Econometric Analysis of Potential Outcomes Time Series: Instruments, Shocks, Linearity and the Causal Response Function

Ashesh Rambachan[1] Neil Shephard[2]

[1]Department of Economics, Harvard University

[2]Department of Economics and Department of Statistics, Harvard University

July 13, 2021

Core question

Stochastic process $(W_t, X_t, Y_t)_{t\geq 1}$, where

- $W_t \in \{0, 1\}$, assignment
- X_t, background
- Y_t, outcome
- $(\mathcal{F}_t)_{t\geq 1}$ natural filtration of data: $\left(w_t^{obs}, x_t^{obs}, y_t^{obs}\right)_{t\geq 1}$.

Build: potential outcome system. Necessary and sufficient conditions for direct prediction

$$\mathrm{E}[Y_t|\mathcal{F}_{t-s-1}, X_{t-s}, (W_{t-s} = 1)] - \mathrm{E}[Y_t|\mathcal{F}_{t-s-1}, X_{t-s}, (W_{t-s} = 0)], \quad s \geq 0,$$

to be causal:

$$\mathrm{E}[\{Y_t(w_{1:t-s-1}^{obs}, 1, W_{t-s+1:t}) - Y_t(w_{1:t-s-1}^{obs}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}, X_{t-s}],$$

Quantifies nonparametric question: when does direct prediction equal causality?

Filtered IRF:

$$\mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\} | \mathcal{F}_{t-s-1}, X_{t-s}],$$

- $Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t})$ time-t potential outcome for past $w^{obs}_{1:t-s-1}$, then $w_{t-s} = 1$, other assignments $W_{t-s+1:t}$.
- Time-t potential outcomes $\{Y_t(w^{obs}_{1:t-s-1}, w_{t-s:t}) : w_{t-s:t} \in \{0,1\}^{s+1}\}$ are unobserved.
- Expectation of $Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) | \mathcal{F}_{t-s-1}, X_{t-s}$: averages over random time-t potential outcomes and random future assignments given past data.

Generalized impulse response function: Koop et al. (1996), Gallant et al. (1992), Gourieroux and Jasiak (2005).

Integrating out the filtration and background variables, then the predictive IRF

$$\mathrm{E}[Y_t|\,(W_{t-s}=1)]-\mathrm{E}[Y_t|\,(W_{t-s}=0)],\quad s\geq 0,$$

equals the causal IRF

$$\mathrm{E}[\{Y_t(W_{1:t-s-1},1,W_{t-s+1:t})-Y_t(W_{1:t-s-1},0,W_{t-s+1:t})\}].$$

Impulse response function: Sims (1980).

LATE-type result for time series

Give conditions for a Wald-like ratio of filtered IRFs

$$\frac{\mathrm{E}[Y_t | (Z_{t-s} = 1), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[Y_t | (Z_{t-s} = 0), \mathcal{F}_{t-s-1}, X_{t-s}]}{\mathrm{E}[W_{t-s} | (Z_{t-s} = 1), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[W_{t-s} | (Z_{t-s} = 0), \mathcal{F}_{t-s-1}, X_{t-s}]},$$

to equal the filtered LATE, where Z_{t-s} is an instrument:

$$\begin{aligned}\mathrm{E}[\{ Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\} \\ |\mathcal{F}_{t-s-1}, X_{t-s}, \{W_{t-s}(1) - W_{t-s}(0) = 1\}],\end{aligned}$$

Nonparametric causal interpretation of instrumented time series.

IV: Angrist and Krueger (2001), Imbens (2014).

IV and measurement error: Durbin (1954), Sargan (1958), Arellano (2002)

LATE in cross-sections: Imbens and Angrist (1994), Angrist et al. (1996)

IV for local projection: Stock and Watson (2018)

Potential outcome system

- Define "potential outcome system" a nonparametric framework for time series causality.
- Extends finite sample experimental time series of Bojinov and Shephard (2019) — to population based observation studies using stochastic processes & background variables.
- Built on four foundational Assumptions.

Time series: Angrist and Kuersteiner (2011), Angrist et al. (2018), Bojinov and Shephard (2019), Blackwell and Glynn (2018), Bondersen et al. (2015).

Natural experiments: Romer and Romer (1989), Cochrane and Piazzesi (2002) and Bernanke and Kuttner (2005), Jorda (2005), Plagborg-Moller and Wolf (2021), Olea and Plagborg-Moller (2021).

Model based strategies: Sims (1980), Kilian and Lutkepohl (2017).

Alternative nonparametric strategies: Granger (1969) (see Lechner (2011)).

Reviews: Stock and Watson (2018), Ramey (2016), Nakamura and Steinsson (2018), Kursteiner (2010).

Panel: Robins (1986), Robins et al. (1999), Hernan and Robins (2021), Boruvka et al. (2017), Bojinov et al. (2021).

Potential outcome system: Assumption 1

An assignment process $\{W_t\}_{t\geq 1}$ has for each t, $W_t \in \mathcal{W}_W$. The joint background and potential output process is, for any deterministic $\{w_t\}_{t\geq 1}$ with $w_t \in \mathcal{W}_W$,

$$\left\{X_t, Y_t(\{w_s\}_{s\geq 1})\right\}_{t\geq 1},$$

where $Y_t(\{w_s\}_{s\geq 1})$ is time-t potential outcome.

- As $\{w_s\}_{s\geq 1}$ changes, the same X_t is produced but a different potential outcome process is selected.
- Often $\mathcal{W}_W = \{0, 1\}$: assignments are univariate and binary.

Potential outcome system: Assumption 2

For each t

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

almost surely, for all deterministic $\{w_t\}_{t\geq 1}$ and $\{w'_s\}_{s\geq t+1}$, where $w_t \in \mathcal{W}_W$ and $w'_t \in \mathcal{W}_W$.

- Stochastic process analogue of non-interference (Cox (1958), Rubin (1980)), extending Bojinov and Shephard (2019).
- For convenience, we will drop references to future treatments and write

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

Potential outcome system: Assumption 3

The output from this system is assumed to be

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

where $Y_t = Y_t(W_{1:t})$ is called the time-t outcome.

Example (Potential vector moving average (PVMA))

$$\begin{pmatrix} Y_t(w_{1:t}) \\ X_t \end{pmatrix} = \sum_{j=0}^{t-1} \Theta_j \begin{pmatrix} \varepsilon_{Y,t-j}(w_{t-j}) \\ \varepsilon_{X,t-j} \end{pmatrix} + \sum_{j=t}^{\infty} \Theta_j \varepsilon_{t-j}, \quad \Theta_j = \begin{pmatrix} \theta_{Y,Y,j} & \theta_{Y,X,j} \\ 0 & \theta_{X,X,j} \end{pmatrix},$$

for all $w_{1:t} \in \mathcal{W}_W^t$, while

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \sum_{j=0}^{t-1} \Theta_j \begin{pmatrix} \varepsilon_{Y,t-j}(W_{t-j}) \\ \varepsilon_{X,t-j} \end{pmatrix} + \sum_{j=t}^{\infty} \Theta_j \varepsilon_{t-j}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{Y,t} \\ \varepsilon_{X,t} \end{pmatrix}.$$

Here $(\varepsilon_t)_{t\leq 0}$ and $(\varepsilon_{X,t})_{t\geq 1}$ do not vary with $w_{1:t}$.

- No assumptions on the prob properties of $(\varepsilon_t)_{t\leq 0}$, $(\varepsilon_{X,t})_{t\geq 1}$ and $\{\varepsilon_{Y,t}(w_t) : w_t \in \mathcal{W}_W\}_{t\geq 1}$. Upper triangular $(\Theta_j)_{j\geq 0}$ essential to stop the assignment hitting background variable. Relates to the "short-run timing" restriction of Sims (1980).

Example

(a) In the binary assignment PVMA case, then

$$\varepsilon_{Y,t}(w_t) = \varepsilon_{Y,t}(0) + \tau_{\varepsilon,t} w_t, \quad \tau_{\varepsilon,t} = \varepsilon_{Y,t}(1) - \varepsilon_{Y,t}(0),$$

implying

$$\begin{pmatrix} Y_t(w_{1:t}) \\ X_t \end{pmatrix} = \sum_{j=0}^{t-1} \begin{pmatrix} \theta_{Y,Y,j} \tau_{\varepsilon,t} w_{t-j} \\ 0 \end{pmatrix} + \sum_{j=0}^{t-1} \Theta_j \begin{pmatrix} \varepsilon_{Y,t-j}(0) \\ \varepsilon_{X,t-j} \end{pmatrix} + \sum_{j=t}^{\infty} \Theta_j \varepsilon_{t-j}.$$

(b) In the linear assignment PVMA case, where $\varepsilon_{Y,t}(w_t) = \varepsilon_{Y,t}(0) + \tau_\varepsilon w_t$.

(c) In the threshold assignment PVMA case, where $\varepsilon_{Y,t}(w_t) = \varepsilon_{Y,t}(0) + \tau_\varepsilon 1_{w_t < 0} w_t$.

Potential outcome system: Assumption 4

Assignment is sequentially probabilistic, that is

$$0 < P(W_t = w_t | \mathcal{F}_{t-1}, X_t) < 1$$

for all $w_t \in \mathcal{W}_W$.

Potential outcome system in words

1. Background, assignment, potential outcomes. Assignment select potential outcome.
2. Outcomes not functions of future assignments. Non-interference.
3. Observable data: assignment, background and outcome.
4. Probabilistic assignment.

Definition

Any $\{W_t, X_t, \{Y_t(w_{1:t}) : w_{1:t} \in \mathcal{W}_W^t\}\}_{t \geq 1}$ satisfying Assumptions 1-4 is called a "potential outcome system."

Filtered treatment effect

Definition

Assume a potential outcome system and $\mathcal{W}_W = \{0, 1\}$. Define the filtered IRF

$$\tau_{X,t}(s) = \mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}, X_{t-s}].$$

Also write the time-t, lag-s filtered treatment effect as

$$\tau_t(s) = \mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}].$$

Filtering is the sequential estimation of a 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 filter on hidden discrete Markov models (Baum and Petrie (1966), Hamilton (1989)).

$$\tau_{X,t}(s) = \mathrm{E}[\{Y_t(w_{1:t-s-1}^{obs}, 1, W_{t-s+1:t}) - Y_t(w_{1:t-s-1}^{obs}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}, X_{t-s}].$$

To illustrate the filtered treatment effect, return to the PVMA.

Theorem

Assume a PVMA with $w_{t-s} \in \{0, 1\}$ *and* $\mathrm{E}[|\varepsilon_{Y,t-s}(1) - \varepsilon_{Y,t-s}(0)| \,|\, (\mathcal{F}_{t-s-1}, X_{t-s})] < \infty$*. Then*

$$\tau_{X,t}(s) = \theta_{Y,Y,s}\mathrm{E}[\{\varepsilon_{Y,t-s}(1) - \varepsilon_{Y,t-s}(0)\} \,|\, (\mathcal{F}_{t-s-1}, X_{t-s})].$$

Notice that even if $\varepsilon_{Y,t-s}(W_{t-s})$ is a martingale difference, this does not imply that $\varepsilon_{Y,t-s}(1)$ and $\varepsilon_{Y,t-s}(0)$ are.

Focusing on observables

$$\tau_{X,t}(s) = \mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}, X_{t-s}].$$

Potential outcomes are unobserved. Cleave $\tau_{X,t}(s)$ into a term involving only observables and a selection effect which involves unobservables.

$$\tau_{X,t}(s) = \mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\}|\mathcal{F}_{t-s-1}, X_{t-s}].$$

Theorem

For the potential outcome system, $\left\{W_t, X_t, \left\{Y_t(w_{1:t}), w_{1:t} \in \{0,1\}^t\right\}\right\}_{t\geq 1}$, *assume additionally that* $\mathrm{E}[|Y_t|\,|\mathcal{F}_{t-s-1}, X_{t-s}] < \infty$. *Then, at time* t*, lag* s*,*

$$\tau_{X,t}(s) = CIRF_t(s, X_{t-s}) - \Delta SEL_t(s, X_{t-s}),$$

where

$$CIRF_t(s, X_{t-s}) = \mathrm{E}[Y_t|\mathcal{F}_{t-s-1}, (W_{t-s} = 1), X_{t-s}] - \mathrm{E}[Y_t|\mathcal{F}_{t-s-1}, (W_{t-s} = 0), X_{t-s}],$$

$$\Delta SEL_t(s, X_{t-s}) = \frac{\mathrm{Cov}(W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s})}{\mathrm{E}[W_{t-s}|\mathcal{F}_{t-s-1}, X_{t-s}]} - \frac{\mathrm{Cov}(1 - W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s})}{\mathrm{E}[(1 - W_{t-s})\,|\mathcal{F}_{t-s-1}, X_{t-s}]}.$$

Thus direct prediction is causal iff $\Delta SEL_t(s, X_{t-s})$ equals zero.

Occurs iff this condition holds:

$$\begin{aligned}
& \frac{\mathrm{Cov}(W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s})}{\mathrm{E}[W_{t-s}|\mathcal{F}_{t-s-1}, X_{t-s}]} \\
= \quad & \frac{\mathrm{Cov}(1-W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s})}{\mathrm{E}[(1-W_{t-s})\,|\mathcal{F}_{t-s-1}, X_{t-s}]}
\end{aligned}$$

Unconfounded process, for all $w_{t-s:t} \in \mathcal{W}_W^{s+1}$

$$(W_{t-s} \perp\!\!\!\perp Y_t(w^{obs}_{1:t-s-1}, w_{t-s:t}))|\mathcal{F}_{t-s-1}, X_{t-s}]$$

is not enough for this to hold.

Theorem

For the PVMA

$$
\begin{aligned}
&\mathrm{Cov}(W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, w_{t-s}, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s}) \\
&= \theta_{Y,Y,s}\mathrm{Cov}\left(W_{t-s}, \varepsilon_{Y,t-s}(w_{t-s})|\mathcal{F}_{t-s-1}, X_{t-s}\right) \\
&+ \sum_{j=0}^{s-1} (\theta_{Y,Y,j}, \theta_{Y,X,j}) \begin{pmatrix} \mathrm{Cov}\left(W_{t-s}, \varepsilon_{Y,t-j}(W_{t-j})|\mathcal{F}_{t-s-1}, X_{t-s}\right) \\ \mathrm{Cov}\left(W_{t-s}, \varepsilon_{X,t-j}|\mathcal{F}_{t-s-1}, X_{t-s}\right) \end{pmatrix}.
\end{aligned}
$$

Recall, in linear case $\varepsilon_{Y,t-s}(w_{t-s}) = \varepsilon_{Y,t-s}(0) + \tau w_{t-s}$. So former is $\mathrm{Cov}\left(W_{t-s}, \varepsilon_{Y,t-s}(0)|\mathcal{F}_{t-s-1}, X_{t-s}\right)$.

- The foundation question for time series causal studies is: under what conditions is the $\Delta SEL_t(s, X_{t-s})$ term zero?
- The selection effect $\Delta SEL_t(s, X_{t-s})$ involve the unobserved potential outcomes and future treatments. Difficult to make much progress with.
- A sufficient condition for it to be zero is the following:

Assume a potential outcome system and

$$\mathrm{Cov}(W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, w_{t-s}, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s}) = 0, \quad \text{for all} \quad w_t \in \mathcal{W}_W. \quad (1)$$

What are sufficient conditions for (1) to hold?

For the PVMA:

$$\mathrm{Cov}\left(\{W_{t-s}, \varepsilon_{Y,t-s}(w_{t-s})\}\,|\mathcal{F}_{t-s-1}, X_{t-s}\right) = 0, \quad \text{for all} \quad w_{t-s} \in \mathcal{W}_W,$$

and that, writing $\varepsilon_{Y,t-j} = \varepsilon_{Y,t-j}(W_{t-j})$, for $j = 0, ..., s-1$

$$\mathrm{Cov}\left(\{W_{t-s}, \varepsilon_{Y,t-j}\}\,|\mathcal{F}_{t-s-1}, X_{t-s}\right) = 0, \quad \mathrm{Cov}\left(\{W_{t-s}, \varepsilon_{X,t-j}\}\,|\mathcal{F}_{t-s-1}, X_{t-s}\right).$$

The former is classic-like: assignment W_{t-s} is a contemporaneously unconfounded.

Latter will hold if, e.g.,

$$\mathrm{E}[\{\varepsilon_{Y,t-j}, \varepsilon_{X,t-j}\}\,|(\varepsilon_{Y,t-s+1:t-j-1}, \varepsilon_{X,t-s+1:t-j-1}, W_{t-s}, \mathcal{F}_{t-s-1}, X_{t-s})] = 0,$$

"shock" type condition for the "unanticipated structural disturbances" intimated by Stock and Watson (2018).

Shock conditions frequently appear time series studies, but they seem absent in the corresponding longtitudinal causal literature.

New direction

Now focus on case where crucial condition:

$$\mathrm{Cov}(W_{t-s}, Y_t(w^{obs}_{1:t-s-1}, w_{t-s}, W_{t-s+1:t})|\mathcal{F}_{t-s-1}, X_{t-s}) = 0, \quad \text{for all} \quad w_t \in \mathcal{W}_W. \qquad (2)$$

fails to hold.

Instrumented potential outcome system

Assume $W_t \in \mathcal{W}_W$, $Z_t \in \mathcal{W}_Z$, write $V_t = (W_t, Z_t)$ and that

$$\left\{V_t, X_t, \left\{Y_t(v_{1:t}) : v_{1:t} \in (\mathcal{W}_W \times \mathcal{W}_Z)^t\right\}\right\}_{t \geq 1}$$

is a "potential outcome system."

(a) Define the time-t "exclusion condition" as

$$Y_t((w_1, z_1), \dots, (w_t, z_t)) = Y_t((w_1, z_1'), \dots, (w_t, z_t'))$$

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 $Y_t = Y_t(W_{1:t})$.

(b) Define the time-t lag-s "sequential relevance condition" as $(W_{t-s} \not\perp\!\!\!\perp Z_{t-s}) \,| \mathcal{F}_{t-s-1}, X_{t-s}$.

Definition

If (a)+(b) hold, then this is called an "instrumented potential outcome system", written as

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

The exclusion condition means that the instrument process has no direct impact on the selection of the potential outcome.

The sequential relevance condition means that the assignment W_t and instrument Z_t are not conditionally independent.

Local filtered treatment effects

To draw a time series link to the influential literature on local average treatment effects (LATE) of Imbens and Angrist (1994) and Angrist et al. (1996), first assume:

$Z_t \in \{0, 1\}$ and that

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

is an instrumented potential outcome system.

Need an instrument and some additional assumptions.

(a) Assume $W_t(z_{1:t}) = W_t(z'_{1:t-1}, z_t)$, for all $z_{1:t}$ and $z'_{1:t-1}$. Write potential assignments as

$$\{W_t(0), W_t(1)\}$$

and $W_t = W_t(Z_t)$.

the instrument Z_{t-s}, selects between the potential assignments $\{W_{t-s}(0), W_{t-s}(1)\}$. It does not rule out that lagged instruments can impact the law of the potential assignments.

(b) Assume $W_t(1) \geq W_t(0)$, with probability one: assignment is monotonic.

monotonicity, was introduced by Angrist et al. (1996) in their LATE work on heterogeneous treatment effects and instrumental variables for cross-sectional data.

(c) Assume, for all z_{t-s},

$$\mathrm{Cov}(Z_{t-s}, Y_t(w^{obs}_{1:t-s-1}, W_{t-s}(z_{t-s}), W_{t-s+1:t}) | \mathcal{F}_{t-s-1}, X_{t-s}) = 0. \quad (3)$$

(d) Assume $[Z_{t-s} \perp\!\!\!\perp \{W_{t-s}(0), W_{t-s}(1)\}] \, | \mathcal{F}_{t-s-1}, X_{t-s}) = 0$.

(e) Assume

$$\mathrm{E}[W_{t-s} | \left(Z_{t-s} = 1\right), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[W_{t-s} | \left(Z_{t-s} = 0\right), \mathcal{F}_{t-s-1}, X_{t-s}] > 0.$$

Given the conditions (a)-(e) the Wald-like ratio of filtered IRFs

$$IV_t(s) = \frac{\mathrm{E}[Y_t | (Z_{t-s} = 1), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[Y_t | (Z_{t-s} = 0), \mathcal{F}_{t-s-1}, X_{t-s}]}{\mathrm{E}[W_{t-s} | (Z_{t-s} = 1), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[W_{t-s} | (Z_{t-s} = 0), \mathcal{F}_{t-s-1}, X_{t-s}]},$$

equals the local filtered treatment effect (filtered LATE)

$$\begin{gathered}\mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, 0, W_{t-s+1:t})\} \\ |\mathcal{F}_{t-s-1}, X_{t-s}, \{W_{t-s}(1) - W_{t-s}(0) = 1\}],\end{gathered}$$

and

$$\frac{\mathrm{E}[\{Y_t(w^{obs}_{1:t-s-1}, W_{t-s}(1), W_{t-s+1:t}) - Y_t(w^{obs}_{1:t-s-1}, W_{t-s}(0), W_{t-s+1:t})\} |\mathcal{F}_{t-s-1}, X_{t-s}]}{\mathrm{E}[\{W_{t-s}(1) - W_{t-s}(0)\} |\mathcal{F}_{t-s-1}, X_{t-s}\}}.$$

Causal meaning to IV

Given the conditions (a)-(e) the lag-s

$$IV(s) = \frac{\mathrm{E}[Y_t | (Z_{t-s} = 1)] - \mathrm{E}[Y_t | (Z_{t-s} = 0)]}{\mathrm{E}[W_{t-s} | (Z_{t-s} = 1)] - \mathrm{E}[W_{t-s} | (Z_{t-s} = 0)]},$$

equals

$$\mathrm{E}[\psi(\mathcal{F}_{t-s-1}, X_{t-s})\{Y_t(W_{1:t-s-1}, 1, W_{t-s+1:t}) - Y_t(W_{1:t-s-1}, 0, W_{t-s+1:t})\}],$$

where

$$\psi(\mathcal{F}_{t-s-1}, X_{t-s}) = \frac{\mathrm{E}[W_{t-s} | (Z_{t-s} = 1), \mathcal{F}_{t-s-1}, X_{t-s}] - \mathrm{E}[W_{t-s} | (Z_{t-s} = 0), \mathcal{F}_{t-s-1}, X_{t-s}]}{\mathrm{E}[W_{t-s} | (Z_{t-s} = 1)] - \mathrm{E}[W_{t-s} | (Z_{t-s} = 0)]}.$$

weights are positive, unit expectation, proportional to the probability of being a time $t-s$ compiler given $\mathcal{F}_{t-s-1}, X_{t-s}$.

Summary

- Define potential outcome system
- Define filtered treatment effect
- Necessary and sufficient conditions for Predictive IRF equals Causal IRF.
- Predictive IRF need new assumptions beyond usual sequential unconfoundedness.
- Extend to instrumented potential outcome system
- Delivers a LATE intepretation to time series IV.

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., Ò. Jordà, and G. M. Kuersteiner (2018). Semiparametric estimates of monetary policy effects: string theory revisited. *Journal of Business & Economic Statistics 36*, 371–387.

Angrist, J. D. and A. B. Krueger (2001). Instrumental variables and the search for identification: from supply and demand to natural experiments. *Journal of Economic Perspectives 15*, 69–85.

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.

Arellano, M. (2002). Sargan's instrumental variables estimation and the generalized method of moments. *Journal of Business and Economic Statistics 20*, 450–459.

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.

Bernanke, B. and K. N. Kuttner (2005). What explains the stock market's reaction to the Federal Reserve policy? *The Journal of Finance 40*, 1221–1257.

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. *Quantitiative 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.

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

Boruvka, A., D. Almirall, K. Witkiwitz, and S. A. Murphy (2017). Assessing time-varying causal effect moderation in mobile health. *Journal of the American Statistical Association 113*, 1112–1121.

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

Cochrane, J. H. and M. Piazzesi (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. (1954). Errors in variables. *Review of the Institute of International Statistics 22*, 23–54.

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

Gallant, A. R., P. E. Rossi, and G. Tauchen (1992). Stock prices and volume. *Review of Financial Studies 5*, 199–242.

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'Économie 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.

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

Imbens, G. W. (2014). Instrumental variables: an econometrician's perspective. *Statisticial Science 29*, 323–358.

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

Jorda, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*, 161—-182.

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: Cambridge University Press.

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

Kursteiner, G. (2010). Granger-Sims causality. In S. N. Durlauf and L. Blume (Eds.), *Macroeconomics and Time Series Analysis*, pp. 119–134. Palgrave Macmillian.

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

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

Olea, J. L. M. and M. Plagborg-Moller (2021). Local projection inference is simpler and more robust than you think. *Econometrica 89*. Forthcoming.

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

Plagborg-Moller, M. and C. K. Wolf (2021). Local projections and VARs estimate the same impulse reponse functions. *Econometrica 89*, 955–980.

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

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.

Robins, J. M., S. Greenland, and F.-C. Hu (1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of the American Statistical Association 94*, 687–700.

Romer, C. and D. H. Romer (1989). Does monetary policy matter? A new test in the spirit of Friedman and Schwartz. In O. J. Blanchard and S. Fischer (Eds.), *NBER Macroeconomics Annual 1989*, pp. 121–170. Cambridge: MIT Press.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association 75*(371), 591–593.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica 26*, 393–415.

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

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