometrics (e.g., Sims, 1972, 1980). We consider this setting in order to place the direct potential outcome system in this broader context, and illustrate that researchers can recover familiar model-based approaches by introducing functional form restrictions on the potential outcome process.

Taking a step back, quantifying dynamic causal effects is one of the great themes of the broader time series literature. Researchers use a variety of methods such as Granger causality (Wiener, 1956; Granger, 1969; White and Lu, 2010), highly structured models such as DSGE models (Herbst and Schorfheide, 2015), state space modelling (Harvey and Durbin, 1986; Harvey, 1996; Brodersen et al., 2015; Menchetti and Bojinov, 2021) as well as intervention analysis (Box and Tiao, 1975) and regression discontinuity (Kuersteiner et al., 2018). The nonparametric potential outcome framework we develop is distinct. References to some of the more closely related work will be given in the next section. This paper is not focused on estimators and the associated distribution theory: we do not have much to say in that regard which is novel.

Roadmap: Section 2 defines the direct potential outcome system and introduces the main class of dynamic causal effects that we focus on throughout the paper. Section 3 looks at the causal meaning of common statistical estimands based on seeing the realized assignments and outcomes. The instrumented potential outcome system is defined in Section 4, which relates assignments and instruments to outcomes. Section 5 studies the causal interpretation of estimands based on seeing the realized assignments, instruments and outcomes. Section 6 looks at the causal meaning of estimands where only the instruments and outcomes are observed. Section 7 looks at the causal meaning of common statistical estimands where only the outcomes are observed. We collect all proofs in the Appendix.

Notation: For a time series $\{A_t\}_{t\geq 1}$ with $A_t \in \mathcal{A}$ for all $t \geq 1$, let $A_{1:t} := (A_1, \ldots, A_t)$ and $\mathcal{A}^t := \bigtimes_{s=1}^{t} \mathcal{A}$. $A \perp\!\!\!\perp B$ says that random variables A and B are probabilistically independent.

2 The Direct Potential Outcome System and Dynamic Causal Effects

We now introduce the *direct potential outcome system*, which extends the design-based approach developed in Bojinov and Shephard (2019) to stochastic processes. We define a large class of casual estimands that summarize the dynamic causal effects of varying the assignment on future outcomes. As an illustration, we show that the direct potential outcome system nests most leading structural models in macroeconometrics as a special case.

The nonparametric potential outcome framework we develop relates to a vast literature on dynamic treatment effects in small-T, large-N panels. The panel work of Robins (1986) and

Abbring and Heckman (2007), amongst others, led to an enormous literature on dynamic causal effects in panel data (Murphy et al., 2001; Murphy, 2003; Heckman and Navarro, 2007; Lechner, 2011; Heckman et al., 2016; Boruvka et al., 2018; Lu et al., 2017; Blackwell and Glynn, 2018; Hernan and Robins, 2021; Bojinov et al., 2021; Mastakouri et al., 2021). Beyond Bojinov and Shephard (2019), our work is most closely related to Angrist and Kuersteiner (2011) and Angrist et al. (2018). We discuss their work in Section 2.3.

2.1 The Direct Potential Outcome System

There is a single unit. At each time period $t \geq 1$, the unit receives a d_w-dimensional assignment $\{W_t\}_{t\geq 1}$. Associated with this *assignment process*, we observe a d_y-dimensional outcome $\{Y_t\}_{t\geq 1}$. The outcomes are causally related to the assignments through the potential outcome process, which describes what outcome would be observed at time t along a particular path of assignments.

Assumption 1 (Assignment and Potential Outcome)**.** *The assignment process $\{W_t\}_{t\geq 1}$ satisfies $W_t \in \mathcal{W} := \bigtimes_{k=1}^{d_w} \mathcal{W}_k \subseteq \mathcal{R}^{d_w}$. The potential outcome process is, for any deterministic sequence $\{w_s\}_{s\geq 1}$ with $w_s \in \mathcal{W}$ for all $s \geq 1$, $\{Y_t(\{w_s\}_{s\geq 1})\}_{t\geq 1}$, where the time-$t$ potential outcome satisfies $Y_t(\{w_s\}_{s\geq 1}) \in \mathcal{Y} \subseteq \mathbb{R}^{d_y}$.*

The simplest case is when the assignment is scalar and binary $\mathcal{W} = \{0,1\}$, in which case $W_t = 1$ corresponds to "treatment" and $W_t = 0$ is "control."

The potential outcome $Y_t(\{w_s\}_{s\geq 1})$ may depend on future assignments $\{w_s\}_{s\geq t+1}$. Our next assumption rules out this dependence, restricting the potential outcome to only depend on past and contemporaneous assignments.[1]

Assumption 2 (Non-anticipating Potential Outcomes)**.** *For each $t \geq 1$, and all deterministic $\{w_t\}_{t\geq 1}, \{w'_t\}_{t\geq 1}$ with $w_t, w'_t \in \mathcal{W}$,*

$$Y_t(w_{1:t}, \{w_s\}_{s\geq t+1}) = Y_t(w_{1:t}, \{w'_s\}_{s\geq t+1}) \text{ almost surely.}$$

Assumption 2 is a stochastic process analogue of non-interference (Cox, 1958; Rubin, 1980), extending White and Kennedy (2009) and Bojinov and Shephard (2019). It still allows for rich dependence on past and contemporaneous assignments. Under Assumption 2, we drop references to the future assignments in the potential outcome process, and write

$$\{Y_t(\{w_s\}_{s\geq 1})\}_{t\geq 1} = \{Y_t(w_{1:t})\}_{t\geq 1}.$$

The set $\{Y_t(w_{1:t})\colon w_{1:t} \in \mathcal{W}^t\}$ collects all the potential outcomes at time t.

Together, the assignments and potential outcome generate the output of the system.

[1] Such a restriction is one of the nine Bradford Hill (1965) criteria for causality ("temporality").

Assumption 3 (Output)**.** *The output is* $\{W_t, Y_t\}_{t\geq 1} = \{W_t, Y_t(W_{1:t})\}_{t\geq 1}$. *The* $\{Y_t\}_{t\geq 1}$ *is called the outcome process.*

The outcome process is the potential outcome process evaluated at the assignment process.

Finally, we assume that the assignment process is sequentially probabilistic, meaning that any assignment vector may be realized with positive probability at time t given the history of the observable stochastic processes up to time $t-1$. Let $\{\mathcal{F}_t\}_{t\geq 1}$ denote the natural filtration generated by (the σ-algebra of) the realized $\{w_t, y_t\}_{t\geq 1}$.

Assumption 4 (Sequentially probabilistic assignment process)**.** *The assignment process satisfies* $0 < P(W_t = w \mid \mathcal{F}_{t-1}) < 1$ *with probability one for all* $w \in \mathcal{W}$. *Here the probabilities are determined by a filtered probability space of* $\{W_t, \{Y_t(w_{1:t}), w_{1:t} \in \mathcal{W}^t\}\}_{t\geq 1}$.

This is the time series analogue of the "overlap" condition in cross-sectional causal studies. We make this assumption throughout the paper in order to focus attention on the causal interpretation of common time series estimands in the presence of rich dynamic causal effects. Understanding how violations of Assumption 4 affect the causal interpretation and estimation of common time series estimands is an important but separate issue.

By putting these assumptions together, we define a direct potential outcome system.

Definition 1 (Direct Potential Outcome System)**.** *Any* $\{W_t, \{Y_t(w_{1:t})\colon w_{1:t} \in \mathcal{W}^t\}\}_{t\geq 1}$ *satisfying Assumptions 1-4 is a* ***direct potential outcome system****.*

We refer to Definition 1 as a "direct" potential outcome system in order to emphasize that it focuses on nonparametrically modelling the direct causal effects of the assignment process $\{W_t\}$ on the outcomes $\{Y_t\}$. We do not, however, explicitly allow for the assignment W_t to have a causal effect on future assignments W_s for $s > t$. That is, we do not introduce a potential assignment $W_t(w_{1:t-1})$ which would model the assignment W_t that would be realized along the assignment path $w_{1:t-1} \in \mathcal{W}^{t-1}$ and would open an *indirect*, causal mechanism that allows the assignment W_t to indirectly affect future outcomes through its effect on future assignments.[2] The assignment process $\{W_t\}$ in the direct potential outcome system can still nonetheless have rich dependence. Assumption 1 places no restrictions on how W_t, W_s for $s \neq t$ are probabilistically related.

In focusing on the direct causal effects of assignments on outcomes, we adopt a common perspective in both macroeconometrics and financial econometrics. In particular, it is common in macroeconometrics to focus on studying the direct causal effects of underlying economic "shocks" on outcomes, which are thought to be underlying "random causes" the drive economic fluctuations and causally unrelated to one another (Frisch, 1933; Slutzky, 1937; Sims, 1980). In this view, the

[2]Such indirect causal mechanisms are often studied in a large biostatistics literature on longitudinal causal effects and dynamic treatment regimes – e.g., see Chapter 19 of Hernan and Robins (2021).

central goal is to trace out the dynamic causal effects of these primitive, economic shocks $\{W_t\}$ on macroeconomic outcomes $\{Y_t\}$. We refer the readers to Ramey (2016), Stock and Watson (2016), and Stock and Watson (2018) for recent discussions of this perspective in macroeconometrics. We further discuss the connections between the assignments in a direct potential outcome system and economic "shocks" in Section 3.

Remark 1 (Background processes)**.** *We could have further introduced the background process $\{X_t\}_{t\geq 1}$ that is causally unaffected by the assignment process. Such a process would play the same role as pre-assignment covariates in cross-sectional or longitudinal studies.*

2.2 Dynamic Causal Effects

Any comparison of the potential outcome process at a particular point in time along different possible realizations of the assignment process define a *dynamic causal effect*. The dynamic causal effect at time t for assignment path $w_{1:t} \in \mathcal{W}^t$ and counterfactual path $w'_{1:t} \in \mathcal{W}^t$ is $Y_t(w_{1:t}) - Y_t(w'_{1:t})$. Of course, this is an enormous class of dynamic causal effects as there are exponentially many possible paths $w_{1:t} \in \mathcal{W}^t$. We therefore introduce causal estimands that average over these dynamic causal effects along various underlying assignment paths.

To do so, let us introduce some shorthand. For $t \geq 1$, $h \geq 0$, and any fixed $w \in \mathcal{W}$, write the time-$(t+h)$ potential outcome at the assignment process $(W_{1:t-1}, w, W_{t+1:t+h})$ as

$$Y_{t+h}(w) := Y_{t+h}(W_{1:t-1}, w, W_{t+1:t+h}).$$

Notice that $Y_{t+h} = Y_{t+h}(W_t)$ by definition.

Definition 2 (Dynamic causal effects)**.** *For $t \geq 1$, $h \geq 0$, and any fixed $w, w' \in \mathcal{W}$, the time-t, h-period ahead* ***impulse causal effect***, ***filtered treatment effect***, *and* ***average treatment effect*** *are, respectively:*

$$Y_{t+h}(w) - Y_{t+h}(w'), \quad \mathbb{E}[Y_{t+h}(w) - Y_{t+h}(w') \mid \mathcal{F}_{t-1}], \quad \mathbb{E}[Y_{t+h}(w) - Y_{t+h}(w')].$$

The impulse causal effect measures the *ceteris paribus* causal effect of intervening to switch the time-t assignment from w' to w on the h-period ahead outcomes holding all else fixed along the assignment process. The impulse causal effect is a random object since the potential outcome process itself is stochastic as well as the past $W_{1:t-1}$ and future $W_{t+1:t+h}$ assignments are stochastic.

The *filtered treatment effect* averages the impulse causal effect conditional on the natural filtration of assignments and observed outcomes up to time $t-1$. We use the nomenclature "filtered" following the stochastic process literature, where filtering refers to the sequential estimation of time-varying unobserved variables, e.g. Kalman filter (Kalman, 1960; Durbin and Koopman, 2012), particle filter (Gordon et al., 1993; Pitt and Shephard, 1999; Chopin and Papasphiliopoulos, 2020), and hidden discrete Markov models (Baum and Petrie, 1966; Hamilton, 1989). This labelling fits as the potential outcomes are unobserved.[3]

Finally, the *average treatment effect* further averages the filtered treatment effect over the filtration, yielding the unconditional expectation of the impulse causal effect $Y_{t+h}(w) - Y_{t+h}(w')$.

Remark 2. *If new outcome variables were added to an existing causal study, the impulse causal effect and the average treatment effect for the existing variables would not be changed, but the filtered treatment effect might as the new outcome variables would bulk up the filtration and so possibly change the conditional expectation.*

We further define analogous versions of the dynamic causal effects for a particular scalar assignment. For any fixed $w_k \in \mathcal{W}_k$, define

$$Y_{t+h}(w_k) := Y_{t+h}(W_{1:t-1}, W_{1:k-1,t}, w_k, W_{k+1:d_W,t}, W_{t+1:t+h}).$$

The corresponding time-t, h-period ahead impulse causal effect, filtered treatment effect, and average treatment effect for the k-th assignment are, respectively:

$$Y_{t+h}(w_k) - Y_{t+h}(w_k'), \quad \mathbb{E}[\{Y_{t+h}(w_k) - Y_{t+h}(w_k')\} \mid \mathcal{F}_{t-1}], \quad \mathbb{E}[Y_{t+h}(w_k) - Y_{t+h}(w_k')].$$

The dynamic causal effects in Definition 2 summarize the causal effect of discrete interventions to switch the time-t assignments on the outcomes. We finally introduce derivatives that summarize marginal causal effects of incrementally varying the time-t assignment (see, for example, Angrist and Imbens (1995) and Angrist et al. (2000) for analogous definitions in cross-sectional settings).

Definition 3. *If they individually exist,*

$$Y_{t+h}'(w_k) = \frac{\partial Y_{t+h}(w_k)}{\partial w_k}, \quad \mathbb{E}[Y_{t+h}'(w_k) \mid \mathcal{F}_{t-1}], \quad \mathbb{E}[Y_{t+h}'(w_k)],$$

*are called the time-t, h-period ahead **marginal impulse causal effect**, the **marginal filtered treatment effect**, and the **marginal average treatment effect** respectively.*

2.3 Links to macroeconometrics

Before continuing, we highlight how the direct potential outcome system naturally links to several recent developments and debates in macroeconometrics and encompasses many familiar parametric models in that field. We start with the former.

[3] We note that Lee and Salanie (2020) also use the phrase "filtered treatment effect" in analyzing a cross-sectional setting with partially observed assignments.

First, the direct potential outcome system provides a unifying framework to analyze what assumptions must be placed on the assignment process to endow causal meaning to common statistical estimands without resorting to functional form assumptions. Workhorse models in macroeconometrics, such as the structural vector moving average, assume linearity. However, this nullifies state-dependence and asymmetry in dynamic causal effects. Researchers recognize the restrictiveness of linearity, yet attempt to weaken it on a case-by-case basis. For example, on the possible nonlinear effects of oil prices (Killian and Vigfusson, 2011b,a; Hamilton, 2011); on the nonlinear and state dependent effects of monetary policy (Tenreyro and Thwaites, 2016; Jordá et al., 2020; Aruoba et al., 2021; Mavroeidis, 2021), and on state-dependent fiscal multipliers (Auerbach and Gorodnichenko, 2012b,a; Ramey and Zubairy, 2018; Cloyne et al., 2020). Similarly, the direct potential outcome system does not rely on "invertibility" or "recoverability" assumptions about the assignment and potential outcome processes (Chahrour and Jurado, 2021). Understanding what can be identified about dynamic causal effects without relying on these assumptions is an active area (Stock and Watson, 2018; Plagborg-Møller, 2019; Plagborg-Møller and Wolf, 2020; Chahrour and Jurado, 2021).

Second, a rapidly growing body of empirical research in macroeconometrics attempts to estimate dynamic causal efects in settings where researchers directly observe both the assignments and outcomes, $\{w_t^{obs}, y_t^{obs}\}_{t\geq 1}$. In this line of work, empirical researchers creatively construct measures of the underlying economic shocks of interest W_t, and then use these constructed shocks to directly estimate dynamic causal effects on macroeconomic outcomes using reduced-form methods such as local projections (Jordá, 2005) or autoregressive distributed lag models (Baek and Lee, 2021). This line of work has recently been called "direct causal inference" by Nakamura and Steinsson (2018b) in order to contrast it with the dominant model-based approach to causal inference in macroeconomics in the tradition of Sims (1980). We refer the reader to Nakamura and Steinsson (2018b), Goncalves et al. (2021), and Baek and Lee (2021) for recent discussions of this growing empirical literature in macroeconomics. The direct potential outcome system provides a causal foundation for such reduced-form methods in time series, elucidating the assumptions that the constructed shock must satisfy in order for the reduced form estimands to have a nonparametric causal interpretation.

2.3.1 Examples from Macroeconomics

Many leading causal models in macroeconomics can be cast as special cases of the direct potential outcome system that place additional restrictions on the potential outcome process.

Example 1 (Structural vector moving average (SVMA) model)**.** The SVMA model is the leading workhorse model for studying dynamic causal effects in macroeconometrics (e.g., Kilian and

Lutkepohl, 2017; Stock and Watson, 2018). Any infinite-order SVMA model can be expressed as a direct potential outcome system by assuming that the potential outcome process satisfies the functional form restriction

$$Y_t(w_{1:t}) := \sum_{l=0}^{t-1} \Theta_l w_{t-l} + Y_t^*,$$

where $\{W_t\}_{t\geq 1}$ is the assignment process, $\{\Theta_l\}_{0\leq l<t}$ is a sequence of lag-coefficient matrices, and $\{Y_t^*\}_{t\geq 1}$ is a stochastic process that is causally unaffected by the assignment process. In this sense, the SVMA model imposes that the potential outcome process is linear in the assignment process. This mapping requires no assumptions on the dimensionality of the assignment process d_w, the dimensionality of the potential outcome process d_y, nor the lag-coefficient matrices. As discussed in Plagborg-Møller and Wolf (2020), such an infinite-order SVMA model is consistent with all discrete-time Dynamic Stochastic General Equilibrium models as well as all stable, linear structural vector autoregression (SVAR) models. We further discuss the SVMA model in Section 7. ▲

Example 2 (Nonlinear structural vector autoregressions (SVAR))**.** Recent advances in nonlinear SVARs can also be cast as special cases of the direct potential outcome system. As an illustration, consider the motivating example in Goncalves et al. (2021), which is a non-linear SVAR of the form:

$$Y_{1,t}(w_{1:t}) = w_{1,t}, \quad Y_{2,t}(w_{1:t}) = b + \beta Y_{1,t}(w_{1:t}) + \rho Y_{2,t-1}(w_{1:t-1}) + cf(Y_{1,t}(w_{1:t})) + w_{2,t},$$

where f is a nonlinear function. Given a stochastic initial condition $Y_{2,0} := \epsilon_{2,0}$ that is causally unaffected by the assignment process, iterating this system of equations forward arrives at a potential outcome process $Y_{1,t}(w_{1:t}) = w_{1,t}$, and $Y_{2,t}(w_{1:t}) = g_{2,t}(w_{1:t}, \epsilon_{2,0}; \theta)$, where $g_{2,t}$ is a known function and $\theta := (b, c, \beta, \rho)$ are the parameters. This is a direct potential outcome system where (1) $Y_{1,t}(w_{1:t})$ is non-random and only depends on the contemporaneous assignment, (2) the randomness in $Y_{2:t}(w_{1:t})$ is driven by the initial condition. Other recent examples of nonlinear SVARs include Aruoba et al. (2021) and Mavroeidis (2021). ▲

Example 3 (Potential outcome model in Angrist and Kuersteiner (2011), Angrist et al. (2018))**.** Angrist and Kuersteiner (2011) and Angrist et al. (2018) introduce a potential outcome model for time series settings that is a special case of the direct potential outcome system. Using our notation, Angrist and Kuersteiner (2011) introduce a system of structural equations in which for $t \geq 1$,

$$Y_{1,t}(w_{1:t}) = f_{1,t}(Y_{1:t-1}(w_{1:t-1}), w_{1,t}; \epsilon_0), \quad Y_{2,t}(w_{1:t}) = f_{2,t}(Y_{1,t}(w_{1:t}), w_{2,t}, w_{1:t-1}; \epsilon_0),$$

where $f_{1,t}$, $f_{2,t}$ are deterministic functions and ϵ_0 is a random initial condition. These structural equations impose that $w_{1:t}$ only impacts $Y_{1,t}$ through $w_{1,t}$ directly and through $Y_{1,t-1}$ indirectly. Further, $w_{2,1:t}$ only impacts $Y_{2,t}$ contemporaneously. Related thinking includes White and Kennedy (2009) and White and Lu (2010). Through forward iteration of the system starting at $t = 1$, this can also be expressed as a direct potential outcome system. In this system of structural equations, the authors defined the collection of their time-$t{+}h$ potential outcomes, as $\{Y_{t+h}(w^{obs}_{1:t-1}, w, W_{t+1:t+h}) : w \in \mathcal{W}_W\}$ and focused on $\mathbb{E}[Y_{t+h}(w^{obs}_{1:t-1}, w, W_{t+1:t+h}) - Y_{t+h}(w^{obs}_{1:t-1}, w', W_{t+1:t+h})]$, which they called the "average policy effect." ▲

Example 4 (Expectations)**.** Macroeconomists often consider how assignments are influenced by the distribution of future outcomes and how they in turn vary with assignments. For example, consumers and firms are modelled as forward-looking and so, expectations about future outcomes influence behavior today. Consider a simple optimization-based version (e.g., Lucas, 1972; Sargent, 1981) in which the assignment process is given by

$$W_t \in \arg\max_{w_t} \max_{w_{t+1:T}} \mathbb{E}[\, U(Y_{t:T}(w^{obs}_{1:t-1}, w_{t:T}), w_{t:T}) \,|\, y^{obs}_{1:t-1}, w^{obs}_{1:t-1}], \tag{1}$$

where U is a utility function of future outcomes and assignments, while $\mathcal{F}_{t-1}$ is written out in long hand as $y^{obs}_{1:t-1}, w^{obs}_{1:t-1}$. For each possible $w_{t:T} \in \mathcal{W}^{T-t+1}$, the expectation is over the law of $Y_{t:T}(w^{obs}_{1:t-1}, w_{t:T})|y^{obs}_{1:t-1}, w^{obs}_{1:t-1}$. This decision rule delivers the output $\{W_t, Y_t(W_{1:t})\}_{t\geq 1}$. This looks like a direct potential outcome system since Assumption 2 holds. The assignment W_t could be a deterministic function of past data if the optimal choice is unique, which would violate Assumption 4. However, incorporating noise in the decision rule (1) would deliver a direct potential outcome system. ▲

3 Estimands Based on Assignments and Outcomes

In this section, we establish nonparametric conditions under which common statistical estimands based on assignments and outcomes have causal meaning in the direct potential outcome system $\{W_t, \{Y_t(w_{1:t}) : w_{1:t} \in \mathcal{W}^t\}\}_{t\geq 1}$, where researchers observe the realized assignments and realized outcomes $\{w^{obs}_t, y^{obs}_t\}_{t\geq 1}$. We ask if the following statistical estimands have causal meaning: impulse response function, local projection, generalized impulse response function and the local filtered projection. Table 1 defines these estimands and summarizes our main results on their causal interpretation under important restrictions on the assignment process and other technical conditions. The rest of this Section spells out the details.

In this section, there is no loss in generality in assuming the outcome Y_{t+h} is univariate. The more general case is covered by running the analysis equation by equation.

Name	Estimand	Causal Interpretation
Impulse Response Function	$\mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k]$ $-\mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k']$	$\mathbb{E}[Y_{t+h}(w_k) - Y_{t+h}(w_k')]$;
Local Projection	$\frac{Cov(Y_{t+h}, W_{k,t})}{Var(W_{k,t})}$	$\frac{\int_{\mathcal{W}_k} \mathbb{E}[Y'_{t+h}(w_k)]\mathbb{E}[G_t(w_k)]dw_k}{\int_{\mathcal{W}_k} \mathbb{E}[G_t(w_k)]dw_k}$
Generalized Impulse Response Function	$\mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k, \mathcal{F}_{t-1}]$ $-\mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k', \mathcal{F}_{t-1}]$	$\mathbb{E}[Y_{t+h}(w_k) - Y_{t+h}(w_k') \mid \mathcal{F}_{t-1}]$;
Local Filtered Projection	$\frac{\mathbb{E}[\{Y_{t+h} - \hat{Y}_{t+h\mid t-1}\}\{W_{k,t} - \hat{W}_{k,t\mid t-1}\}]}{\mathbb{E}[\{W_{k,t} - \hat{W}_{k,t\mid t-1}\}^2]}$	$\frac{\int_{\mathcal{W}_k} \mathbb{E}\left[\mathbb{E}[Y'_{t+h}(w_k)\mid\mathcal{F}_{t-1}]\mathbb{E}[G_{t\mid t-1}(w_k)\mid\mathcal{F}_{t-1}]\right]dw_k}{\int_{\mathcal{W}_k} \mathbb{E}[G_{t\mid t-1}(w_k)]dw_k}$

Table 1: Top line results for the causal interpretation of common estimands based on assignments and outcomes. Here $h \geq 0$, $w_k, w_k' \in \mathcal{W}_k$, $G_t(w_k) = 1\{w_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t}])$ and $G_{t|t-1}(w_k) = 1\{w_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}])$, while $\hat{Y}_{t+h|t-1} := \mathbb{E}[Y_{t+h} \mid \mathcal{F}_{t-1}]$ and $\hat{W}_{k,t|t-1} := \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}]$. Note that $\mathbb{E}[G_t(w_k)] \geq 0$ and $\mathbb{E}[G_{t|t-1}(w_k) \mid \mathcal{F}_{t-1}] \geq 0$.

3.1 Impulse Response Function

We begin by determining the conditions under which the unconditional *impulse response function* (Sims, 1980) is the h-period ahead average treatment effect. For $h \geq 0$ and deterministic $w_k, w_k' \in \mathcal{W}_k$, the impulse response function is defined by, if it exists,

$$IRF_{k,t,h}(w_k, w_k') := \mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k] - \mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k']. \tag{2}$$

$IRF_{k,t,h}(w_k, w_k')$ can be decomposed into the average treatment effect and a selection bias term.

Theorem 1. *Assume a direct potential outcome system, consider some $k = 1, \ldots, d_w$, $t \geq 1$, $h \geq 0$, fix $w_k, w_k' \in \mathcal{W}_k$ and that $\mathbb{E}[|Y_{t+h}(w_k) - Y_{t+h}(w_k')|] < \infty$. Then,*

$$IRF_{k,t,h}(w_k, w_k') = \mathbb{E}[Y_{t+h}(w_k) - Y_{t+h}(w_k')] + \Delta_{k,t,h}(w_k, w_k'),$$

where

$$\Delta_{k,t,h}(w_k, w_k') := \frac{Cov\left(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\}\right)}{\mathbb{E}[1\{W_{k,t} = w_k\}]} - \frac{Cov\left(Y_{t+h}(w_k'), 1\{W_{k,t} = w_k'\}\right)}{\mathbb{E}[1\{W_{k,t} = w_k'\}]}.$$

The impulse response function is therefore equal to the average treatment effect if and only if the selection bias term $\Delta_{k,t,h}(w_k, w_k') = 0$. A sufficient condition for this to hold is that the two covariance terms are zero.

Notice that these covariance terms depend on how the assignment $W_{k,t}$ covaries with the potential outcome $Y_{t+h}(w_k)$. Since $Y_{t+h}(w_k) := Y_{t+h}(W_{1:t-1}, W_{1:k-1,t}, w_k, W_{k+1:d_W,t}, W_{t+1:t+h})$ by definition, the selection bias therefore depends on how the assignment $W_{k,t}$ relates to

1. past assignments $W_{1:t-1}$,
2. other contemporaneous assignments $W_{1:k-1,t}, W_{k+1:d_W,t}$,
3. future assignments $W_{t+1:t+h}$, and
4. the potential outcome process $Y_{t+h}(w_{1:t+h})$.

By placing further restrictions on the assignment process, we immediately arrive at sufficient conditions for $\Delta_{k,t,h}(w_k, w_k')$ to be zero.

Theorem 2. *Under the same conditions as Theorem 1, if*

$$Cov\left(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\}\right) = 0, \quad Cov\left(Y_{t+h}(w_k'), 1\{W_{k,t} = w_k'\}\right) = 0 \tag{3}$$

then $\Delta_{k,t,h}(w_k, w_k') = 0$*. Moreover, (3) is satisfied if*

$$W_{k,t} \perp\!\!\!\perp Y_{t+h}(w_k), \quad \textit{and} \quad W_{k,t} \perp\!\!\!\perp Y_{t+h}(w_k'), \tag{4}$$

which is in turn implied by

$$W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k) \colon w_k \in \mathcal{W}_k\}, \tag{5}$$

which is in turn implied by

$$W_{k,t} \perp\!\!\!\perp \left(W_{1:t-1}, W_{1:k-1,t}, W_{k+1:d_W,t}, W_{t+1:t+h}, \{Y_{t+h}(w_{1:t+h}) \colon w_{1:t+h} \in \mathcal{W}^{t+h}\}\right). \tag{6}$$

Equation (6) says the selection bias is zero if the assignment $W_{k,t}$ is randomized in the sense that it is independent of all other assignments and the time-$(t+h)$ potential outcomes.

Recent reviews on dynamic causal effects in macroeconometrics by Ramey (2016) and Stock and Watson (2018) argue intuitively that the impulse response function of observed outcomes to "shocks" in parametric structural models, such as the SVMA, are analogous to an average treatment effect in a randomized experiment from cross-sectional causal inference.[4] However, these statements rely on either intuitive descriptions of the statistical properties of shocks[5], or on a specific parametric model for the potential outcome process to link the impulse response function to

[4] Stock and Watson (2018) write on pg. 922: "The macroeconometric jargon for this random treatment is a 'structural shock:' a primitive, unanticipated economic force, or driving impulse, that is unforecastable and uncorrelated with other shocks. The macroeconomist's shock is the microeconomists' random treatment, and impulse response functions are the causal effects of those treatments on variables of interest over time, that is, dynamic causal effects."

[5] Ramey (2016) writes on pg. 75, "the shocks should have the following characteristics: (1) they should be exogenous with respect to the other current and lagged endogenous variables in the model; (2) they should be uncorrelated with other exogenous shocks; otherwise, we cannot identify the unique causal effects of one exogenous shock relative to another; and (3) they should represent either unanticipated movements in exogenous variables or news about future movements in exogenous variables."

an average dynamic causal effect. Theorem 2 clarifies that if the assignment $W_{k,t}$ is randomly assigned in these sense of (6), then the impulse response function nonparametrically identifies an average treatment effect in the direct potential outcome system. In this sense, Theorem 2 provides an interpretation of "shock" in terms of a random assignment assumption on the assignment process in a direct potential outcome system.

Furthermore, Theorems 1-2 clarifies a recent empirical literature that seeks to directly construct measures of the shocks of interest and measure dynamic causal effects through reduced-form estimates of impulse response functions — so called "direct causal inference" (e.g., see Nakamura and Steinsson, 2018b; Baek and Lee, 2021). In order for researchers to causally interpret reduced-form impulse response functions of outcomes on particular constructed shocks as nonparametrically identifying an average treatment effect, then the constructed shocks must be randomized in these sense given in Theorem 2.

3.2 Local Projection Estimand

Under the conditions of Theorem 1, impulse response functions are causal, but nonparametrically estimating impulse response functions is in general challenging. If the assignment is observed by the researcher, it is therefore common to estimate impulse response functions using "local projections" (Jordá, 2005), which directly regresses the h-step ahead outcome on a constant and the assignment. The corresponding local projection estimand is

$$LP_{k,t,h} := \frac{Cov(Y_{t+h}, W_{k,t})}{Var(W_{k,t})}. \tag{7}$$

Theorem 3 establishes that $LP_{k,t,h}$ identifies a weighted average of marginal causal effects of the assignment on the h-step ahead outcome.

Theorem 3. *Under the same conditions as Theorem 1, further assume that:*

i. The support of $W_{k,t}$ is a closed interval, $\mathcal{W}_k := [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}$.

ii. Differentiability: $Y_{t+h}(w_k)$ is continuously differentiable in w_k, as is $\mathbb{E}[Y'_{t+h}(w_k)]$.

iii. Independence: $W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k) \colon w_k \in \mathcal{W}_k\}$.

Then, if it exists,

$$LP_{k,t,h} = \frac{\int_{\mathcal{W}_k} \mathbb{E}[Y'_{t+h}(w_k)]\mathbb{E}[G_t(w_k)]dw_k}{\int_{\mathcal{W}_k} \mathbb{E}[G_t(w_k)]dw_k},$$

where $G_t(w_k) = 1\{w_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t}])$, noting $\mathbb{E}[G_t(w_k)] \geq 0$;

The local projection estimand $LP_{k,t,h}$ is therefore a weighted average of marginal average treatment effects of $W_{k,t}$ on the Y_{t+h}, where the weights $\mathbb{E}[G_t(w_k)]$ are non-negative and sum to one. Thus, if the assignment $W_{k,t}$ is a shock in the sense stated in Theorem 2, the local projection estimand also has a nonparametric causal interpretation.

3.3 Generalized Impulse Response Function

In non-linear time series models, it is common to focus on the conditional version of the impulse response function, the h-period ahead *generalized impulse response function* (Gallant et al., 1993; Koop et al., 1996; Gourieroux and Jasiak, 2005), which is

$$GIRF_{k,t,h}(w_k, w_k' \mid \mathcal{F}_{t-1}) := \mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k, \mathcal{F}_{t-1}] - \mathbb{E}[Y_{t+h} \mid W_{k,t} = w_k', \mathcal{F}_{t-1}]. \quad (8)$$

Mirroring our analysis of the impulse response function, we next show that $GIRF_{k,t,h}$ can be decomposed into the filtered treatment effect and a selection bias term.

Theorem 4. *Assume a direct potential outcome system, some $k = 1, \ldots, d_w$, $t \geq 1$, and $h \geq 0$ and that $\mathbb{E}[|Y_{t+h}(w_k) - Y_{t+h}(w_k')| \mid \mathcal{F}_{t-1}] < \infty$. Then, for any deterministic $w_k, w_k' \in \mathcal{W}$,*

$$GIRF_{k,t,h}(w_k, w_k' \mid \mathcal{F}_{t-1}) = \mathbb{E}[\{Y_{t+h}(w_k) - Y_{t+h}(w_k')\} \mid \mathcal{F}_{t-1}] + \Delta_{k,t,h}(w_k, w_k' \mid \mathcal{F}_{t-1}),$$

where

$$\Delta_{k,t,h}(w_k, w_k' \mid \mathcal{F}_{t-1}) := \frac{Cov\left(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}\right)}{\mathbb{E}[1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}]} - \frac{Cov\left(Y_{t+h}(w_k'), 1\{W_{k,t} = w_k'\} \mid \mathcal{F}_{t-1}\right)}{\mathbb{E}[1\{W_{k,t} = w_k'\} \mid \mathcal{F}_{t-1}]}.$$

Sufficient conditions for the selection bias term $\Delta_{k,t,h}(w_k, w_k' \mid \mathcal{F}_{t-1})$ to equal zero is that the two conditional covariances are zero. Repeating the unconditional case, Theorem 5 provides sufficient conditions such that the selection bias term is equal to zero.

Theorem 5. *Under the same conditions as Theorem 4, if*

$$Cov\left(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}\right) = 0, \quad Cov\left(Y_{t+h}(w_k'), 1\{W_{k,t} = w_k'\} \mid \mathcal{F}_{t-1}\right) = 0, \quad (9)$$

then $\Delta_{k,t,h}(w_k, w_k') = 0$. Moreover, (9) is implied by

$$W_{k,t} \perp\!\!\!\perp Y_{t+h}(w_k) \mid \mathcal{F}_{t-1}, \quad \textit{and} \quad W_{k,t} \perp\!\!\!\perp Y_{t+h}(w_k') \mid \mathcal{F}_{t-1}, \quad (10)$$

which is in turn implied by

$$[W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k) \colon w_k \in \mathcal{W}_k\}] \mid \mathcal{F}_{t-1}, \quad (11)$$

which is in turn implied by

$$[W_{k,t} \perp\!\!\!\perp \left(W_{1:k-1,t}, W_{k+1:d_W,t}, W_{t+1:t+h}, \{Y_{t+h}(w^{obs}_{1:t-1}, w_{t:t+h})\colon w_{t:t+h} \in \mathcal{W}^{h+1}\}\right)] \mid \mathcal{F}_{t-1}. \quad (12)$$

Therefore, under (9), the selection bias $\Delta_{k,t,h}(w_k, w'_k \mid \mathcal{F}_{t-1}) = 0$ and the generalized impulse response function identifies the filtered impulse causal effect. Notice how much weaker (12) is than (6) as it allows the assignment to probabilistically depend flexibly on the past realised potential outcomes and realised assignments.

At first glance, (11) appears analogous to a typical unconfoundedness assumption from cross-sectional causal inference or sequential randomization assumption from longitudinal causal inference. That is, it imposes that conditional on the history up to time $t-1$, the assignment $W_{k,t}$ must be as good as randomly assigned. However, recall that the notation $Y_{t+h}(w_k)$ buries dependence on (i) other contemporaneous assignments $W_{1:k-1,t}, W_{k+1:d_W,t}$; (ii) future assignments $W_{t+1:t+h}$; and (iii) the potential outcomes at time-$(t+h)$. Therefore, (12) in Theorem 5 provides further sufficient conditions under which (11) is satisfied, highlighting that it is sufficient to further impose that the assignment $W_{k,t}$ is jointly independent of all other contemporaneous and future assignments as well as the underlying potential outcomes.

Remark 3. *How do the conditions in Theorem 2 relate to the conditions in Theorem 5? Applying the law of total covariance yields*

$$\begin{aligned} Cov(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\}) &= \mathbb{E}[Cov(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1})] \\ &\quad + Cov(\mathbb{E}[Y_{t+h}(w_k) \mid \mathcal{F}_{t-1}], \mathbb{E}[1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}]), \end{aligned}$$

so $Cov(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\}) = 0$ neither implies or is implied by $Cov(Y_{t+h}(w_k), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1})] = 0$. Hence, the conditional and unconditional cases are non-nested. If we instead work probabilistically, then the condition

$$W_{k,t} \perp\!\!\!\perp \left(W_{1:t-1}, W_{1:k-1,t}, W_{k+1:d_W,t}, W_{t+1:t+h}, \{Y_{1:t+h}(w_{1:t+h})\colon w_{1:t+h} \in \mathcal{W}^{t+h}\}\right),$$

which strengthens (6) to additionally require independence of the full potential outcome process, implies the condition (12). This second point is important practically. The generalized impulse response function tells us the filtered treatment effect provided that $[W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k)\colon w_k \in \mathcal{W}_k\}] \mid \mathcal{F}_{t-1}$. A temporally averaged generalized impulse response function therefore tells us the average treatment effect without the need to employ the harsher condition $[W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k)\colon w_k \in \mathcal{W}_k\}]$ as it sidesteps the use of the impulse response function.

3.4 Generalized Local Projection and Local Filtered Projection Estimands

Again estimating generalized impulse response functions nonparametrically is challenging. Under the same conditions as Theorem 3 but replacing condition (iii) with Equation (11), the *generalized local projection* satisfies

$$\frac{Cov(Y_{t+h}, W_{k,t} \mid \mathcal{F}_{t-1})}{Var(W_{k,t} \mid \mathcal{F}_{t-1})} = \frac{\int_{\mathcal{W}_k} \mathbb{E}[Y'_{t+h}(w_k) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(w_k) \mid \mathcal{F}_{t-1}]dw_k}{\int_{\mathcal{W}_k} \mathbb{E}[G_{t|t-1}(w_k) \mid \mathcal{F}_{t-1}]dw_k},$$

where $G_{t|t-1}(w_k) = 1\{w_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}])$, noting $\mathbb{E}[G_{t|t-1}(w_k) \mid \mathcal{F}_{t-1}] \geq 0$. The generalized local projection is equivalent to a weighted average of conditional average marginal effects of $W_{k,t}$ on Y_{t+h}, where the weights now depend on the natural filtration but still are non-negative and sum to one.

Of more practical importance, is the local projection of $Y_{t+h} - \hat{Y}_{t+h|t-1}$ on $W_{k,t} - \hat{W}_{k,t|t-1}$, where $\hat{Y}_{t+h|t-1} := \mathbb{E}[Y_{t+h} \mid \mathcal{F}_{t-1}]$ and $\hat{W}_{k,t|t-1} := \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}]$. We call the associated estimand the *local filtered projection*, which is defined as

$$\frac{\mathbb{E}[\{Y_{t+h} - \hat{Y}_{t+h|t-1}\}\{W_{k,t} - \hat{W}_{k,t|t-1}\}]}{\mathbb{E}[\{W_{k,t} - \hat{W}_{k,t|t-1}\}^2]}.$$

Under the same conditions as needed for the generalized local projection plus needing the unconditional expectations to exist, the local filtered projection estimand equals

$$\frac{\int_{\mathcal{W}_k} \mathbb{E}\left[\mathbb{E}[Y'_{t+h}(w_k) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(w_k) \mid \mathcal{F}_{t-1}]\right] dw_k}{\int_{\mathcal{W}_k} \mathbb{E}[G_{t|t-1}(w_k)]dw_k},$$

This is a long-run weighted average of the marginal filtered causal effect. The weights are non-zero and average to one over time.

4 The Instrumented Potential Outcome System

We now use a special case of the direct potential outcome system to incorporate instrumental variables for the assignment process. This is useful as a rapidly growing literature in macroeconomics exploits the use of instruments to identify dynamic causal effects (e.g., see Jordá et al., 2015; Gertler and Karadi, 2015; Ramey and Zubairy, 2018; Stock and Watson, 2018; Plagborg-Møller and Wolf, 2020; Jordá et al., 2020, among many others). Section 5 details the case where the researcher observes the assignments, the instruments, and the outcomes. Section 6 considers the case where only the instruments and the outcomes are observed.

4.1 The Instrumented System

We start by setting up an "augmented assignment" V_t, so that $\{V_t, \{Y_t(v_{1:t}) : v_{1:t} \in \mathcal{W}_V^t\}\}_{t\geq 1}$ is a direct potential outcome system.

The instrumented potential outcome system then imposes two further assumptions on the potential outcome system: (i) that $\{V_t\}_{t\geq 1}$ splits into an "instrument" $\{Z_t\}_{t\geq 1}$ and a "potential assignment" $\{W_t(z_t) : z_t \in \mathcal{W}_Z^t\}_{t\geq 1}$ that is only causally affected by the contemporaneous instrument, meaning $V_t = (Z_t, W_t(Z_t))$; (ii) the potential outcome process is only affected by the assignment $W_{1:t}$.

Definition 4 (Instrumented potential outcome system)**.** *Assume $W_t \in \mathcal{W}_W$, $Z_t \in \mathcal{W}_Z$ and write $V_t = (W_t, Z_t)$. Assume $\{V_t, \{Y_t(v_{1:t}) : v_{1:t} \in \mathcal{W}_W^t \times \mathcal{W}_Z^t\}\}_{t\geq 1}$ is a direct potential outcome system. Additionally, enforce three Assumptions:*

i. ***Contemporaneous Instrument****: The "potential assignments" satisfy*

$$W_{k,t}(\{z_s\}_{s\geq 1}) = W_{k,t}(z'_{1:t-1}, z_t, \{z'_s\}_{s\geq t+1})$$
$$W_{1:k-1,t}(\{z_s\}_{s\geq 1}) = W_{1:k-1,t}(\{z'_s\}_{s\geq 1})$$
$$W_{k+1:d_W,t}(\{z_s\}_{s\geq 1}) = W_{k+1:d_W,t}(\{z'_s\}_{s\geq 1})$$

almost surely, for all $t \geq 1$ and all deterministic $\{z_t\}_{t\geq 1}$ and $\{z'_t\}_{t\geq 1}$. Write the potential assignments as $\{W_t(z_t) = (W_{1:k-1,t}, W_{k,t}(z_t), W_{k+1:d_W,t}) : z_t \in \mathcal{W}_Z\}$, while the assignment is $W_t = W_t(Z_t) = (W_{1:k-1,t}, W_{k,t}(Z_t), W_{k+1:d_W,t})$.

ii. ***Potential Outcome Exclusion****:*

$$Y_t((w_1, z_1), ..., (w_t, z_t)) = Y_t((w_1, z'_1), ..., (w_t, z'_t))$$

almost surely for all $w_{1:t} \in \mathcal{W}_W^t$ and $z_{1:t}, z'_{1:t} \in \mathcal{W}_Z^t$. Write the potential outcomes as $\{Y_t(w_{1:t}) : w_{1:t} \in \mathcal{W}_W^t\}$ and outcome as $Y_t = Y_t(W_{1:t})$.

iii. ***Output****: The output is*

$$\{Z_t, W_t, Y_t\}_{t\geq 1} = \{Z_t, W_t(Z_t), Y_t(W_{1:t})\}_{t\geq 1},$$

while Z_t and $\{Z_t\}_{t\geq 1}$ are called the "contemporaneous instrument," and instrument process, respectively.

Any $\{Z_t, \{W_t(z_t), z_t \in \mathcal{W}_Z\}, \{Y_t(w_{1:t}), w_{1:t} \in \mathcal{W}_W^t\}\}_{t\geq 1}$ satisfying (i)-(iii) is an **instrumented potential outcome system***.*

The simplest case is when both the assignment and instrument are scalar and binary, $\mathcal{W}_W = \{0,1\}, \mathcal{W}_Z = \{0,1\}$. In this case, the instrument $Z_t = 1$ corresponds to "intention to treat" and $Z_t = 0$ is "intention to control." There is treatment and control as intended when $W_t(1) = 1$ and $W_t(0) = 0$. But there can be noncompliance when $W_t(1) = 0$ and $W_t(0) = 1$.

Assumption (i) imposes that Z_t is only an instrument for the time-t, k-th assignment. This formalizes common empirical intuition in macroeconometrics where a constructed external instrument is often "targeted" towards a single economic shock of interest – for example, empirical researchers construct proxies for a monetary policy shock (e.g., Gertler and Karadi, 2015; Nakamura and Steinsson, 2018a; Jordá et al., 2020) or a fiscal policy shock (Ramey and Zubairy, 2018). Assumption (ii) is the familiar outcome exclusion restriction on the instrument from cross-sectional causal inference.

To use this structure, we also need a type of "relevance" condition on the instrument. Such conditions will be stated as needed below.

5 Estimands Based on Assignments, Instruments and Outcomes

We now study the conditions under which leading statistical estimands based on assignments, instruments and outcomes have causal meaning in the context of an instrumented potential outcome system $\{Z_t, \{W_t(z_t), z_t \in \mathcal{W}_Z\}, \{Y_t(w_{1:t}), w_{1:t} \in \mathcal{W}_W^t\}\}_{t\geq 1}$. We consider the case in which researcher observes the instruments, the assignments, and the outcomes $\{z_t^{obs}, w_t^{obs}, y_t^{obs}\}_{t\geq 1}$.

Since the assignments themselves are assumed to be directly observable, we focus on dynamic IV estimands that involve taking the ratio of an impulse response function of the outcome on the instrument relative to the impulse response function of the assignment on the instrument. We show that such dynamic IV estimands identify local average impulse causal effects in the sense of Imbens and Angrist (1994), Angrist et al. (1996), and Angrist et al. (2000). Our results in this section are most closely related to Jordá et al. (2020), who used a potential outcome model analogous to that introduced in Angrist et al. (2018), to understand the causal content of local projection-IV with a binary assignment and binary instrument.

In particular, we ask if the following statistical estimands have causal meaning: the Wald estimand, the IV estimand, the generalized Wald estimand, and the filtered IV estimand. Table 2 defines these estimands and summarizes our main results on their causal interpretation under important restrictions on the assignment process and other technical conditions. The rest of this Section spells out the details.

Name	Estimand	Causal Interpretation
Wald	$\frac{\mathbb{E}[Y_{t+h}\|Z_t=z]-\mathbb{E}[Y_{t+h}\|Z_t=z']}{\mathbb{E}[W_{k,t}\|Z_t=z]-\mathbb{E}[W_{k,t}\|Z_t=z']}$	$\frac{\int_{\mathcal{W}}\mathbb{E}[Y'_{t+h}(w_k)\|H_t(w_k)=1]\mathbb{E}[H_t(w_k)]dw_k}{\int_{\mathcal{W}}\mathbb{E}[H_t(w_k)]dw_k}$
IV	$\frac{Cov(Y_{t+h},Z_t)}{Cov(W_t,Z_t)}$	$\frac{\int_{\mathcal{W}_Z}\mathbb{E}[Y'_{t+h}(z_t)]\mathbb{E}[G_t(z_t)]dz_t}{\int_{\mathcal{W}_Z}\mathbb{E}[W'_t(z_t)]\mathbb{E}[G_t(z_t)]dz_t}$
Generalized Wald	$\frac{\mathbb{E}[Y_{t+h}\|Z_t=z,\mathcal{F}_{t-1}]-\mathbb{E}[Y_{t+h}\|Z_t=z',\mathcal{F}_{t-1}]}{\mathbb{E}[W_{k,t}\|Z_t=z,\mathcal{F}_{t-1}]-\mathbb{E}[W_{k,t}\|Z_t=z',\mathcal{F}_{t-1}]}$	$\frac{\int_{\mathcal{W}}\mathbb{E}[Y'_{t+h}(w_k)\|H_t(w_k)=1,\mathcal{F}_{t-1}]\mathbb{E}[H_t(w_k)\|\mathcal{F}_{t-1}]dw_k}{\int_{\mathcal{W}}\mathbb{E}[H_t(w_k)\|\mathcal{F}_{t-1}]dw_k}$
Filtered IV	$\frac{\mathbb{E}[(Y_{t+h}-\hat{Y}_{t+h\|t-1})(Z_t-\hat{Z}_{t\|t-1})]}{\mathbb{E}[(W_{k,t}-\hat{W}_{k,t\|t-1})(Z_t-\hat{Z}_{t\|t-1})]}$	$\frac{\int_{\mathcal{W}_Z}\mathbb{E}[\mathbb{E}[Y'_{t+h}(z_t)\|\mathcal{F}_{t-1}]\mathbb{E}[G_t(z_t)\|\mathcal{F}_{t-1}]]dz_t}{\int_{\mathcal{W}_Z}\mathbb{E}[\mathbb{E}[W'_t(z_t)\|\mathcal{F}_{t-1}]\mathbb{E}[G_t(z_t)\|\mathcal{F}_{t-1}]]dz_t}$

Table 2: Top line results for the causal interpretation of common estimands based on assignments, instruments and outcomes. Here $h \geq 0$, $z, z' \in \mathcal{W}_Z$, $Y_{t+h}(z_t) := Y_{t+h}(W_{1:t-1}, W_{t,1:k-1}, W_k(z_t), W_{t,k+1:d_W}, W_{t+1:t+h})$, $Y'_{t+h}(z_t) := \partial Y_{t+h}(z_t)/\partial z_t$, $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$, $G_t(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t])$ and $G_{t|t-1}(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t] \mid \mathcal{F}_{t-1}])$, while $\hat{Y}_{t+h|t-1} = \mathbb{E}[Y_{t+h} \mid \mathcal{F}_{t-1}]$, $\hat{Z}_{t|t-1} = \mathbb{E}[Z_t \mid \mathcal{F}_{t-1}]$ and $\hat{W}_{k,t|t-1} = \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}]$. Note that $\mathbb{E}[G_t(z_t)] \geq 0$ and $\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}] \geq 0$.

5.1 Wald Estimand

Consider the classic Wald estimand

$$\frac{\mathbb{E}[Y_{t+h} \mid Z_t = z] - \mathbb{E}[Y_{t+h} \mid Z_t = z']}{\mathbb{E}[W_{k,t} \mid Z_t = z] - \mathbb{E}[W_{k,t} \mid Z_t = z']}.$$

The numerator is the impulse response of the outcome Y_{t+h} on the instrument Z_t, which can be thought of as the "reduced-form." The denominator is the impulse response function of the assignment $W_{k,t}$ on the instrument Z_t, which can be thought of as the "first-stage." Our next result establishes that the Wald estimand identifies a weighted average of marginal causal effects for "compliers" provided that (i) the potential outcome process is continuously differentiable in the assignment; (ii) the instrument is independent of the potential assignment and outcome processes; (iii) is a relevance condition; (iv) satisfies a monotonicity condition as introduced in Imbens and Angrist (1994).

Theorem 6. *Assume an instrumented potential outcome system, fix $z, z' \in \mathcal{W}_Z$ and that*

i. *<u>Differentiability</u>: $Y_{t+h}(w_k)$ is continuously differentiable in the closed interval $w_k \in \mathcal{W}_k := [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}$.*

ii. *<u>Independence</u>: The instrument satisfies $Z_t \perp\!\!\!\perp \{W_{k,t}(z) \colon z \in \mathcal{W}_Z\}$ and $Z_t \perp\!\!\!\perp \{Y_{t+h}(w_k) \colon w_k \in \mathcal{W}_k\}$.*

iii. *<u>Relevance</u>: $\int_{\mathcal{W}} \mathbb{E}[1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}]dw_k > 0$.*

iv. Monotonicity: $W_{k,t}(z') \leq W_{k,t}(z)$ *with probability one.*

Then, Wald estimand equals, so long as it exists,

$$\frac{\int_{\mathcal{W}} \mathbb{E}[Y'_{t+h}(w_k)|H_t(w_k) = 1]\mathbb{E}[H_t(w_k)]dw_k}{\int_{\mathcal{W}} \mathbb{E}[H_t(w_k)]dw_k},$$

where $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$.

Provided the instrument is randomly assigned, relevant, and satisfies a monotonicity condition, then the Wald estimand equals a weighted average of the marginal causal effects for "compliers" (i.e., realizations of the potential assignment function for which moving the instrument from z' to z changes the assignment). The marginal causal effect is the derivative of the h-step ahead potential outcome process with respect to the k-th assignment, holding all else constant. The weights are proportional to the probability of the potential assignment function being a "complier," so are non-negative and sum to one.

Since $Y_{t+h}(w_k) := Y_{t+h}(W_{1:t-1}, W_{1:k-1,t}, w_k, W_{k+1:d_W,t}, W_{t+1:t+h})$, Assumption (ii) implicitly restricts the relationship between the instrument Z_t and:

1. other assignments $W_{1:k-1,1:t+h}, W_{k+1:d_W,1:t+h}$,
2. future and past potential assignments $\{W_{k,1:t-1}(z_{1:t-1}), W_{k,t+1:t+h}(z_{t+1:t+h})\colon z_{1:t-1} \in \mathcal{Z}^t, z_{t+1:t+h} \in \mathcal{Z}^h\}$,
3. future and past instruments $Z_{1:t-1}$ and $Z_{t+1:t+h}$, and
4. the potential outcome process $\{Y_{j,t+h}(w_{1:t+h})\colon w_{1:t+h} \in \mathcal{W}^{t+h}\}$.

We could extend Theorem 2 to the instrumented potential outcome system, and show that Assumption (ii) is implied by restricting the instrument Z_t to be independent of each of these quantities.

Remark 4 (Binary Assignment, Binary Instrument Case)**.** *Consider the simplest case with* $W_{k,t} \in \{0, 1\}$*,* $Z_t \in \{0, 1\}$ *and* $z = 1, z' = 0$*. Although the math is different due to the discreteness of the assignment and instrument, under the same conditions as Theorem 8, we can show that the Wald estimand in this case equals*

$$\mathbb{E}[\{Y_{t+h}(1) - Y_{t+h}(0)\} \mid W_{k,t}(1) - W_{k,t}(0) = 1],$$

which is the time-series generalization of the binary assignment, binary instrument local average treatment effect originally derived in Imbens and Angrist (1994).

5.2 IV Estimand

Rather than directly estimating the Wald estimand, it is natural to estimate a two-stage least squares regression of the outcome Y_{t+h} on the assignment $W_{k,t}$ using the instrument Z_t. The associated IV estimand is

$$IV_{k,t,h} := \frac{Cov(Y_{t+h}, Z_t)}{Cov(W_t, Z_t)}.$$

This has a causal interpretation by applying Theorem 3 for the local projection estimand on both the numerator for the local projection of Y_{t+h} on Z_t and the denominator for the local projection of W_t on Z_t. The statement of the results uses the notation

$$Y_{t+h}(z_t) := Y_{t+h}(W_{1:t-1}, W_{t,1:k-1}, W_k(z_t), W_{t,k+1:d_W}, W_{t+1:t+h}),$$

and $Y'_{t+h}(z_t) := \partial Y_{t+h}(z_t)/\partial z_t$.

Theorem 7. *Assume an instrumented potential outcome system. Further assume that*

i. Differentiability: $Y_{t+h}(z)$ and that $W_t(z)$ are continuously differentiable in the closed interval $z \in \mathcal{W}_Z = [\underline{z}, \overline{z}] \subset \mathbb{R}$.

ii. Independence: $Z_t \perp\!\!\!\perp \{W_t(z) \colon z \in \mathcal{W}_Z\}, \quad Z_t \perp\!\!\!\perp \{Y_{t+h}(z) \colon z \in \mathcal{W}_Z\}$,

iii. Relevance: $\int_{\mathcal{W}_Z} \mathbb{E}[W'_t(z_t)]\mathbb{E}[G_t(z_t)]dz_t \neq 0$.

Then, it follows, if it exists, that

$$IV_{k,t,h} = \frac{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{t+h}(z_t)]\mathbb{E}[G_t(z_t)]dz_t}{\int_{\mathcal{W}_Z} \mathbb{E}[W'_t(z_t)]\mathbb{E}[G_t(z_t)]dz_t}$$

where $G_t(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t])$, noting $\mathbb{E}[G_t(z_t)] \geq 0$.

5.3 Generalized Wald Estimand

The generalized Wald estimand is a ratio of a reduced-form generalized impulse response function to a first-stage generalized impulse response function. It is given by, for fixed $z, z' \in \mathcal{W}_Z$,

$$\frac{\mathbb{E}[Y_{t+h} \mid Z_t = z, \mathcal{F}_{t-1}] - \mathbb{E}[Y_{t+h} \mid Z_t = z', \mathcal{F}_{t-1}]}{\mathbb{E}[W_{k,t} \mid Z_t = z, \mathcal{F}_{t-1}] - \mathbb{E}[W_{k,t} \mid Z_t = z', \mathcal{F}_{t-1}]}. \tag{13}$$

Theorem 8. *Assume an instrumented potential outcome system, fix $z, z' \in \mathcal{W}_Z$ and that*

i. Differentiability: $Y_{t+h}(w_k)$ is continuously differentiable in the closed interval $w_k \in \mathcal{W}_k := [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}$.

ii. *Independence:* *The instrument satisfies* $[Z_t \perp\!\!\!\perp \{W_{k,t}(z)\colon z \in \mathcal{W}_Z\}] \mid \mathcal{F}_{t-1}$ *and* $[Z_t \perp\!\!\!\perp \{Y_{t+h}(w_k)\colon w_k \in \mathcal{W}_k\}] \mid \mathcal{F}_{t-1}$.

iii. *Relevance:* $\int_{\mathcal{W}} \mathbb{E}[1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\} \mid \mathcal{F}_{t-1}]dw_k > 0$.

iv. *Monotonicity:* $W_{k,t}(z') \leq W_{k,t}(z)$ *with probability one.*

Then, the generalized Wald estimand equals, so long as it exists,

$$\frac{\int_{\mathcal{W}} \mathbb{E}[Y'_{t+h}(w_k) \mid H_t(w_k) = 1, \mathcal{F}_{t-1}]\mathbb{E}[H_t(w_k) \mid \mathcal{F}_{t-1}]dw_k}{\int_{\mathcal{W}} \mathbb{E}[H_t(w_k) \mid \mathcal{F}_{t-1}]dw_k},$$

where, again, $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$.

The generalized Wald estimand analogously equals a weighted average of the marginal filtered causal effects for "compliers," where the weights are proportional to the probability of the potential assignment function being a "complier" conditional on the filtration.

We next provide a sufficient condition for the instrument to be randomly assigned in terms of conditional independence restrictions on these underlying processes.

Theorem 9. *Assume that the instrument satisfies*

$$Z_t \perp\!\!\!\perp \big(Z_{t+1:t+h}, W_{1:k-1,t:t+h}, \{W_{k,t+1:t+h}(z_{t+1:t+h})\colon z_{t+1:t+h} \in \mathcal{Z}^h\}, W_{k+1:d_W,t:t+h}, \\ \{Y_{t+h}(w_{1:t+h})\colon w_{1:t+h} \in \mathcal{W}^{t+h}\}\big) \mid \mathcal{F}_{t-1}.$$

Then, Assumption (ii) in Theorem 8 is satisfied.

5.4 Generalized IV and Filtered IV Estimands

Estimating generalized Wald estimand is not easy, particularly if Z_t is not discrete. Here we derive a causal interpretation for *generalized IV estimand*

$$\frac{Cov(Y_{t+h}, Z_t \mid \mathcal{F}_{t-1})}{Cov(W_{k,t}, Z_t \mid \mathcal{F}_{t-1})} = \frac{\mathbb{E}[(Y_{t+h} - \hat{Y}_{t+h|t-1})(Z_t - \hat{Z}_{t|t-1}) \mid \mathcal{F}_{t-1}]}{\mathbb{E}[(W_{k,t} - \hat{W}_{k,t|t-1})(Z_t - \hat{Z}_{t|t-1}) \mid \mathcal{F}_{t-1}]}.$$

where $\hat{Y}_{t+h|t-1} = \mathbb{E}[Y_{t+h} \mid \mathcal{F}_{t-1}]$, $\hat{W}_{k,t|t-1} = \mathbb{E}[W_{k,t} \mid \mathcal{F}_{t-1}]$ and $\hat{Z}_{t|t-1} = \mathbb{E}[Z_t \mid \mathcal{F}_{t-1}]$.

No new technical issues arise in dealing with this setup, but Assumption (ii) in Theorem 7 now becomes

$$[Z_t \perp\!\!\!\perp \{Y_{t+h}(z)\colon z \in \mathcal{W}_Z\}] \mid \mathcal{F}_{t-1}, \quad [Z_t \perp\!\!\!\perp \{W_t(z)\colon z \in \mathcal{W}_Z\}] \mid \mathcal{F}_{t-1}. \tag{14}$$

Then, the generalized IV estimand equals

$$\frac{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{t+h}(z_t) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}]dz_t}{\int_{\mathcal{W}_Z} \mathbb{E}[W'_t(z_t) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}]dz_t}$$

where $G_{t|t-1}(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t \mid \mathcal{F}_{t-1}])$, noting $\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}] \geq 0$.

Of more practical importance is the *filtered IV estimand*

$$\frac{\mathbb{E}[(Y_{t+h} - \hat{Y}_{t+h|t-1})(Z_t - \hat{Z}_{t|t-1})]}{\mathbb{E}[(W_{k,t} - \hat{W}_{k,t|t-1})(Z_t - \hat{Z}_{t|t-1})]},$$

which can be estimated by instrumental variables applied to $Y_{t+h} - \hat{Y}_{t+h|t-1}$ on $W_{k,t} - \hat{W}_{k,t|t-1}$ with instruments $Z_t - \hat{Z}_{t|t-1}$. Under the conditions of Theorem 7 but using (14) instead of Assumption (ii), then the filtered IV estimand becomes

$$\frac{\int_{\mathcal{W}_Z} \mathbb{E}[\mathbb{E}[Y'_{t+h}(z_t) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}]]dz_t}{\int_{\mathcal{W}_Z} \mathbb{E}[\mathbb{E}[W'_t(z_t) \mid \mathcal{F}_{t-1}]\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}_{t-1}]]dz_t}.$$

6 Estimands Based on Instruments and Outcomes

In this section, we study the nonparametric conditions under which common statistical estimands based on only instruments and outcomes have causal meaning. We focus on an instrumented potential outcome system

$$\{Z_t, \{W_t(z_t), z_t \in \mathcal{W}_Z\}, \{Y_t(w_{1:t}), w_{1:t} \in \mathcal{W}_W^t\}\}_{t\geq 1},$$

in which the researcher only observes the instruments and the outcomes $\{z_t^{obs}, y_t^{obs}\}_{t\geq 1}$. We will sometimes refer to $\{\mathcal{F}_t^{Z,Y}\}_{t\geq 1}$ as the natural filtration generated by the realized $\{z_t^{obs}, y_t^{obs}\}_{t\geq 1}$.

In this context, it is common for empirical researchers to analyze estimands involving two elements of the outcome vector $Y_{j,t+h}$, $Y_{k,t}$ and the instrument Z_t (therefore, we return to using a explicit subscript on the outcome variable). Consider, for example, an empirical researcher that constructs an instrument Z_t for the monetary policy shock (e.g., an instrument of the form used in Kuttner (2001); Cochrane and Piazessi (2002); Gertler and Karadi (2015) or Romer and Romer (2004)). In this case, the empirical researcher may measure the dynamic causal effect of the monetary policy shock $W_{k,t}$ on unemployment $Y_{j,t+h}$ by estimating the first-stage impulse response function of the federal funds rate $Y_{k,t}$ on the instrument Z_t. See, for example, Jordá et al. (2015); Ramey and Zubairy (2018); Jordá et al. (2020) for recent empirical applications of this empirical strategy.

In particular, we ask if the following estimands have causal meaning: Ratio Wald, Local Projection IV, generalized Ratio Wald, and the local filtered projection IV. We show that such dynamic IV estimands identify "relative" local average impulse causal effect, which is a nonparametrically generalization of the interpretation of such a dynamic IV estimand in existing literature on external instruments (Stock and Watson, 2018; Plagborg-Møller and Wolf, 2020; Jordá et al., 2020). Table 3 defines these estimands and summarizes our main results on their causal interpretation under important restrictions on the assignment process and other technical conditions. The rest of this Section spells out the details.

Name	**Estimand**	**Causal Interpretation**
Ratio Wald	$\frac{\mathbb{E}[Y_{j,t+h}\mid Z_t=z]-\mathbb{E}[Y_{j,t+h}\mid Z_t=z']}{\mathbb{E}[Y_{k,t}\mid Z_t=z]-\mathbb{E}[Y_{k,t}\mid Z_t=z']}$	$\frac{\int_{\mathcal{W}}\mathbb{E}[Y'_{j,t+h}(w_k)\mid H_t(w_k)=1]\mathbb{E}[H_t(w_k)]dw_k}{\int_{\mathcal{W}}\mathbb{E}[Y'_{k,t}(w_k)\mid H_t(w_k)=1]\mathbb{E}[H_t(w_k)]dw_k}$
Local Projection IV	$\frac{Cov(Y_{j,t+h},Z_t)}{Cov(Y_{k,t},Z_t)}$	$\frac{\int_{\mathcal{W}_Z}\mathbb{E}[Y'_{j,t+h}(z_k)]\mathbb{E}[G_t(z_k)]dz_k}{\int_{\mathcal{W}_Z}\mathbb{E}[Y'_{k,t}(z_k)]\mathbb{E}[G_t(z_k)]dz_k}$
Generalized Ratio Wald	$\frac{\mathbb{E}[Y_{j,t+h}\mid Z_t=z,\mathcal{F}^{Z,Y}_{t-1}]-\mathbb{E}[Y_{j,t+h}\mid Z_t=z',\mathcal{F}^{Z,Y}_{t-1}]}{\mathbb{E}[Y_{k,t}\mid Z_t=z,\mathcal{F}^{Z,Y}_{t-1}]-\mathbb{E}[Y_{k,t}\mid Z_t=z',\mathcal{F}^{Z,Y}_{t-1}]}$	$\frac{\int_{\mathcal{W}}\mathbb{E}[Y'_{j,t+h}(w_k)\mid H_t(w_k)=1,\mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[H_t(w_k)\mid\mathcal{F}^{Z,Y}_{t-1}]dw_k}{\int_{\mathcal{W}}\mathbb{E}[Y'_{k,t}(w_k)\mid H_t(w_k)=1,\mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[H_t(w_k)\mid\mathcal{F}^{Z,Y}_{t-1}]dw_k}$
Local Filtered Projection IV	$\frac{Cov(Y_{j,t+h}-\hat{Y}_{j,t+h\mid t-1},Z_t-\hat{Z}_{t\mid t-1})}{Cov(Y_{k,t}-\hat{Y}_{k,t\mid t-1},Z_t-\hat{Z}_{t\mid t-1})}$	$\frac{\int_{\mathcal{W}_Z}\mathbb{E}[\mathbb{E}[Y'_{j,t+h}(z_k)\mid\mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[G_t(z_k)\mid\mathcal{F}^{Z,Y}_{t-1}]]dz_k}{\int_{\mathcal{W}_Z}\mathbb{E}[\mathbb{E}[Y'_{k,t}(z_k)\mid\mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[G_t(z_k)\mid\mathcal{F}^{Z,Y}_{t-1}]]dz_k}$

Table 3: Top line results for the causal interpretation of common estimands based on instruments and outcomes. Here $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$, $G_t(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t])$ and $G_{t|t-1}(z_t) = 1\{z_t \leq Z_t\}(Z_t - \mathbb{E}[Z_t \mid \mathcal{F}^{Z,Y}_{t-1}])$, while $\hat{Y}_{k,t+h|t-1} = \mathbb{E}[Y_{k,t+h} \mid \mathcal{F}^{Z,Y}_{t-1}]$ and $\hat{Z}_{t|t-1} = \mathbb{E}[Z_t \mid \mathcal{F}^{Z,Y}_{t-1}]$. Note that $\mathbb{E}[G_t(z_t)] \geq 0$ and $\mathbb{E}[G_{t|t-1}(z_t) \mid \mathcal{F}^{Z,Y}_{t-1}] \geq 0$.

6.1 Ratio Wald Estimand

The *Ratio Wald Estimand*

$$\frac{\mathbb{E}[Y_{j,t+h} \mid Z_t = z] - \mathbb{E}[Y_{j,t+h} \mid Z_t = z']}{\mathbb{E}[Y_{k,t} \mid Z_t = z] - \mathbb{E}[Y_{k,t} \mid Z_t = z']},$$

which is the ratio of the Wald estimands:

$$\frac{\mathbb{E}[Y_{j,t+h} \mid Z_t = z] - \mathbb{E}[Y_{j,t+h} \mid Z_t = z']}{\mathbb{E}[W_{k,t} \mid Z_t = z] - \mathbb{E}[W_{k,t} \mid Z_t = z']}, \quad \text{to} \quad \frac{\mathbb{E}[Y_{k,t} \mid Z_t = z] - \mathbb{E}[Y_{k,t} \mid Z_t = z']}{\mathbb{E}[W_{k,t} \mid Z_t = z] - \mathbb{E}[W_{k,t} \mid Z_t = z']}.$$

Hence we just need to collect the conditions for the validity of their causal representations, and then apply Theorem 6 twice.

Corollary 1. *Assume an instrumented potential outcome system, $z, z' \in \mathcal{W}_Z$ and that* *i.* *Differentiability:* $Y_{k,t}(w_k), Y_{j,t+h}(w_k)$ *are continuously differentiable in closed interval* $\mathcal{W}_k := [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}$.

ii. *Independence:* $Z_t \perp\!\!\!\perp \{W_{k,t}(z)\colon z \in \mathcal{W}_Z\}$ *and* $Z_t \perp\!\!\!\perp \{Y_{k,t}(w_k), Y_{j,t+h}(w_k)\colon w_k \in \mathcal{W}_k\}$.

iii. *Relevance:* $\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t}(w_k) \mid H_t(w_k) = 1]\mathbb{E}[H_t(w_k)]dw_k \neq 0$.

iv. *Monotonicity:* $W_{k,t}(z') \leq W_{k,t}(z)$ *with probability one.*

Then, the Ratio Wald Estimand equals, if it exists,

$$\frac{\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t+h}(w_k) \mid H_t(w_k) = 1]\mathbb{E}[H_t(w_k)]dw_k}{\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t}(w_k) \mid H_t(w_k) = 1]\mathbb{E}[H_t(w_k)]dw_k},$$

where $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$.

In words, the ratio Wald estimand above identifies a *relative* local average impulse causal effect under the instrumented potential outcome system. The numerator is a weighted average of the marginal causal effects of $W_{k,t}$ on the h-step ahead outcome $Y_{j,t+h}$, where the weights are proportional to the probability of compliance. Similarly, the denominator is a weighted average of the marginal causal effects of $W_{k,t}$ on the contemporaneous outcome $Y_{k,t}$. Therefore, the ratio in Corollary 1 measures the causal response of the h-step ahead outcome $Y_{j,t+h}$ to a change in the treatment $W_{k,t}$ that increases the contemporaneous outcome $Y_{k,t}$ by one unit on impact (among compliers).

This is a nonparametric generalization of the well-known result that in linear SVMA models (without invertibility) the IV based estimands identify relative impulse response functions (Stock and Watson, 2018; Plagborg-Møller and Wolf, 2020). Corollary 1 makes no functional form assumptions nor standard time series assumptions such as invertibility or recoverability. In this sense, Corollary 1 highlights the attractiveness of using external instruments to measure dynamic causal effects in observational time series data. Provided there exists an external instrument for the treatment $W_{k,t}$ that is randomly assigned, relevant and satisfies a monotonicity condition, then the researcher can identify causally interpretable estimands without further assumptions and without even directly observing the treatment itself.

6.2 Local Projection IV Estimand

The *local projection IV estimand*

$$\frac{Cov(Y_{j,t+h}, Z_t)}{Cov(Y_{k,t}, Z_t)},$$

is the ratio of the IV estimands $\frac{Cov(Y_{j,t+h},Z_t)}{Cov(W_t,Z_t)}$ to $\frac{Cov(Y_{k,t},Z_t)}{Cov(W_t,Z_t)}$. Therefore, we once again just need to collect the conditions for the validity of their causal representations, and apply Theorem 7 twice.

Corollary 2. *Consider an instrumented potential outcome system. Further assume that*

i. *Differentiability: $Y_{j,t}(z), Y_{k,t+h}(z), W_t(z)$ are continuously differentiable in the closed interval $z \in \mathcal{W}_Z = [\underline{z}, \overline{z}] \subset \mathbb{R}$.*

ii. *Independence: $Z_t \perp\!\!\!\perp \{Y_{k,t}(z), Y_{j,t+h}(z) \colon z \in \mathcal{W}_Z\}$ and $Z_t \perp\!\!\!\perp \{W_t(z) \colon z \in \mathcal{W}_Z\}$.*

iii. *Relevance: $\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{k,t}(z_t)]\mathbb{E}[G_t(z_t)]dz_t \neq 0$.*

Then, the local projection IV estimand equals

$$\frac{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{j,t+h}(z_k)]\mathbb{E}[G_t(z_k)]dz_k}{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{k,t}(z_k)]\mathbb{E}[G_t(z_k)]dz_k},$$

where $G_t(z_k) = 1\{z_k \leq Z_t\}(Z_t - \mathbb{E}[Z_t])$, noting $\mathbb{E}[G_t(z_k)] \geq 0$.

6.3 Generalized Ratio Wald Estimand

Researchers may also be interested in analyzing the *generalized ratio Wald estimand*:

$$\frac{\mathbb{E}[Y_{j,t+h} \mid Z_t = z, \mathcal{F}^{Z,Y}_{t-1}] - \mathbb{E}[Y_{j,t+h} \mid Z_t = z', \mathcal{F}^{Z,Y}_{t-1}]}{\mathbb{E}[Y_{k,t} \mid Z_t = z, \mathcal{F}^{Z,Y}_{t-1}] - \mathbb{E}[Y_{k,t} \mid Z_t = z', \mathcal{F}^{Z,Y}_{t-1}]},$$

which is the ratio of generalized impulse response functions at different lags and for different outcome variables. Since this is the ratio of two generalized Wald estimands, we immediately arrive at the following corollary by applying Theorem 8 twice.

Corollary 3. *Assume an instrumented potential outcome system, $z, z' \in \mathcal{W}_Z$ and that*

i. *Differentiability: $Y_{k,t}(w_k), Y_{j,t+h}(w_k)$ are continuously differentiable in closed interval $\mathcal{W}_k := [\underline{w}_k, \overline{w}_k] \subset \mathbb{R}$.*

ii. *Independence: $[Z_t \perp\!\!\!\perp \{W_{k,t}(z) \colon z \in \mathcal{W}_Z\}] \mid \mathcal{F}^{Z,Y}_{t-1}$ and $[Z_t \perp\!\!\!\perp \{Y_{k,t}(w_k), Y_{j,t+h}(w_k) \colon w_k \in \mathcal{W}_k\}] \mid \mathcal{F}^{Z,Y}_{t-1}$.*

iii. *Relevance: $\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t}(w_k) \mid H_t(w_k) = 1, \mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[H_t(w_k) \mid \mathcal{F}^{Z,Y}_{t-1}]dw_k \neq 0$.*

iv. *Monotonicity: $W_{k,t}(z') \leq W_{k,t}(z) \mid \mathcal{F}^{Z,Y}_{t-1}$ with probability one.*

Then, the generalized ratio Wald estimand equals

$$\frac{\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t+h}(w_k) \mid H_t(w_k) = 1, \mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[H_t(w_k) \mid \mathcal{F}^{Z,Y}_{t-1}]dw_k}{\int_{\mathcal{W}} \mathbb{E}[Y'_{k,t}(w_k) \mid H_t(w_k) = 1, \mathcal{F}^{Z,Y}_{t-1}]\mathbb{E}[H_t(w_k) \mid \mathcal{F}^{Z,Y}_{t-1}]dw_k},$$

where $H_t(w_k) = 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}$.

The interpretation of Corollary 3 is analogous to the interpretation of the ratio Wald estimand in Corollary 1, except now everything is conditional on the natural filtration.

6.4 Generalized Local Projection IV and Local Filtered Projection IV Estimands

In practice researchers typically estimate generalized impulse response functions using a two-stage least-squares type estimator. This is also sometimes called "local projections with an external instrument" (Jordá et al., 2015). We first analyze this *generalized local projection IV*

$$\frac{Cov(Y_{j,t+h}, Z_t \mid \mathcal{F}_{t-1}^{Z,Y})}{Cov(Y_{k,t}, Z_t \mid \mathcal{F}_{t-1}^{Z,Y})}, \tag{15}$$

which again is a ratio, this time of the Generalized IV estimands at different lag lengths. Using the same arguments as Corollary 2, it has the causal interpretation

$$\frac{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{j,t+h}(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]\mathbb{E}[G_t(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]dz_k}{\int_{\mathcal{W}_Z} \mathbb{E}[Y'_{k,t}(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]\mathbb{E}[G_t(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]dz_k},$$

where $G_{t|t-1}(z_k) = 1\{z_k \leq Z_t\}(Z_t - \mathbb{E}[Z_t \mid \mathcal{F}_{t-1}^{Z,Y}])$.

Of more practical relevance, is the *local filtered projection IV estimand* is

$$\frac{Cov(Y_{j,t+h} - \hat{Y}_{j,t+h|t-1}, Z_t - \hat{Z}_{t|t-1})}{Cov(Y_{k,t} - \hat{Y}_{k,t|t-1}, Z_t - \hat{Z}_{t|t-1})},$$

where recall that, for example, $\hat{Y}_{k,t+h} = \mathbb{E}[Y_{k,t+h} \mid \mathcal{F}_{t-1}^{Z,Y}]$, and $\hat{Z}_{t+h|} = \mathbb{E}[Z_{t+h} \mid \mathcal{F}_{t-1}^{Z,Y}]$. The properties of this are inherited from those of the generalized local projection IV. In particular, it equals

$$\frac{\int_{\mathcal{W}_Z} \mathbb{E}[\mathbb{E}[Y'_{j,t+h}(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]\mathbb{E}[G_t(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]]dz_k}{\int_{\mathcal{W}_Z} \mathbb{E}[\mathbb{E}[Y'_{k,t}(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]\mathbb{E}[G_t(z_k) \mid \mathcal{F}_{t-1}^{Z,Y}]]dz_k}.$$

7 Estimands Based Only on Outcomes

The dominant approach to causal inference in macroeconometrics is a model-based approach in the tradition of Sims (1980). See, for example, Ramey (2016) and Kilian and Lutkepohl (2017) for recent reviews. In that literature, researchers introduce parametric models to study the dynamic causal effects of unobservable "structural shocks," which themselves must be inferred from the outcomes. Here we link this to our setup, mostly to place our work in context and illustrate that the enormous macroeconometric literature on simultaneous equation modelling can be nested in the direct potential outcome system framework. Assume there is a direct potential outcome system

$$\{W_t, \{Y_t(w_{1:t}) : w_{1:t} \in \mathcal{W}_w^t\}_{t\geq 1},$$

where researchers only see the outcomes $\{y_t^{obs}\}_{t\geq 1}$.

7.1 Linear simultaneous equation approach

The causal inference approach of using only time series data on outcomes is in the storied tradition of linear simultaneous equations models developed at the Cowles Foundation (e.g., Christ, 1994; Hausman, 1983). The most essential causal challenges arise without any dynamic causal effects, so we start with a static example as an illustration. Suppose that

$$A_0 Y_t(w_{1:t}) = \alpha + w_t, \quad w_{1:t} \in \mathcal{W}^t, \quad t = 1, 2, ...,$$

where A_0 is a non-stochastic, square matrix. Notice that in this model the potential outcome process is deteministic and linear combinations of the potential outcomes equal the possible assignments for every t. If A_0 is additionally invertible, then

$$Y_t(w_{1:t}) = A_0^{-1}(\alpha + w_t),$$

which implies that the contemporaneous average treatment effect is $\mathbb{E}[Y_t(W_{1:t-1}, w) - Y_t(W_{1:t-1}, w')] = A_0^{-1}(w - w')$, and the marginal average treatment effect is $\mathbb{E}[\frac{\partial Y_t(w_{1:t})}{\partial w_t^{\mathrm{T}}}] = A_0^{-1}$ whatever probabilistic assumption is made about $W_{1:t-1}$.

Furthermore, under this model, if we see $(W_t, Y_t) = \{W_t, Y_t(W_{1:t})\}$, then, if the second moments of the observables exist and $\mathrm{Var}(W_t)$ is non-singular, then for every t,

$$\mathrm{Cov}(Y_t, W_t)\mathrm{Var}(W_t)^{-1} = A_0^{-1},$$

which would make statistical inference rather straightforward. But the point of this simultaneous equations literature is to carry out inference without directly observing the assignments — which is a much harder task.

If, in addition to A_0 being invertible, we assume that $\mathrm{Var}(W_t) < \infty$, then

$$\mathrm{Var}(Y_t) = A_0^{-1}\mathrm{Var}(W_t)\left(A_0^{-1}\right)^{\mathrm{T}},$$

Crucially knowing $\mathrm{Var}(Y_t)$ is not enough to untangle A_0 and $\mathrm{Var}(W_t)$, and so knowledge of the second moments of the observables is not enough alone to learn the contemporaneous average treatment effect. In the linear simultaneous equations literature, this is resolved by a priori imposing more economic structure on the potential outcome process, such as placing more structure on the matrix A_0.

A central a priori constraint is the one highlighted by Sims (1980). He imposed that (a) A_0 is triangular, (b) $\mathrm{Var}(W_t)$ is diagonal. For simplicity of exposition, look at the two dimensional case and write

$$A_0 = \begin{pmatrix} 1 & 0 \\ -a_{21} & 1 \end{pmatrix}, \quad A_0^{-1} = \begin{pmatrix} 1 & 0 \\ a_{21} & 1 \end{pmatrix}, \quad \mathrm{Var}(W_t) = \begin{pmatrix} \sigma_{11}^2 & 0 \\ 0 & \sigma_{22}^2 \end{pmatrix},$$

then the elements within A_0 and $\mathrm{Var}(W_t)$ can be individually determined from $\mathrm{Var}(Y_t)$ if $\mathrm{Var}(Y_t)$ is of full rank. The same holds in higher dimensions. Hence, with additional restrictions on the potential outcome process, the contemporaneous causal effect can be determined from the data on the outcomes, without having observing the assignments (or without the access to instruments). There are alternative a priori constraints to this triangular which also work here and the above structure extends to non-linear systems of equations $g\left(Y_t(w_{1:t})\right) = w_t$.

The linear "structural vector autoregressive " (SVAR) version of the linear simultaneous equation has the same fundamental structure. Focusing on the one lag model with no intercept for simplicity, the SVAR approach assumes that the potential outcome process satisfies

$$A_0 Y_t(w_{1:t}) = w_t + A_1 Y_{t-1}(w_{1:t-1}).$$

Kilian and Lutkepohl (2017) provide a book length review of this model structure and its various extensions and implications. Then $A_0\left(I - \Phi_1 L\right) Y_t(w_{1:t}) = w_t$, where L is a lag operator and $\Phi_1 = A_0^{-1} A_1$. So

$$Y_t(w_{1:t}) = A_0^{-1} w_t + \Phi_1 Y_{t-1}(w_{1:t-1}),$$

which in turn implies that the potential outcome process also has an SVMA model representation

$$Y_t(w_{1:t}) = A_0^{-1} w_t + \Phi_1 A_0^{-1} w_{t-1} + \Phi_1^2 A_0^{-1} w_{t-2} + ... + \Phi_1^{t-1} A_0^{-1} w_1 + \Phi_1^t A_0^{-1} Y_0.$$

In this case, the h-period ahead average treatment effect is

$$\mathbb{E}[Y_{t+h}(W_{1:t-1}, w, W_{t+1:t+h}) - Y_{t+h}(W_{1:t-1}, w', W_{t+1:t+h})] = \Phi_1^h A_0^{-1}\left(w - w'\right)$$

and the h-period ahead marginal average treatment effect is $\mathbb{E}[\partial Y_{t+h}(w_{1:t+h})/\partial w_t'] = \Phi_1^h A_0^{-1}$. The time series parameter Φ_1 can be determined from the dynamics of the observable outcomes if this process is stationary. But again A_0 and $\mathrm{Var}(W_t)$ cannot be separately identified from the observable outcomes, so further structural assumptions are needed.

7.2 Causal meaning of the GIRF of $Y_{k,t}$ on $Y_{j,t+h}$

A broader analysis focuses on the h-step ahead generalized impulse response function of the j-th outcome on the k-th outcome without placing functional form restrictions on the potential outcome process. Here we provide a nonparametric causal meaning to it in terms of potential outcomes. To do so, we will further assume that the potential outcome process is a deterministic function of the assignments and that the assignments are independent across time.

Theorem 10. *Consider a direct potential outcome system, and further assume that*

i. the potential outcome process $Y_t(w_{1:t})$ is deterministic for all $t \geq 0$, $w_{1:t} \in \mathcal{W}^t$.

ii. for all $t \neq s$, $W_t \perp\!\!\!\perp W_s$.

Them, so long as the corresponding moments exist,

$$\mathrm{E}[Y_{j,t+h}|(Y_{k,t} = y_k), \mathcal{F}_{t-1}^Y] - \mathrm{E}[Y_{j,t+h}|(Y_{k,t} = y_k'), \mathcal{F}_{t-1}^Y] \tag{16}$$

$$= \mathrm{E}[\psi_{j,t+h}(W_{1:t})|(Y_{k,t} = y_k), \mathcal{F}_{t-1}^Y] - \mathrm{E}[\psi_{j,t+h}(W_{1:t})|(Y_{k,t} = y_k'), \mathcal{F}_{t-1}^Y], \tag{17}$$

where $\psi_{j,t+h}(w_{1:t}) := \mathrm{E}[Y_{j,t+h}(w_{1:t}, W_{t+1:t+h})]$.

Theorem 10 illustrates that without functional form restrictions on the potential outcome process, the generalized impulse response function of the j-th outcome on the k-th outcome has a causal interpretation in terms of the shifting the entire conditional distribution of the treatments $W_{1:t}$. While this is a non-standard object, it can be interpreted as the causal effect of a stochastic intervention on the assignment path $W_{1:t}$, which has been an object of recent interest in a growing cross-sectional literautre on causal inference in the presence of interference/spillovers across units – see, for example, Munoz and van der Laan (2012), Papadogeorgou et al. (2019), Papadogeorgou et al. (2021), and Wu et al. (2021). Nonetheless, this is a complex causal effect as it measures an average causal effect of simultaneously shifting all assignments from time $t = 1$ to t.

8 Conclusion

In this paper, we developed the nonparametric, direct potential outcome system to study causal inference in observational time series settings. We place no functional form restrictions on the potential outcome process, no restrictions on the extent to which past assignments causally affect the outcomes, nor common time series assumptions such as "invertibility' or "recoverability." The direct potential outcome system therefore nests most leading econometric models used in time series settings as a special case. We then studied conditions on the assignments under which common time series estimands, such as the impulse response functions, generalized impulse response function, and local projections, have a causal interpretation in terms of underlying dynamic causal effects. We further showed that provided the researcher observes an instrument that satisfies an appropriate unconfoundedness and monotonicity condition, then common IV estimands such as local projections instrumenal variables also have causal interpretations in terms of local average dynamic causal effects. Taken together, the potential outcome system provides a flexible, non-parametric foundation for making causal statements from observational time series of outcomes, assignments, and instruments.

References

Abbring, J. H. and J. J. Heckman (2007). Econometric evaluation of social programs, part III: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6, pp. 5145–5303. Amsterdam, The Netherlands: North Holland.

Angrist, J., K. Graddy, and G. Imbens (2000). The interpretation of instrumental variables estimators in simultaneous equation models with an application to the demand for fish. *Review of Economic Studies 67*, 499–527.

Angrist, J. D. and G. W. Imbens (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association 90*(430), 431–442.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*, 444–455.

Angrist, J. D. and G. M. Kuersteiner (2011). Causal effects of monetary shocks: Semiparametric conditional independence tests with a multinomial propensity score. *Review of Economics and Statistics 93*, 725–747.

Angrist, J. D., Óscar Jordá, and G. M. Kuersteiner (2018). Semiparametric estimates of monetary policy effects: String theory revisited. *Journal of Business and Economic Statistics 36*, 381–387.

Aruoba, S. B., M. Mlikota, F. Schorfheide, and S. Villalvazo (2021). SVARs with occasionally-binding constraints. Technical report.

Auerbach, A. J. and Y. Gorodnichenko (2012a). *Fiscal Multipliers in Recession and Expansion*, pp. 63–98. University of Chicago Press.

Auerbach, A. J. and Y. Gorodnichenko (2012b). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy 4*, 1–27.

Baek, C. and B. Lee (2021). A guide to autoregressive distributed lag models for impulse response estimations. Technical report.

Baum, L. E. and T. Petrie (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics 37*, 1554–1563.

Blackwell, M. and A. Glynn (2018). How to make causal inferences with time-series and cross-sectional data. *American Political Science Review 112*, 1067–1082.

Bojinov, I., A. Rambachan, and N. Shephard (2021). Panel experiments and dynamic causal effects: A finite population perspective. *Quantitative Economics*. Forthcoming.

Bojinov, I. and N. Shephard (2019). Time series experiments and causal estimands: Exact randomization tests and trading. *Journal of the American Statistical Association 114*, 1665–1682.

Boruvka, A., D. Almirall, K. Witkiwitz, and S. A. Murphy (2018). Assessing time-varying causal effect moderation in mobile health. *Journal of the American Statistical Association 113*, 1112–1121.

Box, G. E. P. and G. C. Tiao (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association 70*, 70–79.

Bradford Hill, A. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine 58*, 295–300.

Brodersen, K., F. Gallusser, J. Koehler, N. Remy, and S. Scott (2015). Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics 9*, 247–274.

Chahrour, R. and K. Jurado (2021). Recoverability and expectations-driven fluctuations. Technical report.

Chopin, N. and O. Papasphiliopoulos (2020). *An Introduction to Sequential Monte Carlo*. Springer.

Christ, C. F. (1994). The Cowles Commission's contributions to econometrics at Chicago, 1939-1955. *Journal of Economic Literature 32*, 30–59.

Cloyne, J. S., Óscar Jordá, and A. M. Taylor (2020). Decomposing the fiscal multiplier. Technical report.

Cochrane, J. H. and M. Piazessi (2002). The Fed and interest rates - a high-frequency identification. *American Economic Review 92*, 90–95.

Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society, Series B 20*, 215–42.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2 ed.). Oxford: Oxford University Press.

Frisch, R. (1933). *Propagation Problems and Impulse Problems in Dynamic Economics*. London, United Kingdom: Allen and Unwin.

Gallant, A. R., P. E. Rossi, and G. Tauchen (1993). Nonlinear dynamic structures. *Econometrica 61*, 871–907.

Gertler, M. L. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*, 44–76.

Goncalves, S., A. M. Herreraz, L. Kilian, and E. Pesavento (2021). Impulse response analysis for structural dynamic models with nonlinear regressors. Technical report.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993). A novel approach to nonlinear and non-Gaussian Bayesian state estimation. *IEE-Proceedings F 140*, 107–113.

Gourieroux, C. and J. Jasiak (2005). Nonlinear innovations and impulse responses with application to VaR sensitivity. *Annales d'Economie et de Statistique 78*, 1–31.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*, 424–438.

Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357–384.

Hamilton, J. D. (2011). Nonlinearities and the macroeconomic effects of oil prices. *Macroeconomic Dynamics 15*(S3), 364–378.

Harvey, A. C. (1996). Intervention analysis with control groups. *International Statistical Review 64*, 313–328.

Harvey, A. C. and J. Durbin (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. *Journal of the Royal Statistical Society, Series A 149*, 187–227.

Hausman, J. A. (1983). Specification and estimation of simultaneous equation models. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 1, pp. 391–448. North-Holland.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016). Dynamic treatment effects. *Journal of Econometrics 191*, 276–292.

Heckman, J. J. and S. Navarro (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics 136*, 341–396.

Herbst, E. and F. Schorfheide (2015). *Bayesian Estimation of DSGE Models*. Princeton, New Jersey, USA: Princeton University Press.

Hernan, M. A. and J. M. Robins (2021). *Causal Inference*. Boca Raton, Florida, USA: Chapman & Hall. Forthcoming.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*, 467–475.

Jordá, Óscar. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*, 161–182.

Jordá, Óscar., M. Schularick, and A. M. Taylor (2015). Betting the house. *Journal of International Economics 96*, S2–S18.

Jordá, Óscar., M. Schularick, and A. M. Taylor (2020). The effects of quasi-random monetary experiments. *Journal of Monetary Economics 112*, 22–40.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transactions ASMA, Series D 82*, 35–45.

Kilian, L. and H. Lutkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge, United Kingdom: Cambridge University Press.

Killian, L. and R. J. Vigfusson (2011a). Are the responses of the U.S. economy asymmetric in energy price increases and decreases? *Quantitative Economics 2*, 419–453.

Killian, L. and R. J. Vigfusson (2011b). Nonlinearities in the oil price-output relationship. *Macroeconomic Dynamics 15*(S3), 337–363.

Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*, 119–147.

Kuersteiner, G., D. Phillips, and M. Villamizar-Villegas (2018). Effective sterilized foreign exchange intervention? Evidence from a rule-based policy. *Journal of International Economics 118*, 118–138.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics 47*, 523–544.

Lechner, M. (2011). The relation of different concepts of causality used in time series and microeconomics. *Econometric Reviews 30*, 109–127.

Lee, S. and B. Salanie (2020). Filtered and unfiltered treatment effects with targeting instruments. Unpublished paper: Department of Economics, Columbia University.

Lu, X., L. Su, and H. White (2017). Granger causality and structural causality in cross-section and panel data. *Econometric Theory 33*, 263–291.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory 4*, 103–124.

Mastakouri, A., B. Scholkopf, and D. Janzing (2021). Necessary and sufficient conditions for causal feature selection in time series with latent common causes. *Proceedings of Machine Learning Research 139*, 7502–7511.

Mavroeidis, S. (2021). Identification at the zero lower bound. Technical report.

Menchetti, F. and I. Bojinov (2021). Estimating the effectiveness of permanent price reductions for competing products using multivariate Bayesian structural time series models. *Journal of Applied Statistics*. Forthcoming.

Munoz, I. D. and M. van der Laan (2012). Population intervention causal effects based on stochastic interventions. *Biometrics 68*, 541–549.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statisticsl Society B 65*, 331–366.

Murphy, S. A., M. J. van der Laan, J. M. Robins, and C. P. P. R. Group (2001). Marginal mean models for dynamic regimes. *Journal of the American Statistical Association 96*, 1410–1423.

Nakamura, E. and J. Steinsson (2018a). High-frequency identification of monetary non-neutrality: The information effect. *The Quarterly Journal of Economics 133*, 1283–1330.

Nakamura, E. and J. Steinsson (2018b). Identification in macroeconomics. *Journal of Economic Perspectives 32*(3), 59–86.

Papadogeorgou, G., K. Imai, J. Lyall, and F. Li (2021). Causal inference with spatio-temporal data: Estimating the effects of airstrikes on insurgent violence in iraq. Technical report, arXiv preprint, arXiv:2003.13555.

Papadogeorgou, G., F. Mealli, and C. M. Zigler (2019). Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics 75*, 778–787.

Pitt, M. K. and N. Shephard (1999). Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association 94*, 590–599.

Plagborg-Møller, M. (2019). Bayesian inference for structural impulse response functions. *Quantitative Economics 10*(1), 145–184.

Plagborg-Møller, M. and C. K. Wolf (2020). Instrumental variable identification of dynamic variance decompositions.

Ramey, V. A. (2016). Macroeconomics shocks and their propagation. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2A, pp. 71–162. Amsterdam, The Netherlands: North Holland.

Ramey, V. A. and S. Zubairy (2018). Government spending multipliers in good times and in bad: Evidence from US historical data. *Journal of Political Economy 126*, 850–901.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with sustained exposure periods: Application to control of the healthy worker survivor effect. *Mathematical Modelling 7*, 1393–1512.

Romer, C. D. and D. H. Romer (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review 94*, 1055–1084.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association 75*, 591–593.

Sargent, T. J. (1981). Interpreting economic time series. *Journal of Political Economy 89*, 213–248.

Sims, C. A. (1972). Money, income and causality. *American Economic Review 62*, 540–552.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica 48*, 1–48.

Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica 5*, 105–146.

Stock, J. H. and M. W. Watson (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2A, pp. 415–525.

Stock, J. H. and M. W. Watson (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Economic Journal 128*, 917–948.

Tenreyro, S. and G. Thwaites (2016). Pushing on a string: US monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics 8*(4), 43–74.

White, H. and P. Kennedy (2009). Retrospective estimation of causal effects through time. In J. L. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, pp. 59–87. Oxford University Press.

White, H. and X. Lu (2010). Granger causality and dynamic structural systems. *Journal of Financial Econometrics 8*, 193–243.

Wiener, N. (1956). The theory of prediction. In E. F. Beckenbeck (Ed.), *Modern Mathematics*, pp. 165–190. New York, USA: McGraw-Hill.

Wu, X., K. R. Weinberger, G. A. Wellenius, F. Dominici, and D. Braun (2021). Assessing the causal effects of a stochastic intervention in time series data: Are heat alerts effective in preventing deaths and hospitalizations? Unpublished paper: Department of Biostatistics, Harvard T.H. Chan School of Public Health.

A Proofs of Results for Assignments and Outputs

A.1 Proof of Theorem 1

To prove this result, we begin by rewriting $\mathbb{E}[Y_{j,t+h}1\{W_{k,t} = w_k\}]$. Notice that

$$\mathbb{E}[Y_{j,t+h}1\{W_{k,t} = w_k\}]$$

$$= \mathbb{E}[Y_{j,t+h}(W_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h})1\{W_{k,t} = w_k\}]$$

$$= \mathbb{E}[Y_{j,t+h}(W_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h})]\mathbb{E}[1\{W_{k,t} = w_k\}]$$

$$+Cov\left(Y_{j,t+h}(W_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}), 1\{W_{k,t} = w_k\}\right).$$

Therefore, it immediately follows that

$$\mathbb{E}[Y_{j,t+h} \mid W_{k,t} = w_k] = \mathbb{E}[Y_{j,t+h}(W_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h})]$$

$$+\frac{Cov\left(Y_{j,t+h}(W_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}), 1\{W_{k,t} = w_k\}\right)}{\mathbb{E}[1\{W_{k,t} = w_k\}]}.$$

The result is then immediate by (i) applying the same calculation to $\mathbb{E}[Y_{j,t+h}1\{W_{k,t} = w'_k\}]$, (ii) taking the difference, and (iii) applying the definition of $Y_{j,t+h}(w_k)$. $\square$

A.2 Proof of Theorem 3

The style proof extends Angrist et al. (2000) in their analysis of the Wald estimand in a cross-sectional setting. Begin by writing $Y_{t+h} = Y_{t+h}(W_{k,t})$ as

$$\begin{aligned} Y_{t+h} &= Y_{t+h}(\underline{w}_k) + \int_{\underline{w}_k}^{W_{k,t}} \frac{\partial Y_{t+h}(\tilde{w}_k)}{d\tilde{w}_k} \partial\tilde{w}_k \\ &= Y_{t+h}(\underline{w}_k) + \int_{\underline{w}_k}^{\overline{w}_k} \frac{\partial Y_{t+h}(\tilde{w}_k)}{\partial\tilde{w}_k} 1\{\tilde{w}_k \leq W_{k,t}\} d\tilde{w}_k \end{aligned}$$

by the fundamental theorem of calculus. Then, it follows that

$$
\begin{aligned}
&Cov(Y_{t+h}, W_{k,t}) = \mathbb{E}[Y_{t+h}(W_{k,t} - \mathbb{E}[W_{k,t}])] \\
&\overset{(1)}{=} \mathbb{E}[(Y_{t+h} - Y_{t+h}(\underline{w}_k))(W_{k,t} - \mathbb{E}[W_{k,t}])] \\
&= \mathbb{E}\left[\left(\int_{\underline{w}_k}^{\overline{w}_k} \frac{\partial Y_{t+h}(\tilde{w}_k)}{\partial \tilde{w}_k} 1\{\tilde{w}_k \leq W_{k,t}\} d\tilde{w}_k\right)(W_{k,t} - \mathbb{E}[W_{k,t}])\right] \\
&= \int_{\underline{w}_k}^{\overline{w}_k} \mathbb{E}\left[\frac{\partial Y_{t+h}(\tilde{w}_k)}{\partial \tilde{w}_k} 1\{\tilde{w}_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t}])\right] d\tilde{w}_k \\
&\overset{(2)}{=} \int_{\underline{w}_k}^{\overline{w}_k} \mathbb{E}\left[\frac{\partial Y_{t+h}(\tilde{w}_k)}{\partial \tilde{w}_k}\right] \mathbb{E}\left[1\{\tilde{w}_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t}])\right] d\tilde{w}_k
\end{aligned}
$$

where (1) and (2) follow since $W_{k,t} \perp\!\!\!\perp \{Y_{t+h}(w_k)\colon w_k \in \mathcal{W}_k\}$. Interchanging the order of the derivation and the expectation delivers the result. Analogously,

$$
W_{k,t} = \underline{w}_k + \int_{\underline{w}_k}^{W_{k,t}} d\tilde{w}_k = \underline{w}_k + \int_{\underline{w}_k}^{\overline{w}_k} 1\{\tilde{w}_k \leq W_{k,t}\} d\tilde{w}_k,
$$

so

$$
Var(W_{k,t}) = \mathbb{E}[(W_{k,t} - \underline{w}_k)(W_{k,t} - \mathbb{E}[W_{k,t}])] = \int_{\underline{w}_k}^{\overline{w}_k} \mathbb{E}\left[1\{\tilde{w}_k \leq W_{k,t}\}(W_{k,t} - \mathbb{E}[W_{k,t}])\right] d\tilde{w}_k.
$$

The result then follows immediately. To see that the resulting weights are non-negative, observe that for $\tilde{w}_k \in [\underline{w}_k, \overline{w}_k]$

$$
\begin{aligned}
&\mathbb{E}\left[1\{W_{k,t} \geq \tilde{w}_k\}(W_{k,t} - \mathbb{E}[W_{k,t}])\right] \\
&= \mathbb{E}\left[1\{W_{k,t} \geq \tilde{w}_k\} W_{k,t}\right] - \mathbb{E}[1\{W_{k,t} \geq \tilde{w}_k\}]\mathbb{E}[W_{k,t}] \\
&= \left(\mathbb{E}\left[W_{k,t} \mid W_{k,t} \geq \tilde{w}_k\right] - \mathbb{E}[W_{k,t}]\right)\mathbb{P}(W_{k,t} \geq \tilde{w}_k) \geq 0
\end{aligned}
$$

since $\mathbb{E}\left[W_{k,t} \mid W_{k,t} \geq \tilde{w}_k\right] \geq \mathbb{E}[W_{k,t}]$ for $\tilde{w}_k \in [\underline{w}_k, \overline{w}_k]$. $\square$

A.3 Proof of Theorem 4

The proof is analogous to the proof of Theorem 1. We start by rewriting $\mathbb{E}[Y_{j,t+h} 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}]$, noticing that

$$
\begin{aligned}
&\mathbb{E}[Y_{j,t+h} 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}] \\
&= \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}) 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}] \\
&= \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}) \mid \mathcal{F}_{t-1}]\mathbb{E}[1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}]
\end{aligned}
$$

$$+Cov\left(Y_{j,t+h}(w^{obs}_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}\right).$$

Therefore, we have shown that

$$\mathbb{E}[Y_{j,t+h} \mid W_{k,t} = w_k, \mathcal{F}_{t-1}] = \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}) \mid \mathcal{F}_{t-1}]$$

$$+\frac{Cov\left(Y_{j,t+h}(w^{obs}_{1:t-1}, w_k, W_{-k,t}, W_{t+1:t+h}), 1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}\right)}{\mathbb{E}[1\{W_{k,t} = w_k\} \mid \mathcal{F}_{t-1}]}.$$

The result follows by (i) applying the same calculation to $\mathbb{E}[Y_{j,t+h}1\{W_{k,t} = w'_k\} \mid \mathcal{F}_{t-1}]$, (ii) taking the difference, and (iii) applying the definition of the potential outcome $Y_{j,t+h}(w_k)$. $\square$

B Proofs of Results for Assignments, Instruments and Outputs

B.1 Proof of Theorem 6

To prove this result, we first observe that

$$\mathbb{E}[Y_{j,t+h} \mid Z_t = z] = \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, W_{k,t}(z), W_{-k,t}, W_{t+1:t+h}) \mid Z_t = z]$$

$$= \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, W_{k,t}(z), W_{-k,t}, W_{t+1:t+h})]$$

by (iii). Therefore,

$$\mathbb{E}[Y_{j,t+h} \mid Z_t = z] - \mathbb{E}[Y_{j,t+h} \mid Z_t = z']$$

$$= \mathbb{E}[Y_{j,t+h}(w^{obs}_{1:t-1}, W_{k,t}(z), W_{-k,t}, W_{t+1:t+h}) - Y_{j,t+h}(w^{obs}_{1:t-1}, W_{k,t}(z'), W_{-k,t}, W_{t+1:t+h})].$$

Next, we can further rewrite this previous expression as

$$\mathbb{E}[\int_{W_{j,t}(z')}^{W_{j,t}(z)} \frac{\partial Y_{j,t+h}(w_k)}{\partial w_k} dw_k] = \mathbb{E}[\int_{\mathcal{W}} \frac{\partial Y_{j,t+h}(w_k)}{\partial w_k} 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\} dw_k]$$

where we used the definition $Y_{j,t+h}(w_k) := Y_{j,t+h}(W_{1:t-1}, w_{k,t}, W_{-k,t}, W_{t+1:t+h})$. Finally, assuming that we can exchange the order of integration and expectation, we arrive at

$$\int_{\mathcal{W}} \mathbb{E}[\frac{\partial Y_{j,t+h}(w_k)}{\partial w_k} 1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}] dw_k$$
$$= \int_{\mathcal{W}} \mathbb{E}[\frac{\partial Y_{j,t+h}(w_k)}{\partial w_k}, W_{k,t}(0) \leq w_k \leq W_{k,t}(1)] \mathbb{E}[1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}] dw_k.$$

We may apply the same argument to the denominator (again assuming that we can exchange the order of integration and expectation) to arrive at

$$\mathbb{E}[W_{k,t} \mid Z_t = z] - \mathbb{E}[W_{k,t} \mid Z_t = z'] =$$

$$\mathbb{E}[W_{k,t}(z) - W_{k,t}(z')] = \int_{\mathcal{W}} \mathbb{E}[1\{W_{k,t}(z') \leq w_k \leq W_{k,t}(z)\}].$$

Taking the ratio then delivers the desired result. □

B.2 Proof of Theorem 8

The proof is the same as the Proof of Theorem 6, except we must now condition on $\mathcal{F}_{t-1}$ throughout. □

C Proofs of Results for Outputs

C.1 Proof of Theorem 10

Then, if the subsequent moments exist, we have that

$$\begin{aligned}
\mathrm{E}[Y_{j,t+h}|(Y_{k,t} = y_k), \mathcal{F}^Y_{t-1}] &= \mathrm{E}[Y_{j,t+h}(W_{1:t})|(Y_{k,t} = y_k), \mathcal{F}^Y_{t-1}], \quad \text{Assumption (i)} \\
&= \mathrm{E}[\mathrm{E}[Y_{j,t+h}(W_{1:t+h})|(Y_{k,t} = y_k), W_{1:t}, \mathcal{F}^Y_{t-1}]|Y_{k,t} = y_k, \mathcal{F}^Y_{t-1}], \quad \text{Adam's law} \\
&= \mathrm{E}[\mathrm{E}[Y_{j,t+h}(W_{1:t+h})|W_{1:t})]|(Y_{k,t} = y_k), \mathcal{F}^Y_{t-1}], \quad \text{Assumption (i)} \\
&= \mathrm{E}[\psi_{j,t+h}(W_{1:t})|(Y_{k,t} = y_k), \mathcal{F}^Y_{t-1}], \quad \text{Assumption (ii)}
\end{aligned}$$

the last line holds as the future assignments are not informed by the historical ones. Applying this result twice gives the first result. □